

밀리미터파 다중 사용자 MISO 시스템에서의 빔포밍을 위한 심층 강화 학습

이영록, 이주형, 고영채

고려대학교

{spiritica, leejuhung, koyc}@korea.ac.kr

Deep Reinforcement Learning for Beamforming in mmWave MU-MISO Systems

Yeongrok Lee, Ju-Hyung Lee, Young-Chai Ko

Korea Univ.

요약

본 논문에서는 밀리미터파 대역에서 다중 사용자 MISO 통신 시스템에서의 빔포밍을 위한 심층 강화 학습 알고리즘과 그 성능을 소개한다. 특히, 심층 강화 학습 기법 중 Deep Q-Network 방식을 사용하여 사용자들의 sum-rate를 최대화하는 최적의 빔포밍 벡터들을 얻어내는 과정에서, 현재의 채널 정보와 그 부속 정보들을 상태로, 빔포밍 벡터를 행동으로, sum-rate를 보상으로 하는 기법을 제안한다. 시뮬레이션을 통해 코드북 기반 빔포밍에서 최적의 성능을 가지지만 높은 시간 복잡도를 가지는 완전 탐색 알고리즘과의 성능을 비교 및 분석한다.

I. 서론

기술이 발달함에 따라 높은 전송 속도를 요구하는 사용자와 시스템이 증가하고 있다. 이러한 환경에서 밀리미터파 대역에서의 통신은 넓은 대역폭을 활용하여 높은 전송 속도를 얻을 수 있다는 장점이 있어 차세대 기술로 각광받고 있다. 하지만 기존에 사용하던 6 GHz 이하 대역에서의 통신과는 다르게 밀리미터파 대역에서의 통신은 주파수 특성으로 인해 높은 경로 손실과 감쇠, 낮은 투과율과 회절로 인해 성능이 크게 감소할 수 있다.[1].

밀리미터파 대역에서는 이를 해결하기 위해 일반적으로 다수의 안테나 요소가 장착된 배열 안테나를 사용해 성능을 높이는 빔포밍 방식을 사용한다. 특히, 파장이 짧고 강한 지향성을 가지고 있기 때문에 안테나 크기와 그 간격을 획기적으로 줄일 수 있어 다중 안테나 시스템을 구성하기도 유리하다. 또, 다중 사용자를 지원하는 시스템에서는 사용자 간의 간섭 또한 통신 성능에 부정적인 영향을 끼치기 때문에, 성능 향상을 위해서는 반드시 빔포밍 방식을 적용해야 한다.

본 논문에서는 밀리미터파 대역에서 배열 안테나를 통해 빔포밍 기법을 사용하는 기지국이 단일 안테나를 사용하는 다수의 사용자를 커버하는 MU-MISO 시스템을 가정한다. 이 시스템에서 높은 sum-rate를 얻을 수 있는 빔포밍 벡터를 심층 강화 학습 기법을 통해 찾아내는 알고리즘을 제안하고 그 성능을 비교 분석한다.

II. 본론

1. 시스템 모델

본 논문에서는 그림 1과 같이 기지국이 M 개의 안테나 요소로 이루어진 Uniform Linear Array (ULA) 배열 안테나를 통해 빔포밍 기법을 사용하고, 단일 안테나를 사용하는 K 개의 단말기가 존재하는 시스템을 가정한다. 단말기 k 가 받는 수신 신호 y_k 는 (1)과 같이 표현할 수 있다.

$$y_k = \mathbf{h}_k^H \mathbf{w}_k s_k + \sum_{\substack{i=1 \\ i \neq k}}^K \mathbf{h}_k^H \mathbf{w}_i s_i + n_k \quad (1)$$

이 때, \mathbf{h}_k 는 $M \times 1$ 벡터로 기지국으로부터 단말기 k 로 전달되는 채널이고 \mathbf{w}_k 는 $M \times 1$ 벡터로 기지국에서 단말기 k 를 대상으로 한 빔포밍 벡터, s_k 는 스칼라 값으로 단말기 k 로 전송하고자 하는 신호이며 n_k 는 평균이 0이고 분산이 σ^2 인 정규 분포를 따르는 단말기 k 로의 가산 백색 가우시안 잡음이다.

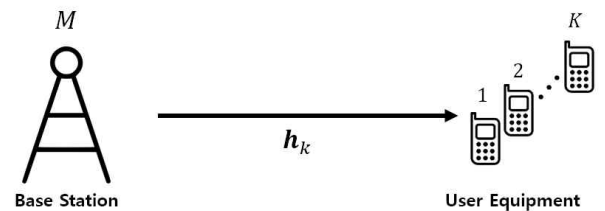


그림 1. 밀리미터파 대역에서의 MU-MISO 시스템 모델

(1)에서 우항의 첫 부분은 단말기 k 로 전송해야 하는 원하는 신호이고, 두 번째 부분은 다른 단말기로 가야하는 간섭 신호이다. 이 때, 단말기 k 의 신호 대 간섭 및 잡음 비 γ_k 는 (2)와 같이 표현할 수 있고 이를 통해 시스템의 sum-rate R 를 계산할 수 있는데, 이는 (3)과 같다.

$$\gamma_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sigma^2 + \sum_{\substack{i=1 \\ i \neq k}}^K |\mathbf{h}_k^H \mathbf{w}_i|^2} \quad (2)$$

$$R = \sum_{k=1}^K \log_2(1 + \gamma_k) \quad (3)$$

본 논문에서는 R 을 최대화하기 위한 빔포밍 벡터를 심층 강화 학습 기법을 통해 정하기로 한다. 이 때, 빔포밍 벡터는 완전한 디지털 빔포밍이 아

년 제한된 위상 변환기를 가지고 있는 시스템에서 코드북 기반으로 선택한다고 가정한다. 또, 해당 코드북은 밀리미터파 대역 채널에 적합하다고 알려알려져 있는 이산 푸리에 변환 기반 코드북을 사용하였다[2].

2. 채널 모델

밀리미터파 대역은 일반적으로 Saleh-Valenzuela 채널 모델을 기반으로 한 채널 모델을 사용한다. 본 논문에서는 기지국 안테나가 M 개 이고, 단일 안테나를 사용하는 단말기가 K 개라고 가정하므로 각 단말기 k 의 채널의 형태는 벡터로 표현할 수 있다. 단말기 k 의 채널 벡터 \mathbf{h}_k 는 (4)와 같다[3].

$$\mathbf{h}_k = \sqrt{\frac{M}{L}} \sum_{i=1}^L \delta_i^k \mathbf{a}(\phi_i^k) \quad (4)$$

이 때, L 은 들어가는 총 광선의 수이며 δ 는 정규 분포를 따르는 복소 이득, ϕ 는 광선의 입사각이며 $\mathbf{a}(\cdot)$ 는 기지국 배열 안테나의 배열 조정 벡터이다.

3. 시뮬레이션 결과

본 논문에서는 심층 강화 학습 기법으로 이산적인 행동을 갖는 환경에 적합한 Deep Q-Network (DQN)을 사용한다[4]. 이 환경에서 행동은 코드북 기반 빔포밍 벡터이고 상태는 채널 정보와 그에 따른 부속 정보들, 보상은 시스템의 sum-rate로 정하였다. 기지국의 배열 안테나 요소의 수, 단말기의 수, 채널의 광선 수는 각각 $M=16, K=4, L=3$ 이고 빔포밍 빔 벡터의 개수는 8이다. 내부 네트워크는 은닉층이 2개인 심층 신경망을 사용했고, 입력, 은닉층1, 은닉층2의 노드 수는 64, 64, 32개이다. 각 에피소드는 모두 동일한 채널로 학습될 수 있도록 하였다.

시뮬레이션 결과는 채널을 10000번 생성한 후 나온 값을 평균하여 나타내었다. 성능 비교 대상은 완전한 디지털 빔포밍 방식 중 제로 포싱 기법을 사용했고, 코드북 기반 빔포밍은 완전 탐색 알고리즘과 제한한 알고리즘이 있다.

그림 2는 송신측 신호 대 잡음비 대비 빔포밍 방식에 따른 sum-rate를 나타낸 그래프이다. 이론상으로도 RF 단의 수가 제한적인 코드북 기반 빔포밍 방식은 완전한 디지털 빔포밍 방식에 비해 성능이 크게 낮다고 알려져 있다. 시뮬레이션 결과 또한 제로 포싱 기법의 특성상 낮은 신호 대 잡음비에서는 오히려 코드북 기반 빔포밍이 더 좋은 성능을 보여주나, 신호 대 잡음비가 높아질수록 코드북 기반 빔포밍 방식들이 제로 포싱 기반 디지털 빔포밍 방식 성능에 크게 못 미치는 것을 알 수 있다. 송신측 신호 대 잡음비가 5 dB 이하에서는 코드북 기반 빔포밍이, 그 이상에서는 제로 포싱 기법의 성능이 더 뛰어났으며, 특히 송신측 신호 대 잡음비가 30 dB인 구간에서는 약 25 bps/Hz의 차이를 보였다.

코드북 기반 빔포밍 방식 간의 성능 비교를 하면 코드북 기반 빔포밍 방식에서는 모든 경우의 수를 계산하므로 완전 탐색 알고리즘이 가장 성능이 좋다고 알려져 있고 해당 코드북에서 임의로 빔 벡터를 선택하는 경우도 생각할 수 있다. 심층 강화 학습을 활용한 에이전트를 통해 빔포밍 벡터를 얻어 냈을 때의 성능도 30 dB에서 1 bps/Hz 이하의 차이로 거의 근접함을 알 수 있고, 임의로 빔 벡터를 선택하는 경우보다는 약 5 bps/Hz 가까이 좋은 성능을 보였다. 이는 본 논문에서 제한한 환경에서는 심층 강화 학습의 행동과 상태의 크기가 크지 않고, 학습을 통해 개선될 부분이 많지 않아 최적에 가까운 결과를 쉽게 얻어냈음을 유추할 수 있다. 다만, 송신측 신호 대 잡음비가 낮은 경우에는 임의로 빔을 선택했을 때와 성능 차이가 거의 없었다. 이는 낮은 송신측 신호 대 잡음비에서, 잡음의 전력

크기로 인해 에이전트가 최적의 빔 벡터를 찾아내지 못하고 올라가지 못한 값으로 수렴해버리는 현상 때문에 임의로 선택하는 것과 큰 차이가 없음을 알 수 있다.

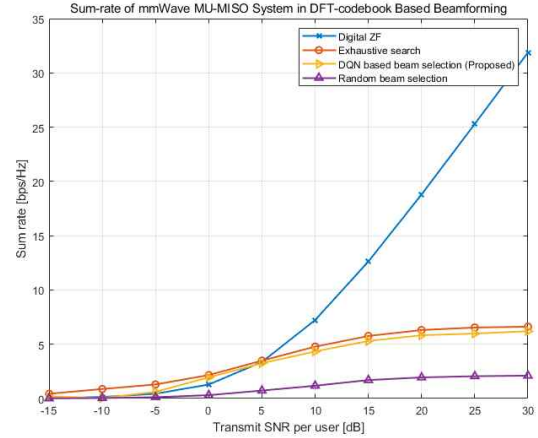


그림 2. 송신측 신호 대 잡음비 대비 빔포밍 방식에 따른 sum-rate

III. 결론

본 논문에서는 배열 안테나가 장착된 기지국에서 단일 안테나를 사용하는 다중 사용자 시스템에서 통신 성능을 높이기 위한 심층 강화 학습 알고리즘을 제안하고 그 성능을 비교 분석하였다. 시뮬레이션 결과 높은 송신측 신호 대 잡음비에서 디지털 빔포밍 환경과 비교하는 것은 무리가 있으나 코드북 기반 완전 탐색 알고리즘과 제한한 알고리즘을 사용한 빔포밍 방식의 성능은 상당히 근접함을 알 수 있었다. 본 논문에서는 알고리즘과 시스템의 한계로 안테나와 단말기의 수가 매우 낮은 상황에서 시뮬레이션을 하였으나, 시스템을 개선하여 안테나와 단말기 등의 수가 훨씬 많은 환경에서 추가적인 연구가 이루어 질 것이다. 추가로, 디지털 빔포밍에 근접한 성능을 보이는 하이브리드 빔포밍 방식에서의 연구 또한 필요하다.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT)(NRF-2020R1A4A1019628)

참 고 문 헌

- [1] T. S. Rappaport et al., "Millimeter wave mobile communications for 5G cellular: It will work!," IEEE Access, vol. 1, pp. 335-349, 2013.
- [2] L. Wan, X. Zhong, Y. Zheng and S. Mei, "Adaptive codebook for limited feedback MIMO system," 2009 IFIP International Conference on Wireless and Optical Communications Networks, 2009, pp. 1-5.
- [3] Stefano Buzzi, Carmen D'Andrea, "On Clustered Statistical MIMO Millimeter Wave Channel Simulation", arXiv:1604.00648 [cs.IT].
- [4] Mnih, V., Kavukcuoglu, K., Silver, D. et al. "Human-level control through deep reinforcement learning." Nature 518, 529 - 533 (2015).