

# 밀리미터파 통신 심층 Q-네트워크 기반 빔 폭 제어 및 도래각 추적 알고리즘

전종현, 박현우, 강정완, 정현진, 김선우  
한양대학교 융합전자공학부

{danny0418, stark95, rkdwdhks77, hyeonjingo, remero}@hanyang.ac.kr

## Deep Q-network based angle tracking with beamwidth adaptation for mmWave communications

Jonghyun Jeon, Hyunwoo Park, Jeongwan Kang, Hyeonjin Chung, Sunwoo Kim  
Department of Electronic Engineering, Hanyang University

### 요약

본 논문에서는 deep Q-network (DQN) 기반 협동 빔 폭 제어 및 도래각 추적 알고리즘을 제안한다. 제안하는 알고리즘은 채널 모델과 단말의 동적 시나리오에 대한 사전 정보 없이 수신신호만을 이용하여 도래각을 추적하고 활성화 안테나 수를 최적화한다. DQN 모델을 사용하면 수신신호에 대한 복잡한 신호 처리가 불필요하므로 계산 복잡도가 대폭 감소한다. 빔 폭 제어 여부에 따른 도래각 추적 성능을 비교/분석하였고, 이동성이 큰 시나리오에서 빔 폭 제어를 통하여 도래각 추적 성능을 향상시킬 수 있음을 확인하였다.

### I. 서론

송신 또는 수신신호를 지정 방향으로 초점을 맞추는 적절한 빔포밍 전략은 5G 밀리미터파 통신을 위한 핵심 기술 전략 중 하나이다 [1]. 빔 추적 기술은 도래각 (angle of arrival, AoA)과 발사각 (angle of departure, AoD)의 변화에 맞추어 빔 방향을 조정하여 오정렬 확률을 낮춘다. 하지만 단말 (user equipment, UE)의 이동성이 높은 경우, 신호가 빔의 주엽을 이탈할 가능성이 증가하여 통신 장애가 발생한다. 따라서 배열 안테나의 부분 활성화를 통하여 빔 폭을 실시간으로 제어하는 기술이 제안되었다 [2]. 본 논문에서는 deep Q-network (DQN) [3] 기반 협동 빔 폭 제어 및 도래각 추적 알고리즘을 제안한다. DQN은 대표적인 심층 강화학습 기술 중 하나이며, Q-테이블을 네트워크화해 연속적 도메인에서 추정 알고리즘으로 활용된다. 시뮬레이션을 통하여 제안하는 알고리즘과 빔 폭 제어를 포함하지 않는 DQN 기반 빔 추적 알고리즘 [4]의 성능을 비교 및 분석하였다.

### II. 시스템 모델

본 논문에서는 배열 안테나로 구성된 기지국이 단일 안테나 단말과 업 링크 통신하는 환경을 고려한다. 단순화를 위하여 기지국 (base station, BS)과 단말은 2차원 평면에 존재하며 기지국은 원점에 위치한다. 하나의 에피소드에는  $K$  개의 타임 슬롯이 존재하며 단말은 하나의 타임 슬롯 당  $L$  개의 파일럿 심볼을 전송한다.  $k$  번째 타임 슬롯에서 단말이 송신하는 파일럿 심볼은  $\mathbf{b}_k = [b_{k,1}, \dots, b_{k,l}, \dots, b_{k,L}] \in C^{1 \times L}$ 이며  $|b_{k,l}| = 1$  이다. 기지국이 수신하는 복소 신호  $z_{k,l}$ 은 다음과 같이 나타낼 수 있다.

$$z_{k,l} = \mathbf{w}_{k,l}^H \mathbf{h}_{k,l} \mathbf{b}_{k,l} + n_{k,l}, \quad (1)$$

$\mathbf{w}_{k,l}$ 는 수신 빔포밍 벡터이며 빔포밍 각도  $\psi_{k,l}$ 와 활성화 안테나 수  $N_{k,l}$ 에 관한 식으로 다음과 같이 나타낼 수 있다.

$$\mathbf{w}_{k,l} = \frac{1}{\sqrt{N_{k,l}}} \mathbf{a}(\psi_{k,l}, N_{k,l}) \in C^{N_{k,l} \times 1}, \quad (2)$$

$\mathbf{a}(\theta, N)$ 는 반 파장 등간격으로 배치된 배열 안테나의 응답 벡터이며 다음과 같다.

$$\mathbf{a}(\theta, N) = [1, e^{-j\pi \cos \theta}, \dots, e^{-j(N-1)\pi \cos \theta}]^T, \quad (3)$$

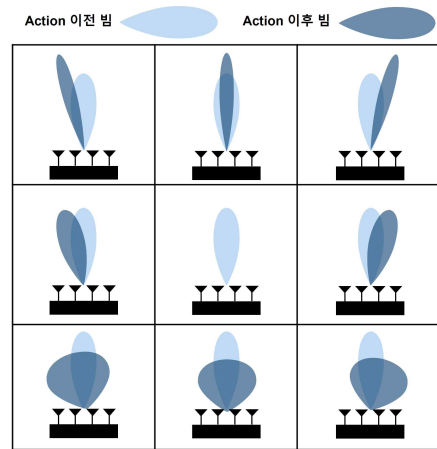


그림 1 제안하는 DQN 모델의 action

(1)에서  $\mathbf{h}_{k,l} \in C^{N_{k,l} \times 1}$ 는 기지국과 단말 사이의 채널 벡터이고  $n_{k,l}$ 은 백색 가우시안 잡음이며  $CN(0, \sigma_n^2)$ 의 분포를 따른다. 송신 signal-to-noise ratio (SNR)은 다음과 같이 정의한다.

$$\text{SNR}[\text{dB}] = 20 \log_{10} \left| \frac{b_{k,l}}{\sigma_n} \right| = 20 \log_{10} \frac{1}{\sigma_n}. \quad (4)$$

### III. DQN 기반 협동 빔 폭 제어 및 도래각 추적 알고리즘

DQN의 에이전트 (agent)는 매 타임 슬롯 state  $\mathbf{s}_k$ 를 관측하고 reward  $r_k$ 의 누적값을 최대화하는 action  $\mathbf{a}_k$ 를 선택하도록 학습한다. 성공적인 빔 폭 제어 및 도래각 추적을 위하여 state, action, reward의 디자인이 중요하다. 제안하는 알고리즘은 실시간으로 단말의 이동성에 의하여 발생하는 빔 오정렬에 대응하여 적절한  $\mathbf{a}_k$ 를 선택한다. 그림 1은 제안하는 DQN 모델의 action을 나타낸다. 제안하는 알고리즘은 단말의 이동성에 따라서 실시간으로 도래각 추적과 빔 폭 제어를 동시에 또는 각각 수행한다. [4]는 활성화 안테나 수  $N$ 이 고정이다. 하지만 제안하는 DQN 모델은 매 타임 슬롯마다 활성화 안테나 수를 조정하여 빔 폭을 제어한다. State  $\mathbf{s}_k$ 는 다음과 같다.

$$\mathbf{s}_k = \begin{bmatrix} \text{Real}(s_{k,1:L}^*); & \text{Imag}(s_{k,1:L}^*); \\ \text{Real}(s_{k-1,L+1:L}^*); & \text{Imag}(s_{k-1,L+1:L}^*); \hat{N}_{k-1} \end{bmatrix}, \quad (5)$$

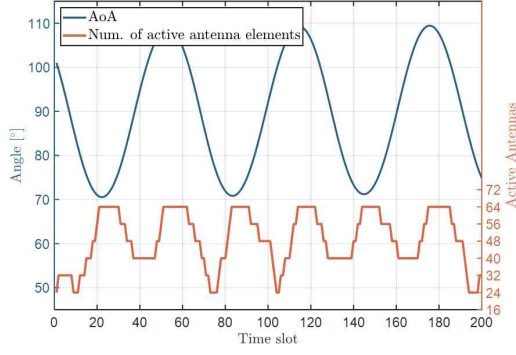


그림 2 타임 슬롯 당 도래각과 기지국의 활성화 안테나의 수

표 1 시뮬레이션 파라미터

파라미터	값	파라미터	값
할인율	0.9	학습률	0.000125
메모리 크기	4,000	타겟 네트워크 업데이트 간격	1,000
훈련 간격	4	훈련 에피소드 수	20,000
미니 배치 크기	32	-	-

$$s_{k,1:L^*} = \frac{1}{L^*} \sum_{l=1}^{L^*} z_{k,l} b_{k,l}^H, \quad (6)$$

$$s_{k-1,L^*+1:L} = \frac{1}{L-L^*} \sum_{l=L^*+1}^L z_{k-1,l} b_{k-1,l}^H, \quad (7)$$

이때  $L^*$ 는 action 이전의 파일럿 심볼의 수이다.  $\hat{N}_k$ 는 알고리즘이 추정한 최적 활성화 안테나 수를 나타낸다. Action  $\mathbf{a}_k = [a_{N,k}, a_{\theta,k}]$ 는 기지국의 활성화 안테나 수와 수신 빔 각도를 조정한다.  $a_{N,k}, a_{\theta,k}$ 는 각각  $\{-1, 0, 1\}$ 의 원소 중 하나이며 다음과 같은 동작을 수행한다.

$$\hat{N}_k = \hat{N}_{k-1} + \Delta_N a_{N,k}, \quad (8)$$

$$\hat{\theta}_k = \hat{\theta}_{k-1} + (\Delta_{\theta} / \hat{N}_{k-1}) a_{\theta,k}, \quad (9)$$

$\Delta_N, \Delta_{\theta}$ 는 각각 활성화 안테나 수와 도래각 추적을 위한 조정 크기이며  $\hat{\theta}_k$ 는 추정 도래각이다. Reward  $r_k$ 는 다음과 같으며 action 이후 수신 RSS(Received signal strength)가 증가할 때 큰 값을 가진다.

$$r_k = \begin{cases} 2, & \text{if } \frac{L^* \sum_{l=L^*+1}^L |z_{k,l}|^2}{(L-L^*) \sum_{l=1}^{L^*} |z_{k,l}|^2} > 1.01 \\ 1, & \text{if } 0.99 \leq \frac{L^* \sum_{l=L^*+1}^L |z_{k,l}|^2}{(L-L^*) \sum_{l=1}^{L^*} |z_{k,l}|^2} \leq 1.01 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

#### IV. 실험 결과 및 분석

시뮬레이션 초기 상황에는 빔이 정렬되어있는 상태를 가정한다. 표 1은 DQN 모델의 네트워크 학습에 사용한 파라미터이다. 에피소드의 최초 기지국 활성화 안테나 수는 24개이며 송신 SNR은 10 dB이다.  $\Delta_N, \Delta_{\theta}$ 는 각각 8개,  $70^\circ$ 이고  $K, L, L^*$ 은 각각 200, 2, 1개이다. 한 타임 슬롯의 길이는 10 ms이다. 그림 2는 한 타임 슬롯 당 최대 도래각 변화량  $\delta_{\theta, \max}$ 가  $2^\circ$ 인 환경에서 타임 슬롯 당 실제 도래각과 기지국의 활성화 안테나 수를 보여준다. 타임 슬롯의 길이를 고려했을 때 극도로 빠른 동적

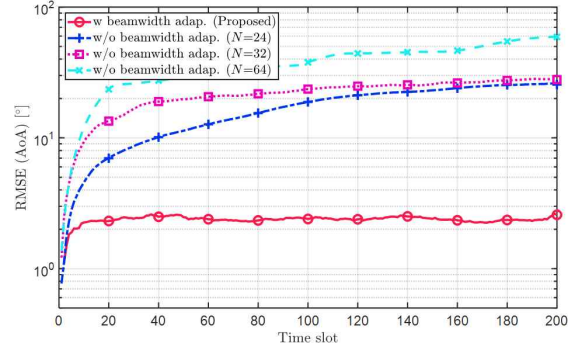


그림 3 타임 슬롯 당 알고리즘의 RMSE 비교

시나리오임을 알 수 있다. 도래각의 변화량이 증가하면 작동 안테나의 수를 줄이며 빔 폭을 넓히는 것을 확인할 수 있다. 추가로 도래각의 변화량이 감소할 때 빔 폭을 좁히면서 빔 이득을 높이는 action을 수행한다. 따라서, 제안하는 DQN 모델은 빔 오정렬 정도에 따라 적정한 빔 폭 제어를 통하여 도래각 추적 성능을 향상시킨다.

그림 3에서는  $\delta_{\theta, \max}$ 가  $2^\circ$ 인 환경에서 제안하는 알고리즘과 [4]의 성능 비교 및 분석을 위하여 각도 추정 root mean square error (RMSE)를 확인한다. RMSE 계산을 위하여 5,000개의 테스트 에피소드를 사용한다.  $N$ 이 고정된 경우 ( $N=24, 32, 64$ ), 시간에 따라 RMSE가 증가하며  $N$ 이 클수록 낮은 성능을 보인다. 또한  $N$ 이 작아지면 성능이 개선되지만, 여전히 약  $10^\circ$  정도의 RMSE를 보인다. 제안하는 알고리즘은 빔 폭 제어를 수행하여 약  $2.5^\circ$ 의 RMSE 가장 높은 성능 보여준다. 따라서 제안 알고리즘은 이동성이 큰 환경에서도 좋은 성능을 보이는 것을 확인할 수 있다.

#### V. 결론

본 논문에서는 DQN 기반 협동 빔 폭 제어 및 도래각 추적 알고리즘을 제안하고 빔 폭 제어 유무에 따른 성능을 비교/분석하였다. 제안하는 알고리즘에서 기지국은 수신신호만을 이용하여 도래각을 추적하고 단말의 이동성을 고려하여 활성화 안테나 수를 최적화한다. 시뮬레이션을 통하여 이동성이 큰 시나리오에서 빔 폭 제어의 유용성을 검증하였다. 향후 학습 시 발생하는 action 겹침 현상으로 인한 모델의 성능 저하를 해결하기 위한 DQN 구조를 연구할 계획이다.

#### ACKNOWLEDGMENT

이 논문은 2023년도 4단계 두뇌한국21 사업(4단계 BK21 사업)에 의하여 지원되었음

#### 참고 문헌

- [1] S. Jayaprakasam, X. Ma, J. W. Choi and S. Kim, "Robust Beam-Tracking for mmWave Mobile Communications," *IEEE Commun. Lett.* vol. 21, no. 12, pp. 2654-2657, Dec. 2017.
- [2] H. Chung, J. Kang, H. Kim, Y. M. Park and S. Kim, "Adaptive Beamwidth Control for mmWave Beam Tracking," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 137-141, Jan. 2021.
- [3] V. Mnih et al., "Human-level Control through Deep Reinforcement Learning," *Nature*, vol. 518, pp. 529-533, Feb. 2015.
- [4] H. Park, J. Kang, S. Lee, J. W. Choi, and S. Kim, "Deep Q-network based beam tracking for mobile millimeter-wave communications," *IEEE Trans. Wireless Commun.*, pp. 1 - 1, 2022.