



UAV-Assisted NOMA Network Power Allocation Under Offshore Multi-energy Complementary Power Generation System

Wenyi Shao, Fei Lin^(✉), Ren Xu, Fanping Meng, and Haoran Wang

School of Information and Automation, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250300, China

linfei@qlu.edu.cn

Abstract. To solve the shortage of non-renewable energy sources, the development and utilization of abundant renewable energy sources at sea are gradually attracting attention. For the acquisition and analysis of offshore energy sources, we propose a new offloading framework for grid communication under offshore multi-energy power generation systems. This framework can enhance the performance of offshore communication and provide a good basis for command transmission of multi-energy complementary power generation systems. In this paper, we consider network offloading with the help of unmanned aerial vehicles (UAVs), while adopting non-orthogonal multiple access (NOMA) techniques on each UAV. The system hardware loss and incomplete successive interference cancellation (SIC) are jointly optimized for UAV trajectory and power allocation to minimize the system energy loss. To solve the non-convex problem, we use a two-step Deep Reinforcement Learning (DRL) based algorithm. Numerical results are based on the number of iterations and the variation of the signal-to-noise ratio magnitude to evaluate the effectiveness of the proposed algorithm in the system in terms of system energy consumption, transmission rate, interruption probability, and error rate. This research was funded by the National Natural Science Foundation of China, grant number U2006222 and Natural Science Foundation of Shandong Province, grant number ZR2020MF138.

Keywords: NOMA · Power allocation · Multi-agent · UAV

1 Introduction

After the 1950s, the global energy shortage gradually emerged, which has attracted continuous attention from all over the world [1]. The ocean covers about 71% of the earth's surface area, and the abundant ocean energy is very valuable for research in renewable energy.

In recent years 5G communication technology has been deeply integrated with land-based power IoT. Edge offloading technologies are widely used in it [2]. However, because renewable energy sources such as offshore wind and solar are too dispersed and

far from onshore power equipment and base stations, it is impossible to achieve edge offloading network coverage by relying only on terrestrial base stations [3].

The use of orthogonal multiple access (OMA) for land-based energy communication with UAVs has been widely used, and considering that the number of access terminals will grow immeasurably in the future 6G network with the integration of earth, sea, air, and sky, OMA is difficult to meet the maritime communication scenario with scarce spectrum resources [4]. Non-orthogonal multiple access (NOMA) allows multiple terminals to use the common resource block simultaneously, and the combination with the mobile UAV model can meet the problem of multi-node access to marine networks [5].

Therefore, based on the multi-agent DQN approach, we mainly make the following innovations:

- Considering the communication patterns under the offshore multi-energy complementary power generation system and using a three-path model that fits the characteristics of the sea, we propose a NOMA-based UAV-assisted offloading framework.
- We first invoke the K-means algorithm to classify user clusters. Then we use a multi-intelligent deep Q-network algorithm to reduce the system energy consumption by optimizing the action trajectory and power allocation of UAVs.
- Based on the optimal energy consumption, we analyze the outage probability and transmission rate of the system under different multi-access techniques and evaluate the effectiveness of the proposed framework.

2 System Model

2.1 System

We consider a maritime offload network consisting of U UAVs carrying mini-MEC servers and mobile ship users [6]. We assume that NOMA technology is used to support downlink communication, where each UAV can be associated with K users [7]. The collection of users and drones is denoted by $k \in \mathbb{K}$ and $u \in \mathbb{U}$. The considered UAV-assisted wireless network is shown in Fig. 1.

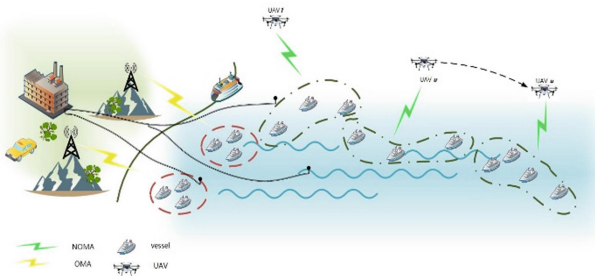


Fig. 1. NOMA-based framework for maritime UAV-assisted cellular networks

Considering the special atmospheric refractive index structure in the marine atmospheric environment easily forms evaporative waveguides, which provide conditions

for long-range communication. Assume that the evaporative duct layer is horizontally homogeneous. When reflected rays from the sea surface exist outside the LoS (Line of Sight) path, it is assumed that near-sweep incidence exists on the sea surface and, eventually, the reflection coefficient of the vertically polarized waves approaches -1 [8]. Using these assumptions, the 3-ray path loss model can be simplified to

$$L_k^u(t) = 10 \lg \left\{ \left(\frac{\lambda}{4\pi d_k^u(t)} \right)^2 [2 \times (1 + \Delta)]^2 \right\}$$

$$\Delta = 2 \times \sin \left(\frac{2\pi h_u h_k}{\lambda d_k^u(t)} \right) \times \sin \left(\frac{2\pi (h_e - h_u)(h_e - h_k)}{\lambda d_k^u(t)} \right) \quad (1)$$

where $d_k^u(t)$ is the distance between UAV and the user in the 3D model, $d_k^u(t) = \sqrt{h_u^2 + (x_u - x_k)^2 + (y_u - y_k)^2}$, h_u denotes the flight altitude of UAV u

It is assumed that the channel gain between the user and the UAV remains constant during each time slot t . Then the channel gain from the UAV u to the user k can be calculated as

$$g_k^u(t) = 10^{-\frac{L_k^u(t)}{10}} \quad (2)$$

2.2 UAV-User Group Association Problems

Assume that in the initial state, the users associated with the UAV are randomly distributed in the maritime cellular network. Let the location state of the UAV and users remain the same in each time slot. It is assumed that each user can be associated with only one UAV within a set of time slots, and the UAV also provides service only to the associated user.

Let $s_k^u(t)$ be the association between the UAV u and the user k . If UAV u is associated with user k , then denote $s_k^u(t) = 1$, otherwise $s_k^u(t) = 0$. Thus there is the following constraint:

$$\sum_{u=1}^U s_k^u(t) = 1, \forall k \in \mathbb{K}, u \in \mathbb{U} \quad (3)$$

Let P denotes the total power of the UAV u . Then the transmit power assigned to user k satisfies the constraint

$$P^u(t) = \sum_{k=1}^K P_k^u(t) s_k^u(t) \quad (4)$$

2.3 NOMA Downlink Communications

We assume that K users are in a cluster. Each user receives and decodes the data transmitted downlink from the associated UAV. Inter-cluster interference is generated when

different UAVs are offloaded to different users and reuse the same channel. If different users reusing the same channel under the same UAV association are considered, intra-cluster interference will occur. NOMA protocol applies SIC receiver at the receiving terminal to implement multi-user detection to eliminate the interference. The receiver first decodes the signal with large channel gain, subtracts the multi-access interference generated by that user's signal from the combined signal, and then judges the remaining users again, and so on, until all the interference is eliminated [9]. Therefore, the signal-to-noise ratio of user k :

$$SINR_k^u(t) = \frac{|g_k^u(t)|^2 P_k^u(t) s_k^u(t)}{I_{in\ k}^u(t) + I_{on\ k}^u(t) + z_k^u} \quad (5)$$

Here, " $I_{on\ k}^u(t) = \sum_{s=1, s \neq u}^U |g_k^s(t)|^2 P^s(t)$ " is the inter-cluster interference to user k from UAVs other than UAV u , where $g_k^s(t)$ denotes the channel gain between UAVs other than UAV u and user k , $s \neq u$. " $I_{in\ k}^u(t) = \sum_{i=k+1}^K |g_i^u(t)|^2 P_i^u(t) s_i^u(t)$ " is the intra-cluster interference generated within the same range served by the UAV. z_k^u is the additive white Gaussian noise (AWGN), $z_k^u \sim CN(0, \sigma^2)$.

With the above reasoning and Shannon's formula, the corresponding data rate can be calculated as

$$R_k^u(t) = \sum_{u=1}^U \sum_{k=1}^K B \log 2(1 + SINR_k^u(t)) \quad (6)$$

The time delay required for the UAV service at moment t is

$$T_k^u(t) = \frac{I_n}{R_k^u(t)} \quad (7)$$

The energy efficiency consumed is

$$E_k^u(t) = P_u(t) \cdot T_k^u(t) \quad (8)$$

2.4 System Outage Probability Analysis

To facilitate the discussion of system mid-range performance, it is necessary to define the system interrupt event first. If any user in the group does not decode successfully, the system is considered to be interrupted. Assuming that the target rate of transmission is r_k , the system interruption probability can be expressed as:

$$P^{out} = 1 - P^s \quad (9)$$

where denotes the probability of successful transmission and has

$$P^s = P^r \{ SINR_k^u \geq SINR_k^r, k \in \mathbb{K} \}$$

$$SINR_k^r = 2^{r_k} - 1 \quad (10)$$

Let $\alpha_k^u(t)$ be the power allocation factor for user k . Then the following relationship exists between user power and transmit power $P_k^u(t) = P^u(t)\alpha_k^u(t)$ and there is a constraint $\sum_{u=1}^U \sum_{k=1}^K \alpha_k^u(t) = 1$.

When $0 < \alpha_k^u(t) < \frac{1}{1+SINR_k'}$, the probability of successful transmission can be expressed as

$$P^s = \Pr \left\{ SINR_k^u(t) \geq \frac{SINR_k'(t)}{\alpha_k^u(t) - \sum_{i=0}^k \alpha_i^u(t) SINR_k'(t)} \right\}, \forall k \in \mathbb{K} \quad (11)$$

3 Question Formula

For the offshore renewable energy generation system, considering the different demands of communication users, we take minimizing the system energy consumption as the optimization objective, including jointly optimizing the trajectory and power allocation of UAVs. We express the positions of UAV u and user k in time slot t by L^u and L^k , respectively, and $L = \{x(t), y(t), h(t), 0 \leq t \leq T\}$. Assume that the UAV and user positions are fixed within each time slot t and that each UAV position is different.

Consequently, the optimization issue may be expressed as

$$\begin{aligned} & \min E(t) \\ & s.t. C1 : L_{\min} \leq L^u(t) \leq L_{\max}, \forall u \in \mathbb{U}, \\ & \quad C2 : L_{\min} \leq L^k(t) \leq L_{\max}, \forall k \in \mathbb{K}, \\ & \quad C3 : L_i^u \neq L_j^u, \quad i, j \in \mathbb{U}, \forall t, \\ & \quad C4 : \sum_{k=1}^{\mathcal{K}} s_k^u(t) P_k^u(t) \leq P^u, \forall k \in \mathbb{K}, u \in \mathbb{U} \end{aligned} \quad (12)$$

C1 ~ C3 denote the location constraints of the UAV and the user in the stereoscopic space. According to C4, the entire amount of power provided to the user cannot be higher than the transmit power.

Since the class of problems in (11) was shown in [6] to be non-convex and NP-hard for this optimization problem, and due to high computational complexity and randomly varying channel conditions, it is difficult to obtain a globally optimal solution in practice. Therefore, this paper invokes an RL-based algorithm that interacts with the environment and learns from the interaction experience.

4 Solutions

This section focuses on the solution proposed to solve the above objective function, which is roughly divided into two steps. After the edge server calculates the offload, it first uses the K-means algorithm to cluster and associate users based on their location and the drones to determine each drone cluster and the users to be served. The multi-agent DQN method is then used to jointly optimize the trajectory and power allocation of the UAVs to accomplish the objective of minimizing system energy consumption [10]. Figure 2 shows a block diagram of joint optimization using a multi-agent DQN scheme. The solution will be described in detail in two areas next.

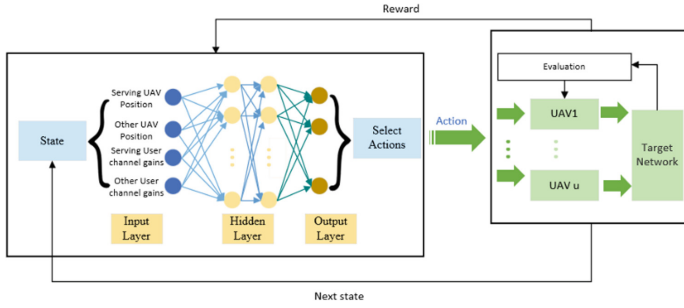


Fig. 2. Block diagram of the joint optimization of a multi-intelligent DQN scheme.

4.1 Clustering of User Based on K-Means

Spatial association of users using K-means clustering with an upper limit on cluster membership allows the offloaded users to be divided into multiple clusters to suppress inter-cluster interference. Within the system, we first divide, at random, the location of each user into U groups and randomly select U users as cluster centers for the clusters. The distance between the other users and each cluster center is calculated and these remaining users are assigned to the closest cluster of users. The number of users in each cluster is calculated, and if there are redundancies or vacancies, users are reassigned based on the nearest user cluster. The clustering center and the assigned users then represent a clustering cluster. For each user assigned, the position of the cluster center is recalculated based on the center distance of the existing users in the cluster. This process is repeated until all users have been allocated, or the cluster centers no longer change and the mean sum of squares of errors is minimized. Once users are allocated, the nearest UAV is selected for each user cluster to complete the clustering.

4.2 Deployment and Power Allocation of Multi-agent DQN Algorithm

In this subsection, we use a multi-intelligent DQN algorithm based on a user sub-clustering strategy to jointly optimize the UAV cluster trajectory and power allocation to minimize the system energy consumption. That is, multiple UAVs are deployed in the maritime cellular offload network, abstracting the scenario as intelligences that can choose their actions, and the UAVs exist independently of each other and do not know the choice of actions.

For each time t , the UAV as an agent decent to a state when it needs to select an action based on the state, receive a reward $r_t = r(s_t, a_t)$, and observe a new state s_{t+1} . Each intelligent body exists independently in the environment. Combined with the power allocation scenario, each part of the task is represented in the multi-intelligent body DQN framework as follows:

State: In the system multi-agent DQN model, each UAV and the cluster of users it serves is treated as an agent, and the state of the intelligent body consists of the cluster's location and channel gain [11]. Each agent interacts with the environment independently, allowing different agent to connect to the same neuron. As shown in Eq. (13):

$$S_t^u = \{L_u(t), L_s(t), g_k^u(t), g_k^s(t)\}, u, s \in \mathcal{U}, s \neq u, k \in \mathcal{K}$$

$$S_t = \{S_t^1, S_t^2, \dots, S_t^u\} \quad (13)$$

Actions: At time t , each agent $u \in U$ observes the current state of the environment $S_t \in S$ and chooses an action $A_t^u \in A^u$ based on a random policy π_u , and the actions of the intelligent agents form a joint action A_t . Each independent agent needs to choose the UAV flight action and the power allocation decision. The action usage is given by the following equation:

$$A_t = \begin{cases} A_L = \{\text{Left, right, front, back, up, down, stationary}\} \\ A_P = \{P_1, P_2, \dots, P_k\} \end{cases} \quad (14)$$

Reward: As a result of this joint action, all intelligences receive a reward signal $R_t = r(S_t, A_t)$ [12] and the environment is transferred to a new state $S_{t+1} \in S$ according to a transfer probability function $P(S_{t+1}|S_t, A_t)$ [13].

A natural approach to such a fully observable, cooperative multi-intelligence RL problem is to consider a “meta-intelligence” that chooses a joint action A_t based on π , π being a vector containing strategies π_u , $u \in U$, i.e. $\pi = (\pi_1, \dots, \pi_n)$. This meta-agent learns the Q function $Q(s, a) = E^\pi[G_t|S_t = s, A_t = a]$, which is conditional on the states and joint actions of all agents.

In summary, in a multi-intelligent DQN model, the intelligences can perform multiple discrete actions, and each network is trained independently without interfering with each other. The drone cluster feeds state information into the evaluation network after establishing appropriate connections with the neurons. After several iterations, the scheme results in higher reward values.

5 Analysis of Result

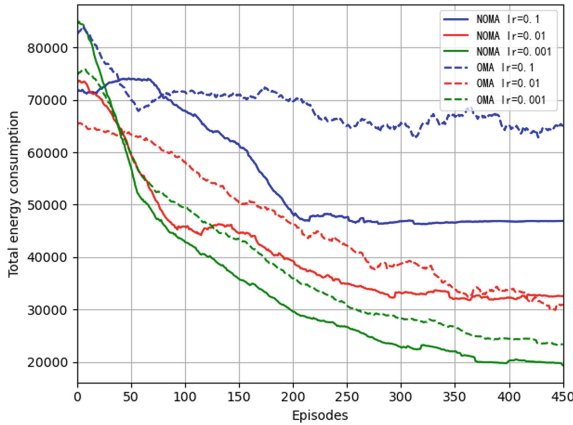
To determine the success of the suggested method and the gain of each component, numerical results are presented in this section. In the simulation scenario, we assume 3–20 UAVs and 10–60 sea vessel users distributed in a 3D space of $500 * 500 * 150$. Users are randomly distributed in the same horizontal plane in space. The hovering position of the UAV depends on the centroids of the user clusters and is optimized. Each UAV serves 2–10 users through NOMA. The neural network in use comprises three layers and a hidden layer with forty nodes. The mean square error is used as the loss function, while the activation function is a corrected linear unit. The neural network is trained using the Adam optimizer [14]. Table 1 includes a list of the additional default simulation parameters. Without a special justification, the simulation uses its default settings.

From Fig. 3, we can easily observe the convergence of the multi-agent model, with the increase of the number of algorithm iterations, the energy consumption decreases correspondingly, and the reduction of energy consumption is more significant for NOMA compared to OMA. However, we can see that as the learning rate increases it may cause the intelligences to update too quickly and not share information as quickly, therefore leading to a lower convergence of the curve.

Figure 4 shows the relationship between the number of errors in user association and the number of training sets, with NOMA also slightly outperforming OMA in terms of error rate.

Table 1. Simulation parameters

Parameter	Value
wavelength in meters λ	0.15
The height of evaporation duct layer h_e	20
Carrier frequency f_c	2 GHZ
Bandwidth	15 kHz
AWGN power	-100 dBm/Hz
Transmitting power of UAV	20 dBm

**Fig. 3.** Total energy consumption versus training episodes for OMA and NOMA.

It is assumed that a UAV can only serve two users. To facilitate the simulation of the degree of association in the 3D model, the UAV' is used to denote the projection of the UAV in the xy-axis plane, and a region of the same color is used to denote the cluster of UAVs and the users they serve under the same cluster. According to the above algorithm, and so that the multi-access method and energy consumption are selected differently, the UAV and user location state information is shown in Fig. 5. Any user under the same cluster with optimal and average energy consumption is selected for outage probability and transmission rate analysis. The analysis is shown below:

Figure 6 gives the outage probability performance of users within the same cluster for NOMA and OMA transmission methods with different energy consumptions. The decoding order of users under the same cluster within user 1 and 2 systems. to see that since each user needs to decode the interference of all previous users during the decoding process, user 2 can obtain a larger outage probability value for a suitable signal-to-noise ratio value. The reason for this is that since User 2 fails to adequately decode and remove the signal from User 1, the interfering signal from User 1 will gradually accumulate throughout the communication process, which will result in higher interference and thus higher outage probability performance for the highest-level user.

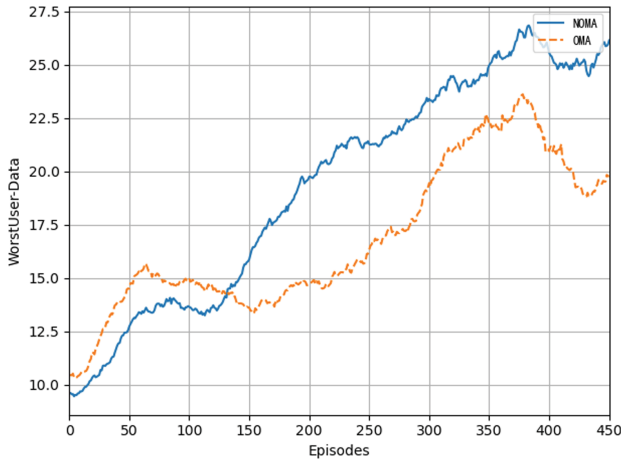


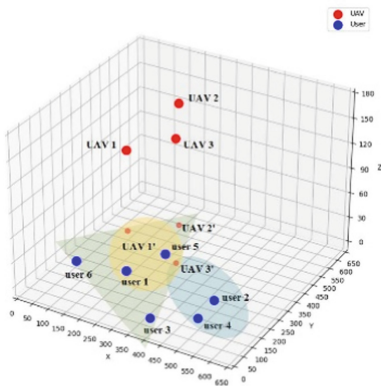
Fig. 4. Worst user-data rate in test episode with re-clustering.

And regardless of the user, the outage probability of interruption at optimal energy is lower than the outage probability at average energy consumption. This is due to the different co-power allocation at different energy consumptions, which results in different outage probabilities.

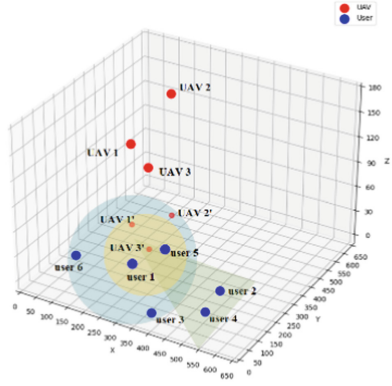
To verify the outage probability results in Fig. 6, the transmission rates of users within the same cluster are given in Fig. 7. As the signal-to-noise ratio increases, the transmission rate of the first user outperforms that of the second user, even if there are hardware defects and interference within the system. From the results, the user starts to have a more significant rate increase at a signal-to-noise ratio of 10dB under optimal energy consumption, while the user under average energy consumption has a change at around 15 dB. This indicates that the system is capable of accomplishing low-energy and high-rate transmissions in maritime communications.

6 Summary

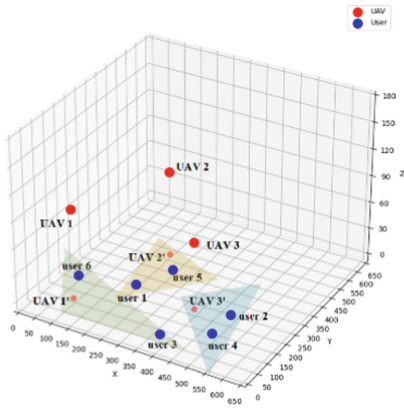
In order to improve the efficiency of offshore energy systems and ensure a tight connection and reliable operation of the sea-land grid. In this paper, a NOMA-based UAV-assisted offloading network is designed for optimizing communication performance under offshore multi-energy complementary power generation systems. The system uses UAVs equipped with edge servers and antennas to perform task offloading calculations and resource allocation tasks for offshore users. The system uses the K-means algorithm to make clustering associations between users and UAVs, and then uses a multi-intelligent deep Q-network algorithm to minimize system energy consumption by optimizing UAV cluster movement trajectories and user power allocation while considering system hardware loss and SIC incompleteness. The simulation results evaluate the performance of the proposed method through numerical results of energy consumption and error rate, comparing in terms of convergence, multi-access schemes, and learning rate. The cluster location analysis in the average energy consumption state in the selected



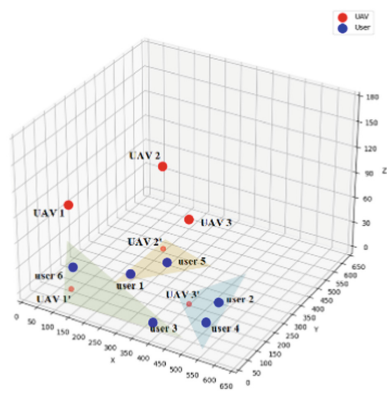
(a) The NOMA association case at optimal energy consumption.



(b) The OMA association case at optimal energy consumption.



(c) The NOMA association case at average energy consumption.



(d) The OMA association case at average energy consumption.

Fig. 5. Location and association information in the average state.

number of iterations and the simulation analysis of the interruption rate and transmission rate of users in this state for the same cluster and different multiple access techniques are given. These results demonstrate the superiority of the NOMA framework in the context of maritime UAV-assisted offload networks.

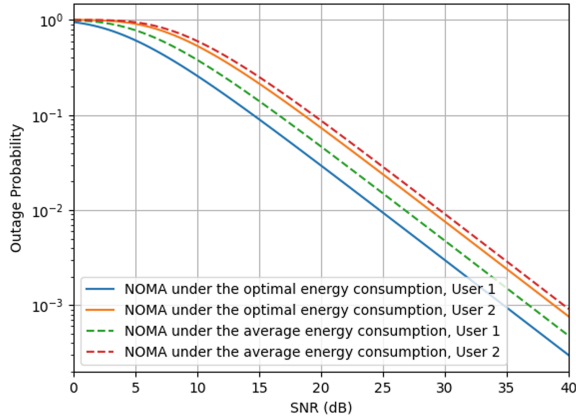


Fig. 6. Outage probability versus SNR.

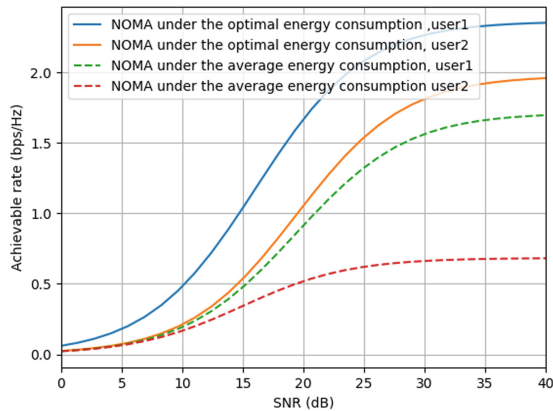


Fig. 7. Achieve rates within the same cluster.

Acknowledgment. This research was funded by the National Natural Science Foundation of China, grant number U2006222 and Natural Science Foundation of Shandong Province, grant number ZR2020MF138.

References

1. Gan, L., et al.: Balancing of supply and demand of renewable energy power system: a review and bibliometric analysis. *Sustain. Futures* **2**(100013), 100013 (2020). <https://doi.org/10.1016/j.sfr.2020.100013>. Accessed 23 May 2020
2. Shafi, M., Molisch, A.F., Smith, P.J., et al.: 5G: a tutorial overview of standards, trials, challenges, deployment, and practice. *IEEE J. Sel. Areas Commun.* **35**(6), 1201–1221 (2017)
3. Pei, H., Guo, S., He, Y., Wang, J.: Capacity optimization of concentrated solar power-photovoltaic wind Power combined generation system. *E3S Web Conf.* **118**, 02060 (2019)

4. Zhang, Y., Wang, X., Xu, Y.: Energy-efficient resource allocation in uplink NOMA systems with deep reinforcement learning. In: 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP), Xi'an, China, pp. 1–6 (2019)
5. Zhong, R., et al.: Multi-agent reinforcement learning in NOMA-Aided UAV networks for cellular offloading. *IEEE Trans. Wireless Commun.* 1 (2021). <https://doi.org/10.1109/twc.2021.3104633>. Accessed 30 Sep 2021
6. Guo, S., Zhao, X.: Multi-agent deep reinforcement learning based transmission latency minimization for delay-sensitive cognitive satellite-UAV networks. *IEEE Trans. Commun.* **71**(1), 131–144 (2023)
7. Nguyen, M.: Interference and resource management techniques for wireless networks. *EspaceINRS*. <https://espace.inrs.ca/id/eprint/13131/> (1970)
8. Lee, Y.H., Dong, F., Meng, Y.S.: Near sea-surface mobile radio wave propagation at 5 GHz: measurements and modeling. *Radio Eng.* **23**, 824–830 (2014)
9. Li, Z.-H., Li, N.-N.: Unmanned aerial vehicle landing point recognition based on natural landmark image. *J. Comput. Appl.* **30**(4), 971–973 (2010). <https://doi.org/10.3724/sp.j.1087.2010.00971>. Accessed 22 Nov 2019
10. Zhang, S., Zhang, H., Di, B., Song, L.: Cellular UAV-to-X communications: design and optimization for multi-UAV networks. *IEEE Trans. Wirel. Commun.* **18**(2), 1346–1359 (2019)
11. Wang, R.Y., Xu, L.: Multi-agent dam management model based on improved reinforcement learning technology. *Appl. Mech. Mater.* **198–199**, 922–926 (2012). <https://doi.org/10.4028/www.scientific.net/amm.198-199.922>. Accessed 24 Sep 2019
12. Liu, W., He, Y., Zhang, J., Qiao, J.: Deep reinforcement learning-based MEC offloading and resource allocation in uplink NOMA heterogeneous network. In: 2021 Computing, Communications and IoT Applications (ComComAp), Shenzhen, China, pp. 144–149 (2021)
13. Melo, D.C., Máximo, M.R.O.A., da Cunha, A.M.: Push recovery strategies through deep reinforcement learning. In: 2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE), Natal, Brazil, pp. 1–6 (2020)
14. Haworth, B., et al.: Deep integration of physical humanoid control and crowd navigation. *Motion, Interact. Games* (2020). <https://doi.org/10.1145/3424636.3426894>