DQN-based Handover Decision Algorithm for mmWave Networks

Yekaterina Kim, Igbafe Orikumhi, Nakyung Lee, and Sunwoo Kim
Department of Electronic Engineering, Hanyang University, South Korea
{kattykimchi, oigbafe, nagyeong2379, remero}@hanyang.ac.kr

Abstract

This paper proposes a Deep Q-Network-based handover approach in dense urban millimeter-wave (mmWave) networks. While mmWave systems suffer from frequent and often unnecessary handovers due to sudden signal blockages, prior studies often neglect realistic deployment scenarios. To address this, we utilize 3D ray-tracing-based signal strength measurements to capture the complex propagation characteristics of urban mmWave environments. The proposed method effectively reduces redundant handovers and enhances network stability. Simulation results show that it significantly outperforms conventional handover mechanisms in terms of per-user handover rate.

***Index Terms*—Deep Q-network, millimeter wave, handover, urban dense environment, ray tracing**

I. Introduction

The growing densification of networks in urban environments, coupled with the deployment of millimeter-wave (mmWave) technologies in 5G and beyond, has introduced significant challenges related to handover performance. These challenges include frequent cell switching, elevated handover failure rates, increased latency due to handover interruptions, and the ping-pong effect [1], [2]. To mitigate these issues, various approaches have been proposed. However, many of these solutions rely on simplified statistical models, limiting their real-world applicability and leading to discrepancies between simulation results and deployment expectations. With machine learning emerging as a powerful tool for solving complex decision-making problems, its integration into handover management has become increasingly important. In this context, we propose a Deep Q-Network (DQN)-based approach to reduce handover rates in a realistic 3D urban environment modeled after the mmWave Urban Micro Street Canyon scenario defined in [3], using ray-tracing-based signal measurements.

II. System model

We consider a mmWave communication network operating in an outdoor 5G small cell environment [3], as shown in Fig.1. Orthogonal frequency division multiplexing (OFDM) is used in the system on the carrier frequency 28 GHz over 100 MHz bandwidth. For the sake of simplicity, single-input single-output system is considered in this study; however, the proposed approach is not limited to such scenarios. The downlink received signal at the k-th subcarrier can be written as

$$Y_k = H_k X_k + N_k, \tag{1}$$

where $N_k \sim \mathcal{CN}(0, \sigma^2)$ denotes the additive white Gaussian noise, X_k is a symbol transmitted with a power budget P_t under the constraint of $\mathbb{E}[|X_k|^2] = P_t$, and $H_k \in \mathbb{C}^{1\times 1}$ is the channel between the BS and the UE including path loss, small scale fading, and shadowing. The signal's level is evaluated based on the channel state information reference signal received power (CSI-RSRP) measurements [4]. The average received signal power P_r on the k-th subcarrier is given by

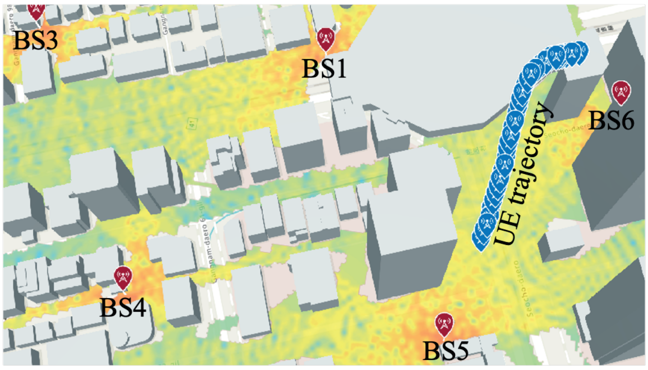

Fig. 1. mmWave outdoor Urban Micro Street Canyon deployment scenario with a UE mobility trajectory

$$P_{r,k} = \left|H_k\right|^2 P_t. \tag{2}$$

III. DQN-based handover algorithm

We propose a DQN-based handover decision algorithm, that minimizes unnecessary handovers in dense urban mmWave environment, where frequent line-of-sight obstructions and severe signal attenuation significantly impact link stability. The proposed DQN method aims to derive the optimal policy π that maximizes the cumulative future reward, following Bellman's optimization equation based on the current observation:

$$\mathcal{Q}^*(\mathbf{s}_t, a_t) = \max_{\pi} \mathcal{Q}_\pi(\mathbf{s}_t, a_t),$$

where $\mathbf{s}_t$ and a_t denote the state and action at time t, respectively. The largest value of $\mathcal{Q}^*(\mathbf{s}_t, a_t)$ defines the optimal action, so that the optimization problem can be written as

$$\pi^* = \arg\max_{a_t \in A} \mathcal{Q}^*(\mathbf{s}_t, a_t), \forall \mathbf{s}_t \in S. \tag{3}$$

The state $\mathbf{s}_t \in S$ observed by the agent at time t is the set of the measured RSRP values s_t^n corresponding to the measurement from the n-th BS. The state is given by

$$\mathbf{s}_t = \{s_t^1, s_t^2, \ldots, s_t^N\}, \tag{4}$$

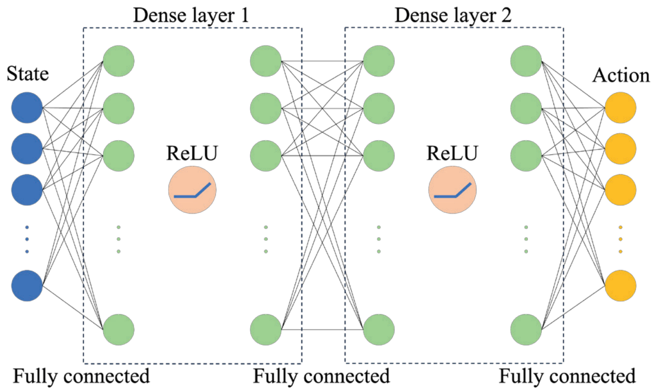

Fig. 2. Q-network for optimal handover decision

where N denotes the total number of BSs deployed.

The Q-network adopted in this work is shown in Fig. 2. It consists of two fully connected dense layers with ReLU activation function each. The output is the set of Q-values with respect to all possible actions a_t. The reward function guides the algorithm toward minimizing handovers by penalizing unnecessary ones and rewarding actions that prevent handover failures.

IV. Simulation Results

The simulation setup is developed using MATLAB R2024a. To validate the proposed approach for stochastic attenuation of the mmWave signal in realistic environment, we perform ray tracing in 3D environment (3D map source: OpenStreetMap) of a dense urban area of the Gangnam district in Seoul, South Korea. The 3D map reflects the real landscape of the urban area including buildings, roads, and BS sites deployment. The BSs are deployed using urban micro small cells with maximum inter-site distance 200 m [5] operating at 28 GHz [3].

The DQN agent is implemented in Python using TensorFlow and Keras, with two fully connected hidden layers of size 64 each and ReLU as the activation function. The learning progress of the proposed algorithm is shown in Fig. 3. The rising trend from episodes 0 to 75 indicates learning, while the plateau thereafter suggests policy convergence. The increase in cumulative rewards from -3000 to over 3500 reflects the development of a more optimal policy.

Fig. 4 shows the performance improvement of the proposed method over the conventional A3-based approach in terms of handover rate. The pedestrian experienced a higher handover rate than the 10 and 30 km/h vehicles due to a trajectory through narrow alleys with severe signal blockage. In contrast, vehicles followed road-based paths where handover rates were influenced more by speed than trajectory. Consequently, the low-speed vehicles (10 and 30 km/h) had lower handover rates compared to the pedestrian and the high-speed (60 km/h) vehicular user.

V. Conclusion

In this paper, we proposed a DQN-based approach for handover decision in mmWave urban micro network deployment for various user mobility scenarios. For more precise and fair

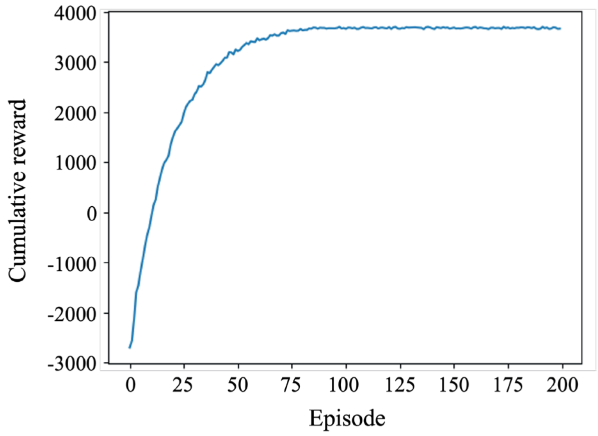

Fig. 3. Training performance: cumulative reward per training episode

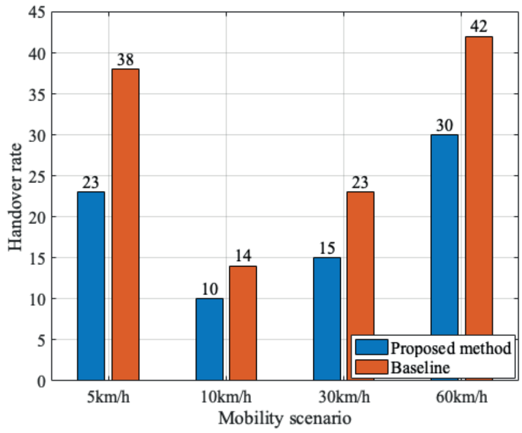

Fig. 4. Handover rate per user across four mobility scenarios: pedestrian (5 km/h) and three vehicles (10, 30, 60 km/h)

evaluation, we considered realistic 3D environment using ray tracing measurements. The simulation results showed that the proposed method can improve mobility latency by reduction of unnecessary handovers.

Acknowledgment

This work was supported by Institute of information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No. RS-2024-00397216, Development of the Upper-mid Band Extreme massive MIMO(E-MIMO))

References

[1] D. Pjanić, A. Sopasakis, A. Reial and F. Tufvesson, "Early-Scheduled Handover Preparation in 5G NR Millimeter-Wave Systems," *IEEE Open J. Commun. Soc.*, vol. 5, no. 10, pp. 6959–6971, 2024.

[2] H. -S. Park, Y. Lee, T. -J. Kim, B. -C. Kim and J. -Y. Lee, "Handover Mechanism in NR for Ultra-Reliable Low-Latency Communications," *IEEE Netw.*, vol. 32, no. 2, pp. 41–47, 2018.

[3] "Study on channel model for frequencies from 0.5 to 100 GHz (Release 18) 3GPP TR38.901 v18.0.0 (2024-03)," *3GPP*, 2024.

[4] "NR; Requirements for support of radio resource management (Release 18) 3GPP TS38.133 v18.8.0 (2024-12)," *3GPP*, 2024.

[5] "NR; NR and NG-RAN Overall Description; Stage 2 (Release 18) 3GPP TS38.300 v18.4.0 (2024-12)," *3GPP*, 2024.