Toward Energy-Efficient UAV-Assisted Wireless Networks Using an Artificial Intelligence Approach

Shu Fu, Meng Zhang, Min Liu, Chen Chen, and Fei Richard Yu

Abstract

This article studies the application of artificial intelligence (AI) approach in UAV-assisted wireless networks to cope with a large number of parameters impacting energy-efficiency in the sixth generation wireless network. In order to improve the energy efficiency for UAV-assisted wireless networks, we focus on the following three aspects: the UAVs trajectory planning; caching, computing, and communication resource allocation of UAVs; and 3D hovering location decision of UAVs. We discuss each aspect and reveal the corresponding optimization problem of energy efficiency. We also explore several promising deep-learning-based AI methods, which include pointer network, federated deep learning, and multi-agent deep deterministic policy gradient, to solve these optimization problems. Through case studies, we verify the superiority of the proposed AI methods to save UAVs' energy and decrease the system delay.

Introduction

Toward ubiquitous coverage and smart connection in the sixth generation (6G) wireless network, the traditional cellular networks with fixed topology among base stations (BSs) have weak coverage for the network edge, especially for the scenarios of deserts, oceans, mountain areas, and so on. To extend the network edge, unmanned aerial vehicles (UAVs) can provide a promising solution [1]. Configured with caching, computing, and communication (3C) resources, UAVs could be deployed to the network edge in a quick and effective manner, where the heterogeneous wireless services can be provided in a real-time manner.

Deploying multiple UAVs in the wireless network will bring a large number of optimization parameters such as the UAVs' flying paths and 3D locations, leading to difficulties in the network optimization. Recently, artificial intelligence (AI) has been deemed key to catering for the ever more complex and scaled wireless networks due to its powerful ability of data processing and analysis [2]. UAVs employ energy-limited onboard batteries for flying, hovering, and communication, resulting in a limited task duration and even poor performance. Thus, energy-efficient UAV-based dynamic wireless networks with AI methods should be given special attention.

Based on the above challenges, there are three main research directions regarding UAVs' energy efficiency: the UAVs' trajectory planning, the 3C resource allocation of UAVs, and the 3D hovering location decision. To cope with the large number of parameters, combined with big data training, AI techniques including reinforcement learning (RL), federated learning (FL) and deep learning (DL), and so on can tame the network complexity and implement UAVs' services in an energy-efficient manner.

A number of works have contributed to UAV-assisted network using traditional methods. In [3], Gong *et al*. investigated a UAV-based data collection mechanism for achieving the shortest flight time in an 1D wireless network. In [4], multiple UAVs were employed to collect the data from Internet of Things (IoT) devices, where the UAVs' trajectories were designed to maximize the minimum transmission rate of IoT devices. The work in [5] investigated fast UAV deployment problems considering different flying speeds, operating altitudes, and wireless coverage radii of UAVs, where the maximum deployment delay among all UAVs and the total deployment delay were minimized.

As the scale and complexity of the optimization problem increase, traditional methods require more human intervention to simplify and approximate the original problems, resulting in less adaptability. AI methods allow UAVs to learn from past experiences and build a self-organized way to adapt to the network environment and achieve autonomous optimization to minimize human intervention. On the other hand, AI methods can analyze and process massive amounts of data and time-related data. Thus, they possess the potential to solve large-scale optimization problems, and to track and respond to the dynamics of networks to provide real-time services. Therefore, many researchers have turned to AI techniques.

Fu *et al.* [6] used a Q-learning-based algorithm to generate the UAV trajectory with wireless chargers for collecting data from users in grids. In [7], an RL-based method was proposed to optimize the deployment and trajectory of UAVs. The work in [8] used a multi-agent RL framework to design the UAVs' trajectory and resource allocation in a downlink multi-UAV cellular network. Hu *et al.* [9]

Shu Fu (corresponding author), Meng Zhang, Min Liu (corresponding author) and Chen Chen are with Chongqing University, China; Fei Richard Yu is with Carleton University, Canada.

Digital Object Identifier: 10.1109/MWC.105.2100389

 1536-1284/22/$25.00 © 2022 IEEE

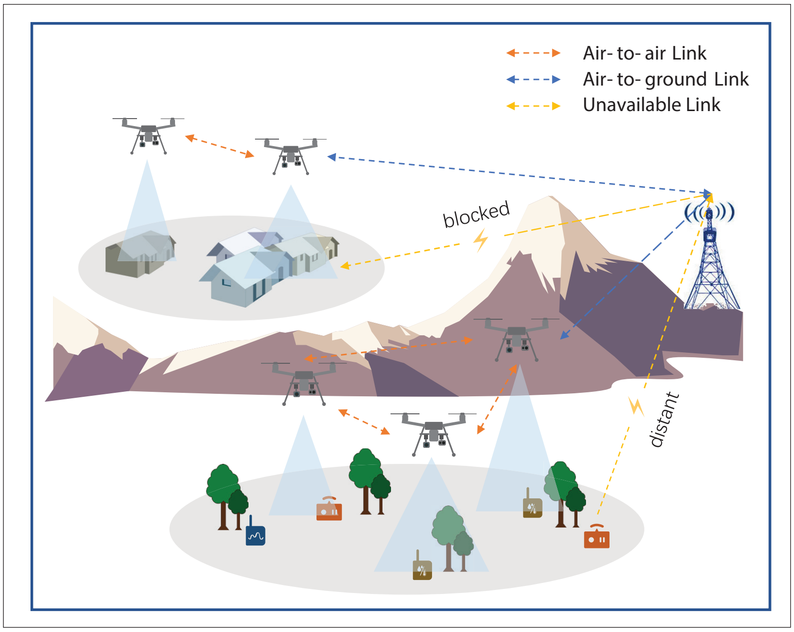

FIGURE 1. A UAV-assisted wireless network.

investigated the UAVs' trajectory in a decentralized manner to coordinate the UAVs implementing the real-time sensing application based on a multi-UAV Q-learning algorithm. The work in [10] used a deterministic policy gradient (DPG) algorithm to learn the UAV trajectory without user-side information.

To efficiently utilize the energy of UAVs, UAVs' flying trajectory, the served users on the trajectory, and the 3C resource allocation of UAVs should be jointly optimized by choosing appropriate AI methods. The large number of parameters should be sufficiently revealed and learned to dynamically generate wireless networks by re-arrangement of UAVs' hovering locations according to the requirements of data traffic.

Motivated by the above observations, this article focuses on the application of AI in UAVs' services from the energy efficiency perspective. Based on UAV-assisted wireless networks, we specifically discuss three main research directions of UAVs' services — the UAVs' trajectory planning, 3C resource allocation of UAVs, and 3D hovering location decision — as well as the applications of deep-learning-based AI methods in them. The contributions of this article are as follows:

- We show the basic architecture of the UAV-assisted wireless network, in which four aspects that primarily affect the performance of the UAVs are introduced, including air-to-ground channel, 3C resource carried by UAV, user pattern, and re-arrangement of UAVs. Some feasible AI approaches to solve the challenges in these aspects are also given.
- For the complex optimization problems involving trajectory planning, 3C resource allocation, and 3D hovering location decisions of UAVs, we provide promising DL-based AI approaches to solve them, respectively, which are the pointer network (PN), federated deep learning (FDL), and multi-agent deep deterministic policy gradient (MADDPG). The architecture, principles, and challenges of these AI approaches are introduced.
- Based on two special case studies, we validate the effectiveness of DL-based AI approaches in solving complex UAV optimization problems. One case is UAVs' flying path planning with PN; the other is UAVs' 3D hovering location decision with MADDPG. Simulation results show that the AI approaches can improve the energy efficiency by increasing the data amount collected by the UAVs at shorter total flying distance and enhance the system throughput.

The remainder of this article is organized as follows. We first introduce the basics of the UAV-assisted wireless network. Then AI-based UAV optimization is discussed to provide UAVs' services with energy efficiency awareness. Next, we study two special cases to verify the effectiveness of the corresponding AI techniques. Finally, we conclude this article.

Basics of UAV-Assisted Wireless Networks

As illustrated in Fig. 1, multiple UAVs as flying BSs to assist wireless networks can enable the enhancement of coverage, capacity, and connectivity in remote areas, post-disaster emergency communication scenarios, IoT, and so on, or form mobile 3C resource pools to provide heterogeneous services such as hot content caching and edge computing. To achieve these benefits, the challenges lie in the three main research directions of UAV optimization, whose performance is primarily affected by the four aspects in UAV-assisted wireless networks: air-to-ground (A2G) wireless channel, 3C resource carried by UAVs, user pattern, and re-arrangement of UAVs.

Air-to-Ground Channel

Lacking full knowledge of the specific environment, air-to-air (A2G) channels have been widely modeled as a probabilistic line-of-sight (LoS) model [11]. Such a model considers the impact of both LoS and non-line-of-sight (NLoS) on wireless channel gain. The probability of an LoS link occurrence between a UAV and a user is modeled as a sigmoid function (S-curve) with respect to α, β, and θ. Here, α and β are related to the statistical environment parameters such as the ratio of built-up land area to the total land area. θ represents the elevation angle from the user to the UAV. Thus, the total path loss between UAVs and users can be expressed as the expectation of path loss caused by LoS and NLoS links.

Considering the specific geographical environments, the extensive measurements in real environments to build universally applicable wireless channel models and the corresponding parameters should be further revealed [1]. AI techniques such as generative adversarial network (GAN), among others, can be effective to cater for the complex features of massive wireless channel parameters [12]. In GAN, two neural networks (NNs) are employed. One NN is used to generate the channel parameters, which are discriminated by the other NN. The real nonlinear channel can be well simulated when the two NNs reach the minimum Nash equilibrium through training with the measurement of parameters. Such a method can reduce the complexity and improve the accuracy of channel modeling when the wireless channels exhibit extreme irregularity.

3C Resources Carried by UAVs

In the heterogeneous services supported by the 3C resources pool formed by UAVs, such as data collection, hot content caching, edge computing, wireless coverage, and capacity enhancement, two connotations exist: the different amount of 3C resources at a UAV and the different hovering locations of UAVs that constitute a dynamic wireless network with different 3C resource distribution. The DL-based AI approach enables the UAVs to collect and analyze wireless environment and user-side information to predict their changes and provide a basis for intelligent 3C resource allocation.

Combining the advantages of deep neural network (DNN) and recurrent neural network (RNN), the deep echo state network (ESN) has the capability to process massive amounts of time-related and high-dimensional data to capture dynamic temporary information of users [13]. Hence, it can be used to predict the content of interest to users and then cluster the users with high similarity of requested content so that the cache resources of UAVs can be allocated efficiently. In similar fashion, deep ESN can be used to predict the requirements of classified computational tasks, thus enabling UAVs to schedule computational resources reasonably and quickly to reduce latency. Combined with DNN, the deep reinforcement learning (DRL) approach enables UAVs to learn the channel characteristics and user behavior from the interactions with the wireless environment and ground users. Then UAVs can find the optimal policy to make 3D deployment decision, channel selection, spectrum division, power allocation, and user association.

User Pattern

The user pattern contains two parts, which are the user distribution and the user dynamics. When users to be served by UAVs present uneven distribution, a common and efficient pre-processing method is traffic-oriented user clustering according to the heterogeneous traffic requirements, where several intelligent clustering algorithms such as *K*-means and density-based spatial clustering of applications with noise (DBSCAN), among others, can be employed [2]. As an instance, unlike the traditional clustering algorithms generating circular clusters, DBSCAN employs a density threshold to generate clusters with different shapes according to the user distribution and traffic requirements.

By user clustering algorithms, a UAV can hover above a cluster to serve the users simultaneously. From the viewpoint of energy efficiency, the number of users in a cluster and the coverage range of the cluster largely impact the energy consumption of wireless transmission between the UAV and users. A large number of users in a cluster will also increase the system complexity and the number of antennas at a UAV. This will increase the energy consumption. Hence, parameters such as the maximum number of users per cluster and the maximum distance between two arbitrary users in a cluster should be jointly considered in intelligent clustering algorithms.

Traditionally, the number of users and user distribution in networks are generally fixed. However, toward 6G, in IoT networks and so on, the requirements of users, the number of users, and the locations of users may be dynamically generated and rapidly changing. The high user dynamics pose a significant impact on the optimization of UAV-assisted wireless networks, such as UAVs' serving path and hovering location invalidation, outdated cached hotspot content, wasted computational resources, and unreasonable allocation of communication resources. To address these issues, the deep ESN method can track the users in real time with the flexibility of UAVs and collect the time-related information such as user mobility and content requests to predict the user location and content caching at UAVs in the next time slot. Thus, the planning and optimization of the UAV-assisted wireless network can be dynamically implemented to provide real-time and reliable services.

Rearrangement of UAVs

One trade-off in controlling and re-arranging UAVs as different UAV-based dynamic wireless networks is the number of rearrangements and the maintenance time for one wireless network topology. Under the constraint of the maximum available energy of a UAV, the larger number of rearrangements can bring more topologies of UAV-based wireless networks. This can provide 3C resources to serve users with higher energy efficiency and lower delay. However, after the serving time of a wireless network, UAVs will consume a configuration time for the rearrangement, where UAVs will fly to updated hovering locations; the UAVs cannot provide services during the configuration time. This suggests that when the UAVs frequently update the wireless network topologies, the available total serving time will be decreased under the constraint of the maximum available energy of a UAV.

Such a trade-off leads to new parameters that should be optimized. UAVs' cooperation to form a computing pool supporting part of the computing load for machine learning will be further studied in future work. One possible AI method is multi-agent RL, such as multi-agent deep deterministic policy gradient (MADDPG), which is a potential tool to manage a great number of UAVs [14]. By MADDPG, multiple UAVs can be trained centrally with the same target to learn the optimal policy based on the data of all UAVs. Then UAVs perform actions in an independent and decentralized manner, based on which an intelligent UAV swarm with strong flexibility and stability can be formed.

AI-Based UAV Optimization

In this section, we focus on the potential of DL-based AI methods in solving complex multi-UAV optimization problems, involving the flying path planning, 3C resource allocation, and 3D hovering location decision of UAVs. Aimed at these problems, suitable DL methods are introduced separately, including their architectures, principles, and challenges.

Flying Path of UAVs

Assume that there are several clusters of users distributed on the ground requesting data collection services from UAVs. Each user cluster on the ground is treated as a node. When the user clusters to be served by a UAV are pre-determined, optimizing the paths to minimize the total flying distance of the UAVs can be directly modeled as a traditional traveling salesman problem (TSP) model, which can be solved easily by a genetic algorithm.

Traditionally, the number of users and user distribution in networks are generally fixed. However, toward 6G, in IoT networks and others, the requirements of users, the number of users, and the locations of users may be dynamically generated and rapidly changing.

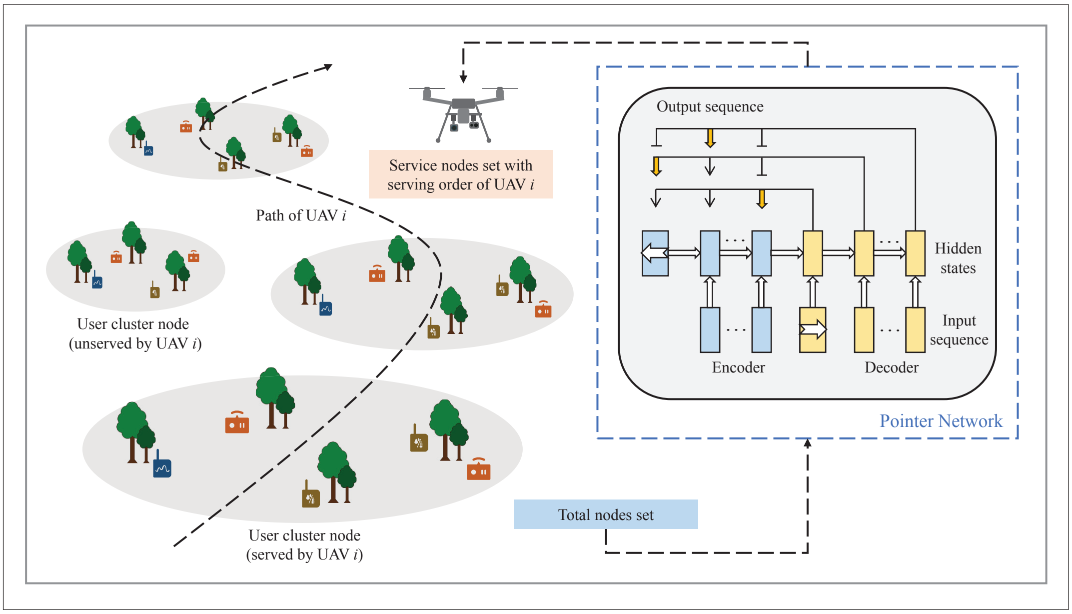

FIGURE 2. UAV flying path planning based on a pointer network.

However, in practice, the limitation of UAVs' energy leads to being able to serve only a set of user clusters. On the other hand, to select the set of user clusters for UAVs' serving, the traffic priorities, serving time, and so on, should be jointly considered, where the selection of user clusters and UAVs' flying trajectory will be integratedly optimized. To cope with such an integrated optimization problem, first, a user cluster will be given a weight value that denotes the priority of the user cluster to be served by a UAV. Then, constrained by the maximum total available serving time of a UAV determined by its maximum energy, the user cluster selection for UAVs' serving can be modeled by a knapsack problem (KP) to maximize the sum weight values of the served user clusters.

KP and TSP can be combined as the optimization problem of UAVs' serving, defined as an orienteering problem (OP). Denote the initial set of all the user clusters by CA. To solve the problem of OP, we should find an appropriate AI algorithm to generate a set of user clusters denoted by C_I that need to be served and the order of service for an arbitrary UAV *i*. A PN based on RNN can be effectively migrated to solve the OP due to its features of adaptability in handling time- or sequence-dependent problems [15].

As shown in Fig. 2, PN mainly consists of an encoder and a decoder, which are composed of one or more layers of long short-term memory (LSTM) networks. To obtain the path of an arbitrary UAV *i*, the total set of nodes C_A, as the input sequence, is encoded into intermediate vectors via feature extraction by the encoder. According to the attention mechanism, different weights are assigned to the hidden states of the encoder and decoder, where the weights are the network parameters to be trained. Based on the assigned weights, the probability distribution of the input nodes is then obtained using the softmax function. Finally, the output sequences (i.e., the service nodes and service order C_i of UAV *i*) are decoded from intermediate vectors by the decoder.

Although PN offers a promising solution for UAV path planning, there are still some challenges. First, in the scenario with a wireless charger to alleviate the energy shortage of UAVs, the placement of chargers should be studied before implementing PN for UAVs' flying trajectories. Such planning involves complex factors including the average user distribution, geographical conditions, a UAV's maximum available energy, and so on. Second, obstacle avoidance among UAVs should be considered in the PN. Furthermore, when the traffic of a user cluster has been served by a UAV, the user will not be served again. This suggests that the required amount of 3C resources will be balanced among UAVs in the PN model, which will be an open problem in future work.

3C Resource Allocation of UAVs

The task-oriented clustering of users is performed based on the demand for 3C resources. The heterogeneous service requirements can be satisfied through UAVs' rearrangement, where 3C resources will be re-allocated according to the traffic requirements to minimize the energy consumption and system delay. Toward the highly dynamic network in 6G, the number of users and traffic requirements change in real time. 3C resource allocation should be implemented based on the heterogeneous traffic demand, time-varying user distribution, and so on.

DL-based AI algorithms can enable UAVs to predict the dynamic traffic requirements and user distribution. However, the existing DL algorithms are generally implemented in a centralized manner, which challenges the protection of the private data. On the other hand, the processing and wireless transmission of a large amount of data for the central computing in DL lead to large latency, invalidating the real-time predictions. By combining FL and DL, a distributed DL framework, federated deep learning (FDL), can be formed to tackle this challenge [14].

As in Fig. 3, K UAVs constitute a set of FDL participants $\mathcal{K} = \{1, 2, ..., K\}$. Each UAV $k \in \mathcal{K}$ has a local dataset LDS_k for local training containing private data such as user location, user behavior, UAV location, and so on. By using the local dataset LDS_k, UAV *k* trains a local model LM_k and sends the model parameters to the FDL server. All received local models LMk are aggregated into a global model $GM = \cup_{i\in\mathcal{K}} LM_k$ through aggregation algorithms such as FedAvg. To update the parameters of the global model *GM*, an FDL server sends *GM* as a shared model back to each UAV to guide the next round of local model training. Based on the distributedly iterative learning, FDL can effectively decrease the system delay and safeguard the data privacy, because the parameters of the trained local models are sent to the FDL server instead of sending a large amount of parameters and data. By FDL, 3C resources of UAVs can be efficiently allocated to the network in a dynamic and real-time manner.

However, FDL also faces challenges in robustness and convergence. In terms of robustness, UAVs may drop out of the FDL training due to deteriorated channel conditions and energy shortage. Thus, FDL needs to be able to predict such situations and keep the training results valid. In terms of convergence, the distributed learning and the different configured 3C resources of UAVs make FDL convergence difficult. These two challenges contribute to an open problem of 3C resources allocation based on FDL to maximize energy efficiency and minimize latency.

3D Hovering Locations of UAVs

When a UAV arrives over a user cluster to be served, the 3D hovering location of the UAV needs to be carefully designed to optimize the A2G channel gains between the UAV and users. Thus, the communication between the UAV and users can achieve a large signal-to-noise ratio to improve the system throughput.

For the given environment parameters, A2G channel gain based on the probabilistic LoS model can be maximized by the optimal elevation angle between the UAV and the user when only a single user exists. However, when multiple UAVs serve a user cluster with multiple users, the optimization of 3D hovering locations for UAVs is still an open problem. Furthermore, when the wireless interference among users is considered, the problem will be more complex. In this case, finding the optimal hovering locations of UAVs is intractable, and appropriate AI algorithms adaptive to the large number of parameters should be studied in future work.

Reinforcement learning enables UAVs to learn from past experiences and make intelligent decisions [2]. To compress the state and action spaces of RL, DNNs are used to approximate the value function. Then a framework of DRL can be achieved to handle high-dimensional continuous variables, including UAVs' hovering locations and velocities, wireless transmitting power, and so on. As shown in Fig. 4, the MADDPG is employed to learn the optimal 3D hovering locations and power allocation for the UAV-based dynamic wireless network. Each UAV has a DDPG network, consisting of an evaluation network and a target network, each of which has an actor-critic network. The actor network implements the mapping from states to actions, while the critic network scores the actions outputted by the actor [14].

To ensure the effectiveness of joint training of multiple agents, centralized control is required where the states and actions of other agents need to be considered in the critic network of each agent. Once the training is completed, each agent can execute the action only according to their own state. Take UAV *i* for example; at time step *t*, UAV *i* observes a state s_t of the current location, and then obtains a reward r_t from the environment,

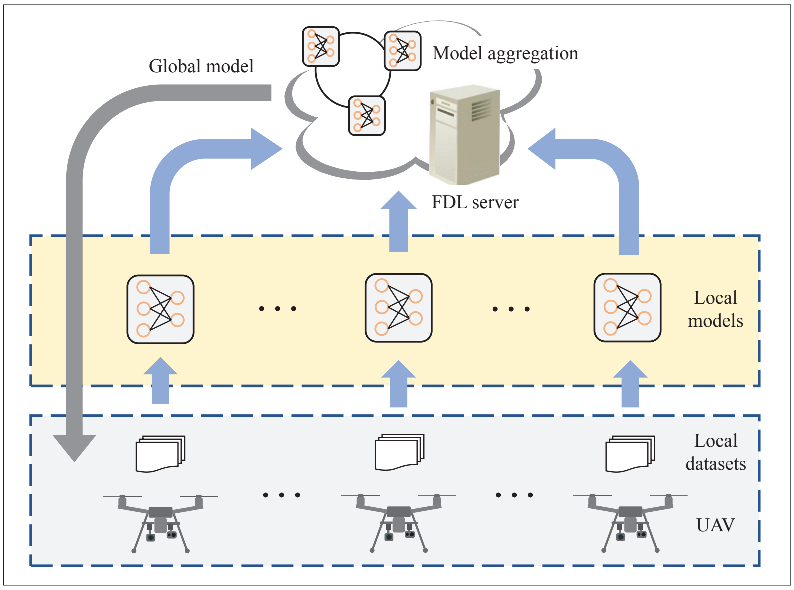

FIGURE 3. The learning process of UAV resource reconfiguration based on federated deep learning.

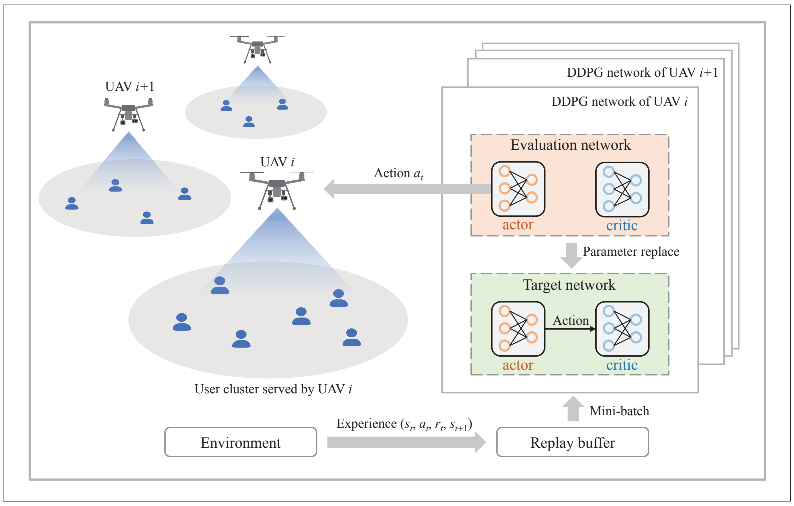

FIGURE 4. The 3D hovering location decision of UAVs based on multi-agent deep deterministic policy gradient.

which is designed based on the specific objective such as maximizing throughput. For a state s_t, the actor network outputs an action a_t of the UAV's displacement to the next location and the allocated power for users. By executing the action a_t, UAV *i* transfers to the next state s_{t+1}. The corresponding tuple of (s_t, a_t, r_t, s_{t+1}) is stored as an experience in the reply buffer. Only the evaluation network will be trained by mini-batch gradient descent using the sampled experiences from the replay buffer. The target network is updated by copying a fraction of the parameters from the evaluation network.

In addition to the issues of low sample utilization and complex reward function design that DRL inherently carries, some challenges remain for DRL to be applied in the 3D deployment of UAVs. For instance, the high dynamics of the network, especially the mobility of the users, lead to the static deployment of UAVs losing optimality. Therefore, dynamic deployment of UAVs is claimed so that UAVs can maintain optimal performance by adjusting their 3D hovering locations and power allocation according to the locations of users and traffic requirements in real time.

Method	U = 50			U = 100		
	DA_{all} (Gb)	FD_{all} (km)	EE (bit/J)	DA_{all} (Gb)	FD_{all} (km)	EE (bit/J)
PN-AS	23.64	11.6066	6.79×10^4	46.92	16.3550	9.57×10^4
GA	25.68	12.6440	6.77×10^4	51.60	18.6348	9.23×10^4
PN-GP	25.68	12.6632	6.75×10^4	51.60	18.7232	9.18×10^4
PN	13.85	14.5041	3.18×10^4	25.28	29.9044	2.82×10^4

TABLE 1. The simulation results of the multi-UAV flying path case.

Case Study

Multi-UAV Flying Path

We assume that U nodes denoting user clusters are randomly distributed in a 1 km × 1 km area, each node with a randomly data amount with the unit of gibabit as a weight value. We consider the scenario in which three UAVs provide uplink data collection services for these nodes. Due to the energy limitation, the three UAVs can only serve some of the U nodes before returning to the depot, where UAVs can be charged and the data can be offloaded.

The general expression for energy efficiency (EE) is the ratio of system throughput to energy consumption. In this case, we can ignore the energy consumed by the data transmission between the UAVs and user nodes because the communication power is generally much smaller than the propulsion power of the UAVs. Assume that the energy consumed by the UAVs per kilometer of flight is fixed. To improve the EE of UAVs, we simultaneously maximize the total amount of collected data, denoted by DA_{all}, and minimize the total flying distance of the UAVs, denoted by FD_{all} through a PN with active search (PN-AS) policy.

In PN-AS, the nodes are formed into multiple batches of input sequences though permutation and combination, and then inputted into PN for training. The first and last nodes in output sequences are fixed to satisfy the constraint that the UAVs both start and end at the depot. The idea of Tabu Search is adopted to avoid repetitive services. Three baseline methods are used to compare the performance: the PN with greed prize (PN-GP) policy, genetic algorithm (GA), and the traditional PN without policy, respectively. In PN-GP and GA, the nodes with the largest amount of data are first selected, and then the serving order of these nodes is optimized by PN and GA, respectively.

In Table 1, we give the DA_{all}, FD_{all}, and EE of the three UAVs according to the flying paths planned by the methods of PN-AS, GA, PN-GP, and PN with U = 50 and U = 100, respectively. Comparing PN, we can observe that the traditional PN has the poorest performance, which suggests that PN needs to combine appropriate policies for better performance. We can also observe that based on GA and PN-GP, the UAVs can serve the nodes with the largest amount of data. Thus, GA and PN-GP are able to collect the largest data amount. However, this may lead to an increase in the flying distance of UAVs and complete loss of user fairness. PN-AS is close to GA and PN-GP in terms of data amount with shorter flying distance of UAVs. This is because PN-AS considers both the joint optimization of maximizing the data amount and minimizing the flying distance of UAVs. Hence, PN-AS has the best energy efficiency. Additionally, the user fairness of PN-AS can be improved compared to that of PN-GP.

Multi-UAV 3D Hovering Locations

We consider a 4 km × 4 km area with four user clusters distributed in the four quadrants, each of which has 100 users. We assume that four UAVs are dispatched to provide downlink services for each user cluster. The probabilistic LoS channel model is adopted. We consider the urban environmental parameters [11].

To improve the EE of UAVs in this case, which is dealt as the ratio of the system throughput to the total transmitting power of the UAVs, we use MADDPG to learn the distribution of users and the channel conditions, based on which the UAVs can find the optimal 3D hovering locations and power allocation that maximizes the system throughput. Thus, given a fixed total transmitting power of the UAVs, the EE is also maximized.

To apply MADDPG, we design the system EE as the body of the reward function plus the constraints in the model (e.g., serving area limit, transmitting power limit, distance limit for arbitrary two UAVs) as additional penalty reward, the UAVs' 3D locations as the state, and the UAVs' displacement to the next state and the power allocation as the action. We compare the MADDPG-based 3D deployment with two traditional methods: altitude optimization (AO) and GA. In AO, the horizontal locations of the UAVs are fixed in the center of the user clusters, and the users are allocated equal power. Then the UAVs' altitudes are optimized by gradient descent. In GA, the 3D locations of the UAVs are optimized with equal power allocation.

In Table 2, we demonstrate the optimal 3D deployment and EE of the four UAVs providing downlink traffic for the four user clusters derived by MADDPG, GA, and AO, respectively. We observe that compared to AO and GA, the MADDPG-based 3D deployment of UAVs results in a significant increase in EE. This is because MADDPG enables the UAVs to learn the features of the user distribution and wireless channel in the 3D direction to optimize not only the 3D hovering locations but also the high-dimensional power allocation variables with the objective of maximizing the system throughput and EE. We also observe that the UAVs' altitudes in AO are generally much higher than in MADDPG and GA. This is because the UAVs' horizontal locations are fixed in AO. Then the UAVs have to fly higher to cover the ground users under the limitation of the minimum elevation angle between users and a UAV.

Method	Cluster 1		Cluster 2		Cluster 3		Cluster 4	
	Deployment (km)	EE (bps/W)	Deployment (km)	EE (bps/W)	Deployment (km)	EE (bps/W)	Deployment (km)	EE (bps/W)
MADDPG	(1.31, 1.22, 0.26)	5.55×10^9	(–0.72, 1.31, 0.24)	5.48×10^9	(–0.79, –0.67, 0.23)	5.22×10^9	(1.33, –0.70, 0.24)	5.26×10^9
GA	(1.31, 1.19, 0.27)	5.28×10^9	(–0.71, 1.26, 0.28)	5.18×10^9	(–0.75, –0.72, 0.35)	4.93×10^9	(1.36, –0.72, 0.29)	4.93×10^9
AO	(1.00, 1.00, 0.52)	4.35×10^9	(-1.00, 1.00, 0.55)	4.15×10^9	(-1.00, -1.00, 0.57)	4.13×10^9	(1.00, -1.00, 0.54)	3.68×10^9

TABLE 2. The simulation results of the multi-UAV 3D hovering location case.

Conclusion

This article has studied an AI-based energy-efficient UAV network for 6G. We have studied the energy efficiency from multiple significant aspects. First, we have studied the combined KP and TSP problem in UAVs' flying paths, where pointer network deep learning was used to generate the UAVs' trajectory. Second, we have studied 3C resource allocation for UAVs, where federated deep learning is discussed. We propose the FDL-based 3C resource allocation for maximizing energy efficiency and minimizing latency as an open problem, which will be further investigated in our future work. Third, we have studied the 3D hovering locations of UAVs, where MADDPG is used to determine the optimal 3D hovering locations of UAVs. These can improve the energy efficiency of UAVs' dynamic wireless networks from different aspects. Finally, in a case study, we have also verified the effectiveness of our proposed AI methods.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62271093 and Grant U21A20448, and in part by the Fundamental Research Funds for the Central University under Grant 2022CDJJJ-003.

References

[1] M. Mozaffari *et al.*, "A Tutorial on UAVs for Wireless Networks: Applications, Challenges, and Open Problems," *IEEE Commun. Surveys & Tutorials*, vol. 21, no. 3, 2019, pp. 2334–60.

[2] H. Yang *et al.*, "Artificial-Intelligence-Enabled Intelligent 6G Networks," *IEEE Network*, vol. 34, no. 6, 2020, pp. 272–80.

[3] J. Gong *et al.*, "Flight Time Minimization of UAV for Data Collection Over Wireless Sensor Networks," *IEEE JSAC*, vol. 36, no. 9, 2018, pp. 1942–54.

[4] W. Luo *et al.*, "Joint 3D Trajectory and Resource Optimization in Multi-UAV Enabled IoT Networks with Wireless Power Transfer," *IEEE IoT J.*, 2020.

[5] X. Zhang and L. Duan, "Fast Deployment of UAV Networks for Optimal Wireless Coverage," *IEEE Trans. Mobile Computing*, vol. 18, no. 3, 2019, pp. 588–601.

[6] S. Fu *et al.*, "Energy-Efficient UAV Enabled Data Collection via Wireless Charging: A Reinforcement Learning Approach," *IEEE IoT J.*, 2021.

[7] X. Liu, Y. Liu, and Y. Chen, "Reinforcement Learning in Multiple-UAV Networks: Deployment and Movement Design," *IEEE Trans. Vehic. Tech.*, vol. 68, no. 8, 2019, pp. 8036–49.

[8] S. Yin and F. R. Yu, "Resource Allocation and Trajectory Design in UAV-aided Cellular Networks Based on Multi-Agent Reinforcement Learning," *IEEE IoT J.*, 2021.

[9] J. Hu, H. Zhang, and L. Song, "Reinforcement Learning for Decentralized Trajectory Design in Cellular UAV Networks with Sense-and-Send Protocol," *IEEE IoT J.*, vol. 6, no. 4, 2019, pp. 6177–89.

[10] S. Yin *et al.*, "Intelligent Trajectory Design in UAV-Aided Communications with Reinforcement Learning," *IEEE Trans. Vehic. Tech.*, vol. 68, no. 8, 2019, pp. 8227–31.

[11] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP Altitude for Maximum Coverage," *IEEE Wireless Commun. Letters*, vol. 3, no. 6, 2014, pp. 569–72.

[12] Y. Yang *et al.*, "Generative-Adversarial-Network- Based Wireless Channel Modeling: Challenges and Opportunities," *IEEE Commun. Mag.*, vol. 57, no. 3, Mar. 2019, pp. 22–27.

[13] C. Gallicchio and A. Micheli, "Deep Reservoir Computing: A Critical Analysis," *Proc. Euro. Symp. Artificial Neural Net. Comp. Intelligent Machine Learn*ing, Bruges, Belgium, Apr. 2016, pp. 1–6.

[14] D. Kwon *et al.*, "Multiagent DDPG-Based Deep Learning for Smart Ocean Federated Learning IoT Networks," *IEEE IoT J.*, vol. 7, no. 10, 2020, pp. 9895–9903.

[15] Y. Zhang *et al.*, "A Pointer Network Based Deep Learning Algorithm for User Pairing in Full-Duplex Wi-Fi Networks," *IEEE Trans. Vehic. Tech.*, vol. 69, no. 10, 2020, pp. 12,363–68.

Biographies

Shu Fu (shufu@cqu.edu.cn) received his Ph.D. degree in communication and information systems from the University of Electronic Science and Technology of China, Chengdu, in June 2016, focusing on resource allocation in wireless networks and wavelength-division-multiplexing-based optical networks. He is an associate professor with the School of Microelectronics and Communication Engineering, Chongqing University, China. His research interests include the next generation of wireless networks, integrated networks, and artificial intelligence.

Meng Zhang (201912131041@cqu.edu.cn) received his M.Eng. degree from Chongqing University in 2022.

Min Liu (liumin@cqu.edu.cn) received her B.Eng. and M.Eng. degrees from Chongqing University in 1997 and 2000, respectively, and her Ph.D. degree from Nanyang Technological University, Singapore, in 2004. She is currently a professor and the Vice Dean of the School of Microelectronics and Communication Engineering, Chongqing University. Her main research interests include optical fiber communication and wireless-optical networks.

Chen Chen [M] (c.chen@cqu.edu.cn) received his B.S. and M.Eng. degrees from the University of Electronic Science and Technology of China in 2010 and 2013, respectively, and his Ph.D. degree from Nanyang Technological University, Singapore, in 2017. He was a postdoctoral researcher with the School of Electrical and Electronic Engineering, Nanyang Technological University from 2017 to 2019. He is currently a tenure track assistant professor with the School of Microelectronics and Communication Engineering, Chongqing University. His research interests include optical wireless communication, optical access networks, the Internet of Things, and machine learning.

Fei Richard Yu [F] (richardyu@cunet.carleton.ca) received his Ph.D. degree in electrical engineering from the University of British Columbia, Vancouver, Canada, in 2003. From 2002 to 2006, he was with Ericsson, Lund, Sweden, and a startup in California. In 2007, he joined Carleton University, Ottawa, Ontario, Canada, where he is currently a professor.His research interests include wireless cyber-physical systems, connected/autonomous vehicles, security, distributed ledger technology, and deep learning. He received the IEEE Outstanding Service Award in 2016, the IEEE Outstanding Leadership Award in 2013, the Carleton Research Achievement Award in 2012, the Ontario Early Researcher Award (formerly Premiers Research Excellence Award) in 2011, the Excellent Contribution Award at IEEE/IFIP TrustCom 2010, the Leadership Opportunity Fund Award from the Canada Foundation of Innovation in 2009, and the Best Paper Awards at IEEE VTC 2017-Spring, IEEE ICC 2014, IEEE GLOBECOM 2012, IEEE/IFIP TrustCom 2009, and the International Conference on Networking 2005. He is a registered Professional Engineer in Ontario, Canada, and a Fellow of the Institution of Engineering and Technology. He has served as the Technical Program Committee Co-Chair of numerous conferences. He serves on the Editorial Boards of several journals, including as Co-Editor-in-Chief for *Ad Hoc and Sensor Wireless Networks*, a Lead Series Editor for *IEEE Transactions on Vehicular Technology*, *IEEE Transactions on Green Communications and Networking*, and *IEEE Communications Surveys & Tutorials*. He is a Distinguished Lecturer, Vice President (Membership), and an Elected Member of the Board of Governors of the IEEE Vehicular Technology Society.