Sum Rate Maximization in UAV-aided FL Systems with MC-NOMA

Manh Cuong Ho, Anh Tien Tran, 이충현, Thanh Phung Truong, 김가현, 서정택*, 권구락**, 조성래

{hmcuong, attran, chlee, tptruong, ghkim}@uclab.re.kr, seojt@gachon.ac.kr, grkwon@chosun.ac.kr, srcho@cau.ac.kr

Chung-Ang Univ., *Gachon Univ., **Chosun Univ.

ABSTRACT

Federated learning (FL) assumes a crucial role in forthcoming Internet of Things (IoT) networks, emerging as a decentralized machine learning technique with promise. It ensures the security and privacy of user data while judiciously utilizing communication resources. This paper addresses the challenge of maximizing the sum rate in UAV-aided FL systems with MC-NOMA. We jointly optimized UAV location, power allocation, CPU frequency, and sub-channel association variable allocation to achieve this goal. We formulated the optimization problem as a Markov decision process (MDP) and employed a deep deterministic policy gradient (DDPG) reinforcement learning algorithm to solve it.

I. INTRODUCTION

In UAV-assisted FL networks, user devices collaborate to execute a learning task by uploading their local learning models to a UAV without the necessity to share all their training data. Nevertheless, the constraints posed by limited wireless resources, finite device energy, and the dynamic movement of UAVs or users are significant challenges that may impair FL performance. Hence, we focus on maximizing the sum rate in UAV-aided FL systems using MC-NOMA by optimizing UAV location, power allocation, CPU frequency, and sub-channel association allocation in this work. The structure of this paper is as follows: Section II describes the system model and problem formulation. Section III explains the DDPG algorithm used to optimize the sum rate. Finally, Section IV concludes this work.

Ⅱ. SYSTEM MODEL AND PROBLEM FORMULATION

We consider an UAV-based FL network where a single antenna UAV serves N single antenna users. The UAV can operate as an aerial server by carrying a cloud server and hover at a fixed point in the aerial. We assume UAV and users have sufficient batteries to complete whole FL training process. Each user device is equipped with an on-board computing processor to train the local model. The operation period of the system T is divided into time slots with duration time of each time slot t, and each slot is indexed by in T = {1, 2,..., T}. The FL round is divided into steps. Firstly, at the beginning of time slot t, the server broadcasts the global parameter to all users. Each user n uses its data to train a local model based on its dataset. Then, each user n sends the local parameter to the UAV in the uplink. After receiving the last local model, the server fuses all local models to build a new global model and broadcast the global parameter to all users, then start the next time slot. We assume that there is a dedicated control channel for the simple slot and UAV can feedback all SINR info to all user through the dedicated channel. Denote by $q = \{a, b\} \in R^{2\times1}$ and $q_{n,t} = \{a_n, b_n\} \in R^{2\times1}$ are the x, y coordinates of the UAV and user n. The distance $d_{n,t}$ between the UAV and user n at time slot t is defined as:

$$d_{n,t} = \sqrt{(H-h_n)^2 + ||q-q_{n,t}||_2^2},$$

where H, q, h_n, $q_{n,t}$ are the height of UAV, the coordinate of the UAV, the height of user n, and the coordinate of the user n, respectively. We assume that the UAV moves with fixed altitude H within a circle of radius C to transmit signal to users. Let $g_{n,t}$ denote the channel coefficient between the UAV and user n at time slot t. In this paper, we assume that links between the UAV and users are NloS links. For user k in sub-channel m at time slot t, the achievable rate in bits per second (bps) is given as

$$r_{k,m,t} = \nabla_{k,m,t} b_t \log_2\left(1 + \frac{|g_{k,m,t}|^2 p_{k,t}}{\sum_{k'=1}^{k-1} \nabla_{k',m,t}|g_{k',m,t}|^2 p_{k',t+\sigma^2 b_t}}\right),$$

where $b_t, \nabla_{k,m,t}, p_{k,t}, \rho^2, g_{k,m,t}$ are the bandwidth allocated to each sub-channel, the sub-channel association variable, the transmit power of user k, the power spectral density of the Gaussian noise at time slot t, and the channel gain between UAV and user k. The rate of user k must meet a minimum quality of service (QoS). The sub-channel association variable indicates user selection for each sub-channel. If user k selects the sub-channel m, we set the sub-channel association variable $\nabla_{k,m,t}$= 1. Otherwise, $\nabla_{k,m,t}$ = 0.

Let's denote the model size of user n is denoted as Z_n. Then the uploading time is calculated as

$$t_{k,m,t}^{u} = \frac{Z_n}{r_{k,m,t}}.$$

The completion time of a FL round should be lower than the total time of system in a time slot. The completion time of a round FL of user n at time slot t can be stated as

$$t^{ct} = max_{n\in N}\,(t_{n,t}^{u} + t_{n,t}^{c}),$$

where $t_{n,t}^{u}, t_{n,t}^{c}$ are the uploading time, and the computation time to train the local models.

We consider the optimization of location of UAV, users, power allocation, CPU frequency allocation, and sub-channel association variable allocation to maximize the sum rate of UAV-aided FL systems with MC-NOMA.

$$\max_{q,p,f,\nabla} \sum_{t=1}^{T} \left(\sum_{n=1}^{N} r_{n,t} \right)$$

$$
\begin{aligned}
s.t.\quad & 0 \le p_{n,t} \le p_{n,t}^{max}, \forall n \in N, \forall t \in T \\
& f_{n,t}^{min} \le f_{n,t} \le f_{n,t}^{max}, \forall n \in N, \forall t \in T \\
& ||q||_2^2 = a^2 + b^2 \le C^2 \\
& q, q_{n,t} \ge 0, \forall n \in N, \forall t \in T \\
& t^{ct} < t^{to} \\
& r_{n,t} \ge r_Q \\
& \textstyle\sum_{m=1}^{M} \nabla_{n,m,t} \in \{1,0\}, \forall n \in N, \forall t \in T
\end{aligned}
$$

III. DDPG-BASED SUM RATE EFFICIENT FL ALGORITHM

To address the problem, we reformulated the optimization challenge as a MDP and devised a DDPG-based reinforcement learning algorithm to effectively tackle it. This model is represented as a tuple (S, A, R), where S denotes the set of states of the environment, A denotes the set of possible actions, and R: S × A → R defines the reward function. This function specifies the reward the agent at server receives from the environment for taking action a in state s. We clarify the definitions of the state space, action space, reward, and penalty as follows:

State: At the outset of each time slot t, the agent observes the network environment, allowing it to ascertain the system state. This includes the location of UAV, the location of user, the bandwidth, and the model size.

$$s = \{\boldsymbol{q}, \boldsymbol{q_{n,t}}, \boldsymbol{b_t}, \boldsymbol{Z_{n,t}}\}, \forall n \in N, \forall t \in T$$

Action: The action taken in response to the current state s_t. This includes the optimal location of UAV, the transmit power, the CPU frequency, and the sub-channel association variable.

$$a = \{\boldsymbol{q'}, \boldsymbol{p_{n,t}}, \boldsymbol{\nabla_{n,t}}, \boldsymbol{f_{n,t}}\}, \forall n \in N, \forall t \in T$$

Reward: Since our goal is maximization of system sum rate, the decision reward function is

$$r_d = r_{n,t}$$

To satisfy the constraints, we set time penalty and sum rate penalty as

$$p_t = \begin{cases} 1, if\ t^{ct} \ge t^{to} \\ 0, otherwise \end{cases}$$

$$p_r = \begin{cases} 1, if\ r_{n,t} < r_Q \\ 0, otherwise \end{cases}$$

Therefore, the total decision reward function of the model at the time section t_e is formulated as

$$r = r_d - p_t - p_r$$

We introduce an energy-efficient FL algorithm based on DDPG as outlined in Algorithm 1, to effectively address our optimization problem. The core elements of the DDPG algorithm involve employing two primary deep neural networks (DNNs): a actor network and a critic network. We aim to determine the optimal policy, defined as

$$\mu^*(s|\theta^\mu) = argmax_{a \in A}\{Q(s_t, a_t)|\theta^Q\},$$

where Q is the value of chosen action a_t at state s_t. Thus, we find the optimal joint action a^* that maximinxes the expected cumulative Q-value Q^*. We achieve this solving the following Bellman equation:

$$Q^*(s_t, a_t)|\theta^Q = E\left[max_{a \in A}\left[r_t + \gamma Q^*(s_{t+1}, a_{t+1})|\theta^Q\right]\right].$$

This actor network takes the current state as input and generates the corresponding action as an output. The actor network and the critic network are updated using a policy gradient, and a loss function, respectively.

The DDPG-based sum rate efficient algorithm is as follow

Algorithm 1: DDPG-based Sum rate Efficient FL Algorithm

Initialize hyperparameters.
Initialize the actor network, $\mu(s|\theta^\mu)$, and the critic network, $Q(s, a|\theta^Q)$ with weights θ^μ and θ^Q.
Initialize the target actor network $\mu'(s|\theta^{\mu'})$ and the target critic network $Q'(s, a|\theta^{Q'})$ with weights $\theta^{\mu'} \leftarrow \theta^\mu$ and $\theta^{Q'} \leftarrow \theta^Q$.
Initialize a replay buffer $\mathcal{H}$.
for each episode $e = 1, 2, ..., e_{max}$ **do**
    Initialize system state s_1 and random action nose χ_0.
    for $t = 1, 2, , , , , T$ **do**.
        Observe s_t.
        Execute overall action a_t^*.
        Observe reward r_t and next state s_{t+1}.
        Store experience in the buffer $\mathcal{H}$ and uniform sample a batch of H to train the DNNs.
        Update the actor network and the critic network.
        Update the target actor network and the target critic network.
    end for
end for

IV. CONCLUSIONS

This paper addresses sum rate maximization in UAV-assisted FL systems using MC-NOMA. Initially, we jointly optimized UAV location, power allocation, CPU frequency, and sub-channel association variable allocation to maximize sum rate. Subsequently, we transformed this joint optimization problem into a MDP formulation and proposed a DDPG-based reinforcement learning algorithm to solve it.

ACKNOWLEDGMENT

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터사업의 연구결과와 2024 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행되었음”(IITP-2024-RS-2022-00156353, No. 2021-0-00493, 5G Massive 차세대 사이버공격 기만기술 개발)

References

[1] Ho, M. C., Tran, A. T., Lee, D., Paek, J., Noh, W., & Cho, S. (2023). A DDPG-based energy efficient federated learning algorithm with SWIPT and MC-NOMA. ICT Express.

[2] Truong, T. P., Dao, N. N., & Cho, S. (2022). HAMEC-RSMA: Enhanced aerial computing systems with rate splitting multiple access. IEEE Access, 10, 52398-52409.