DRL-Based Satellite Network Slice Planning and Handover in the Korean Peninsula Scenarios

Seonghoon Kim *, Taeyeoun Kim †, Jeongho Kwak †, and Jihwan P. Choi *
*Dept. of Aerospace Engineering, Korea Advanced Institute of Science & Technology, Korea
†Dept. of Electrical Engineering and Computer Science, Daegu Gyeongbuk Institute of Science & Technology, Korea
{seonghoon, jhch}@kaist.ac.kr, {kim77ty, jeongho.kwak}@dgist.ac.kr

***Abstract*—This paper introduces a novel approach using deep reinforcement learning (DRL) to enhance network slicing planning and handovers in satellite networks. We propose a proactive handover trigger based on remaining service time and employ the deep deterministic policy gradient (DDPG) algorithm to maximize the utility of virtual networks (VNs). Focusing on the Korean Peninsula, we simulate a low earth orbit (LEO) satellite network based on Starlink satellite specifications and demonstrate the superiority of our intelligent network management technique compared to baseline methods, particularly in terms of latency performance and the number of handovers.**

***Index Terms*—Satellite network, satellite network slicing planning, handover, deep reinforcement learning.**

I. Introduction

Low Earth orbit (LEO) satellite networks, based on inter-satellite links and satellite onboard computing, are expected to play a pivotal role as a core feature of 6G mobile communications by offering global coverage. In the context of managing the satellite constellations, the concept of satellite network slicing has emerged [1]. This approach virtually creates independent virtual networks (VNs) within a shared physical network infrastructure and enables the provisioning of dedicated services to each slice customer, heralding its potential to facilitate a range of user-tailored services with diverse service requirements [2]. However, the end-to-end connection and maintenance performance of satellite network slice planning (SNSP) can manifest diversely depending on the employed VN embedding (VNE) schemes [1]. Thus, the effects of the dynamic network topology due to the mobility of LEO satellites should be addressed to ensure stable service provision.

In this paper, we introduce a novel approach based on deep reinforcement learning (DRL) to choose adequate schemes of SNSP and determine the handover timing. Additionally, this study assumes slice customers making VN requests within a localized scope of the Korean Peninsula, while the satellite network is simulated based on the specifications of the Starlink satellite constellation.

II. DRL-Based SNSP and Handover

In this section, the SNSP process of reserving the VN resource for each slice customer, i.e. VNE, and handover process of VNs are introduced based on DRL. Instead of selecting access nodes among the visible satellites for node embedding and searching end-to-end paths for link embedding, the DRL agent that manages satellite networks can select an VNE method among baseline schemes, which is illustrated in fig. 1.

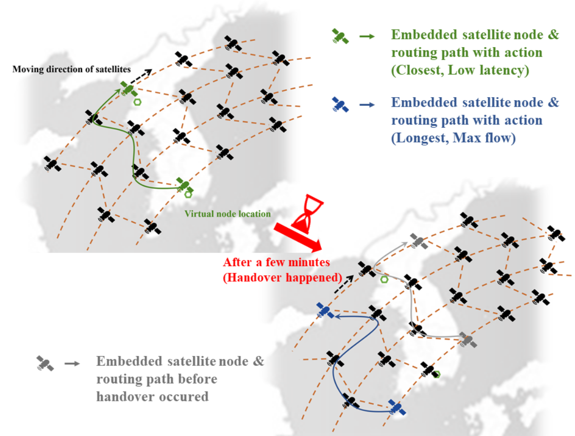

Fig. 1. VNE methods around the Korean Peninsula.

Among the node embedding methods, one method, referred to as (Closest), chooses the nearest visible satellite as the virtual node, resulting in the shortest one-hop propagation delay. Another approach, known as (Longest), selects the satellite with the longest available service time. Additionally, the DRL agent considers two routing strategies: (Max Flow), which focuses on network capacity maximization, and (Low Latency), which emphasizes finding the shortest path. Subsequent to the determination of each embedding method by the DRL agent, it is possible to evaluate the service available time and the end-to-end latency. Through these metrics, the reward r_i for the i-th VNE is defined as

$$r_i = (1 - \alpha_i) \cdot \frac{\log(1 + t_i)}{\mathscr{L}_i}, \tag{1}$$

where $t_i, \mathscr{L}_i$ represent the service available time and the end-to-end latency from action a_i of the i-th VNE, respectively. Lastly, instead of maintaining the service until the access node exits the line of sight, we define a handover trigger α_i to proactively initiate handovers through DRL. The handover trigger employs an algorithm that enforces handovers on the VNs when there is a remaining service available time of $\alpha_i \cdot t_i$. Therefore, the reward linearly decreases as considering the proactive handover.

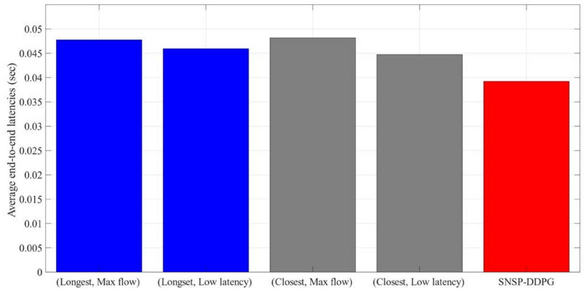

Fig. 2. Comparison of end-to-end latency performance for each method.

Because using all the network states acquired can be problematic for DRL, the state is defined as simplified observable features to implement efficient SNSP, and is expressed as follows:

$$s_i = (t_{i,C}, t_{i,L}, \mathscr{L}_{i,CMF}, \mathscr{L}_{i,LMF}, \mathscr{L}_{i,CLL}, \mathscr{L}_{i,LLL}), \quad (2)$$

where $t_{i,C}$ is the service available time with the closest node based embedding methods that is evaluated until the embedded node is out of the maximum accessible distance, while $t_{i,L}$ is with the longest node based embedding methods. $\mathscr{L}_{i,CMF}, \mathscr{L}_{i,LMF}, \mathscr{L}_{i,CLL}$ and $\mathscr{L}_{i,LLL}$ represent the end-to-end latency regarding the embedding methods (Closest, Max Flow), (Longest, Max Flow), (Closest, Low Latency) and (Longest, Low Latency), respectively. Based on the given states, the DRL agent determines suitable actions for SNSP and handover, which are expressed as

$$a_i = (\mathrm{Pr}_{i,\text{node}}, \mathrm{Pr}_{i,\text{link}}, \alpha_i), \quad (3)$$

where $\mathrm{Pr}_{i,\text{node}}$ and $\mathrm{Pr}_{i,\text{link}}$ respectively denote the parameters of Bernoulli distributions for selecting closest node embedding and Max flow link embedding for the i-th VNE. The actual embedding method is chosen through a coin toss. Thus, after the embedding technique is stochastically determined, the reward (1) for the i-th VNE is computed based on service available time, t_i, and end-to-end performance, $\mathscr{L}_i$, of the selected embedding method in (2).

The aforementioned SNSP and handover problems require making continuous-valued actions and the deep deterministic policy gradient (DDPG) algorithm [3], which is applicable to such problem formulations, can be used. In the upcoming section, we present the results of training neural network parameters using the SNSP-DDPG algorithm to maximize the long-term rewards of the satellite network around the Korean Peninsula.

III. Simulation results

Fig. 2 depicts average end-to-end latency for each method, while the blue and gray bars represent longest node based embeddings and closest node based embeddings, respectively. Although (Low Latency) methods demonstrate favorable latency performance in each benchmark case, SNSP-DDPG outperforms other methods in terms of latency performance by selecting appropriate methods, indicated by the red bar. Fig. 3 shows the numbers of handovers. The number of handovers is solely influenced by the node embedding methods and (Closest) methods exhibit relatively higher handover counts. Meanwhile, SNSP-DDPG strikes a balance in terms of the handover count by intelligently selecting two node embeddings for each VN.

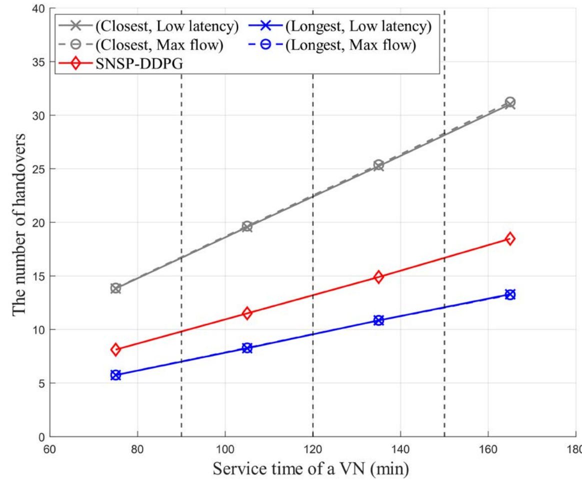

Fig. 3. Comparison of the number of handovers.

IV. Conclusion

In this paper, we have introduced a intelligent method based on DRL for effective SNSP and handover decisions in the Korean Peninsula scenarios. The simulation results have showcased that our approach provides a trade-off in terms of the handover count across the node embedding methods, while excelling in latency performance against all methods. The superiority of our intelligent network management technique over baseline methods affirms its potential for advancing satellite network services.

Acknowledgment

This work was supported by a grant from the Institute of Information & Communications Technology Planning & Evaluation (IITP) funded by the Korean government, Ministry of Science and ICT (No. 2018-0-01658, Key Technologies Development for Next-Generation Satellites). This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00704).

References

[1] T. Kim, J. Kwak, and J. P. Choi, "Satellite network slice planning: Architecture, performance analysis, and open issues," *IEEE Vehicular Technology Magazine*, vol. 18, no. 2, pp. 29 – 38, 2023.

[2] A. Fischer, J. F. Botero, M. T. Beck, H. De Meer, and X. Hesselbach, "Virtual network embedding: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1888–1906, 2013.

[3] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.