Hierarchical QL for Optimal Resource Allocation and UAV Positioning in SAGIN with IAB

Kakyeom Jeon
School of Electronic and Electrical Engineering
Hankyong National University
Anseong, Republic of Korea
kuem0803@hknu.ac.kr

Howon Lee
School of Electronic and Electrical Engineering
Hankyong National University
Anseong, Republic of Korea
hwlee@hknu.ac.kr

Abstract— In this paper, we consider an environment where low-orbit satellites and unmanned aerial vehicles (UAVs) provide downlink communication services to ground devices in satellite-air-ground integrated networks (SAGINs). To provide seamless connectivity in the SAGIN by using limited frequency resources, we consider an integrated access and backhaul architecture and propose the hierarchical Q-Learning algorithm for optimal resource allocation and UAVs' position control considering the propagation delay difference. The proposed algorithm outperforms the various benchmark methods.

Keywords—SAGIN, IAB, hierarchical reinforcement learning, resource allocation, UAV positioning

1. Introduction

A 6G satellite-air-ground integrated network (SAGIN) with integrated access and backhaul (IAB) requires more flexible frequency resource utilization to support 3D network connectivity, resulting in severe co-tier and cross-tier interferences [1]. Hence, considering the interference problem and propagation delay difference in SAGIN, we propose the hierarchical Q-Learning (HQL) algorithm.

2. System Model

We consider a low earth orbit (LEO) satellite with $\boldsymbol{B}$ multiple beams and $\boldsymbol{U}$ unmanned aerial vehicles (UAVs). Each beam provides a downlink communication service to UAVs and $\boldsymbol{G}$ ground devices (GDs). The channel gains from beam $\boldsymbol{b}$ and UAV $\boldsymbol{u}$ to receiver $\boldsymbol{r} \in \{u, g\}$ are represented as follows [2][3]:

$$g_{b,r} = G_b G_r / PL_{b,r}. \tag{1}$$

$$g_{u,g} = \left(Pr_{u,g}^{LoS} \times L_{u,g}^{LoS} + Pr_{u,g}^{NLoS} \times L_{u,g}^{NLoS}\right)^{-1}. \tag{2}$$

Here, $\boldsymbol{G_b}$ and $\boldsymbol{G_r}$ are transmitter antenna gain and receiver antenna gain, respectively. $\boldsymbol{PL_{b,r}}$ is path loss between beam $\boldsymbol{b}$ and receiver $\boldsymbol{r}$. In equation (2), $\boldsymbol{Pr_{u,g}^{LoS}}$ and $\boldsymbol{Pr_{u,g}^{NLoS}}$ denotes line-of-sight (LoS) and Non-LoS (NLoS) probabilities, respectively, and $\boldsymbol{L_{u,g}^{LoS}}$ and $\boldsymbol{L_{u,g}^{NLoS}}$ are propagation losses of LoS and NLoS, respectively. The Signal-to-interference-plus-noise ratio (SINR) $\boldsymbol{\Gamma_{t,r}}$ of receiver $\boldsymbol{r}$ served by a transmitter $\boldsymbol{t} \in \{\boldsymbol{b}, \boldsymbol{u}\}$ is defined as the transmit power $\boldsymbol{P_t}$ and noise power $\boldsymbol{\sigma_r^2}$. Also, the achievable data rate $\boldsymbol{\vartheta_{t,r}}$ is defined by $\boldsymbol{\Gamma_{t,r}}$ with channel bandwidth $\boldsymbol{BW}$ and the number of links $\boldsymbol{l}$ as follows:

$$\Gamma_{t,r} = \frac{P_t g_{t,r}}{\Sigma_{b^*=1/t}^{B} P_{b^*} g_{b^*,r} + \Sigma_{u^*=1/t}^{U} P_{u^*} g_{u^*,r} + \sigma_r^2} \tag{3}$$

$$\vartheta_{t,r} = (BW/l) \times \log_2(1 + \Gamma_{t,r}) \tag{4}$$

The proposed HQL proposes a hierarchical framework to consider the propagation delay difference between LEO link and UAV link. The agents of outer-loop QL and inner-loop QL are a beam and a UAV, respectively. At time-step τ, the state of beam $\boldsymbol{b}$ includes channel status and transmit power strength. The action is channel and power adjustment; $\boldsymbol{s_b(\tau)} = [\boldsymbol{m_b}, \boldsymbol{P_b}]$, $\boldsymbol{a_b(\tau)} \in \{\pm\boldsymbol{\Delta m}, \pm\boldsymbol{\Delta P_b}, \boldsymbol{stay}\}$. The reward of $\boldsymbol{b}$ in the outer-loop QL is the sum-rate of network; $\boldsymbol{r_o(\tau)} = \boldsymbol{\Sigma_{b^*=1}^{B} \vartheta_{b^*,r}}$. In addition, the state of UAV $\boldsymbol{u}$ includes channel, power, and location information. The actions are channel and power adjustment and UAV's movement; $\boldsymbol{s_u(\tau)} = [\boldsymbol{m_u}, \boldsymbol{P_u}, \boldsymbol{x_u}, \boldsymbol{y_u}, \boldsymbol{z_u}]$, $\boldsymbol{a_u(\tau)} \in \{\pm\boldsymbol{\Delta m}, \pm\boldsymbol{\Delta P_u}, \pm\boldsymbol{\Delta x}, \pm\boldsymbol{\Delta y}, \pm\boldsymbol{\Delta z}, \boldsymbol{stay}\}$. Let $\boldsymbol{b_u}$ be the beam in which $\boldsymbol{u}$ resides, and if there are $\boldsymbol{u_b}$ UAVs in $\boldsymbol{b_u}$, the reward of $\boldsymbol{u}$ in the inner-loop QL is the sum-rate of all GDs in $\boldsymbol{b_u}$; $\boldsymbol{r_i(\tau)} = \boldsymbol{\Sigma_{t^*=t_{b_u}} \vartheta_{t^*,g}}$, $\boldsymbol{t_{b_u}} \in \{\boldsymbol{b_u}, \boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_b}\}$.

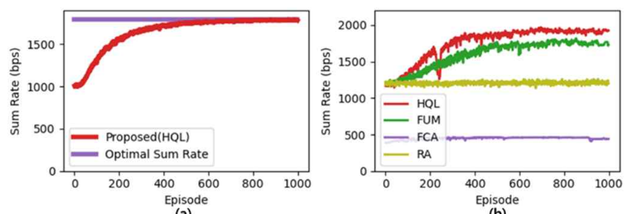

Fig.1. Sum rate vs. episode, (a) HQL and optimal method and (b) FUM, FCA, RA and HQL.

3. Simulation Results and Conclusion

The altitude of the LEO is 300 km and ground devices randomly distributed within the beam coverage. Additionally, the random-walk model is applied for GD's mobility [4][5]. Fig.1(a) illustrates that the proposed HQL algorithm converges to the optimal value obtained by an exhaustive search algorithm under 1 beam-3 UAVs-29 GDs. Also, Fig.1(b) shows the performance comparison of HQL with fixed UAV movement (FUM), fixed channel allocation (FCA), and random action (RA) under 2 beams-6 UAVs-58 GDs. The HQL method, which optimally controls the frequency channel, transmit power, and even the 3D location of the UAVs, outperforms the existing benchmark methods.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2022-0-00704, Development of 3D-NET Core Technology for High-Mobility Vehicular Service).

Reference

[1] H. Lee et al., "Towards 6G hyper-connectivity: Vision, challenges, and key enabling technologies," in Journal of Communications and Networks, vol. 25, no. 3, pp. 344-354, Jun. 2023, doi: 10.23919/JCN.2023.000006.

[2] P. Gu et al., "Dynamic Cooperative Spectrum Sharing in a Multi-Beam LEO-GEO Co-Existing Satellite System," in IEEE Transactions on Wireless Communications, vol. 21, no. 2, pp. 1170-1182, Feb. 2022.

[3] A. Al-Hourani et al., "Optimal LAP altitude for maximum coverage," IEEE Wireless Communications Letters, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[4] S. Lee et al., "Multiagent Q-Learning-Based Multi-UAV Wireless Networks for Maximizing Energy Efficiency: Deployment and Power Control Strategy Design," in IEEE Internet of Things Journal, vol. 9, no. 9, pp. 6434-6442, May. 2022.

[5] S. Lim et al., "Optimal Tethered-UAV Deployment in A2G Communication Networks: Multi-Agent Q-Learning Approach," in IEEE Internet of Things Journal, vol. 9, no. 19, pp. 18539-18549, Oct. 2022.