

강화학습 기반 AIRS 위치 최적화 및 반사 요소 선택 기법

김다은, 윤정우, 김수민, 김준수*

한국공학대학교

yasmin34@tukorea.ac.kr, yoon3932332@tukorea.ac.kr, suminkim@tukorea.ac.kr, *junsukim@tukorea.ac.kr

Reinforcement Learning based AIRS Positioning and Reflection Element Selection Scheme

Da Eun Kim, Jeong Woo Yoon, Su Min Kim, Junsu Kim*
Tech University of Korea

요약

본 논문은 6G 통신에서 주목받는 항공형 지능형 반사 표면(AIRS, aerial intelligent reflecting surface)의 복잡도 완화를 위한 반사 요소 선택 및 최적 위치 선정에 관한 연구를 진행한다. AIRS는 기지국과 사용자 사이에서 신호의 반사를 통해 중계하는 임무를 수행하며 각 반사 요소의 선택 여부는 사용자에게 전달되는 신호의 위상을 기준으로 설정한다. 이에 따른 위치 제어는 Q-learning 기반의 강화학습을 통해 achievable rate을 최대화하도록 진행된다. 결론적으로 위상 선택 범위 조정에 따라 위치 최적화가 이루어지고 통신 성능이 향상하는 것을 확인할 수 있다.

I. 서론

최근 지능형반사표면(IRS, intelligent reflecting surface)과 무인항공기(UAV, unmanned aerial vehicle)를 결합한 모델인 항공형 지능형 반사 표면(AIRS, aerial intelligent reflecting surface)에 관한 연구가 주목받고 있다. IRS는 들어오는 신호에 필요한 위상 변화를 가할 수 있는 수동 반사 요소들의 배열로 구성된 평면 표면으로 반사 신호를 원하는 방향으로 전파하도록 재구성할 수 있다[1]. UAV는 높은 이동성을 가지며 3차원 공간에서 자유롭게 이동할 수 있고, 전파의 가시거리(LoS, line of sight) 확보를 통해 서비스 품질 향상 및 통신 범위에서의 확장 가능성을 가진다. 이들을 결합한 AIRS는 무선 통신 환경에서 기존의 IRS와 달리 이동성을 가지고 다양한 물리적 환경에서 효율적인 통신 경로를 만들어 낼 수 있다. 본 논문에서는 AIRS의 위치 최적화와 위상 선택 범위에 따른 achievable rate을 비교한다.

II. AIRS 위치 최적화 및 위상 선택 기법

A. 시스템 모델

지상의 기지국과 사용자 간의 통신 과정 중 공중에 있는 무작위 위상을 가진 AIRS가 기지국으로부터의 신호를 반사하여 사용자에게 중계한다. AIRS는 M 개의 반사 요소를 갖추고 있으며, 기지국과 사용자 사이 공간의 100m 높이에서 이동한다. 기지국-AIRS, AIRS-사용자 사이는 LoS 환경으로 경로 손실만을 가정하였고, 기지국과 사용자 사이는 NLoS 환경의 Rayleigh fading 채널을 가정했다. 그림 1은 시스템 모델을 나타낸다.

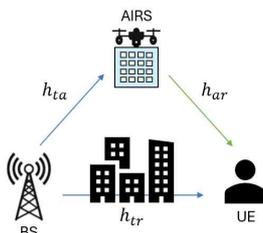


그림 1. 시스템 모델.

$$h_{ta} = \sqrt{\rho d_{ta}^{-\alpha}} [1, e^{-j\frac{2\pi}{\lambda}d\phi_{ta}}, \dots, e^{-j\frac{2\pi}{\lambda}(M-1)d\phi_{ta}}]^T. \quad (1)$$

다음과 같이 경로 손실과 AIRS의 어레이 응답 계산을 통한 채널 이득을 구할 수 있다. 이때, ρ 는 1m당 경로 손실 파라미터로 -20 dB, d_{ta} 는 기지국-AIRS 사이의 거리를 나타낸다[2]. 또한, d 는 $\lambda/2$ 로 antenna separation, λ 는 carrier wavelength, ϕ_{ta} 는 기지국-AIRS로의 도착 위상에 대한 코사인값을 나타낸다. h_{ar} 도 같은 방식으로 구하며 ϕ_{ar} 은 AIRS-사용자에게서의 출발 위상에 대한 코사인값을 가진다. 이후 식(2), (3)을 통해 신호 대 잡음비(SNR, signal-to-noise ratio)와 achievable rate을 계산할 수 있다[3].

$$\gamma_{tr} = \frac{P|h_{tr} + h_{ar}^H \Theta h_{ta}|^2}{\sigma^2}, \quad (2)$$

$$R_{tr} = \log_2(1 + \gamma_{tr}). \quad (3)$$

이때 기지국에서의 전송 전력 P 는 0.01W, AIRS-사용자 사이의 채널 이득은 h_{ar}^H , 기지국-AIRS 사이의 채널 이득은 h_{ta} , noise variance σ^2 는 -80 dBm, AIRS 초기 위상 $\Theta = \text{diag}[e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_M}, \theta_i] \in [0, 2\pi)$ 을 나타낸다.

B. 반사 요소 선택 및 위치 최적화 기법

먼저 반사 요소 선택 기법에 대해 설명한다. 반사 요소 선택에 따른 achievable rate은 반사 요소들의 on/off를 통해 이루어진다.

$$-\alpha \leq \Theta + \phi_{ta} + \phi_{ar} \leq \alpha. \quad (4)$$

식(4)는 사용자가 받을 수 있는 신호의 범위를 나타내는 위상들의 합 표현이다. 사용자와 AIRS 사이에 대한 위상을 기준으로 $\pm\alpha$ 의 범위에 속

하는 반사 요소의 신호만을 고려한다. 이때 α 는 개별 반사 표면을 선택하는 위상의 범위를 나타낸다. 이후 사전에 설정된 α 에 대해 AIRS의 최적 위치를 탐색하기 위한 강화학습 절차를 수행한다.

강화학습은 한 episode 내에서 agent가 step에 따라 action을 반복하여 누적된 reward를 최대화하도록 학습시키는 방법이다. 강화학습 기법 중 Q-learning을 이용하였고 epsilon decay를 적용하여 학습을 진행할수록 exploration을 줄여 수렴하도록 하였다.

표 1. 강화학습 파라미터.

component	parameter
agent	AIRS
state	$x[t], y[t], \text{achievable rate}$
action	$\pm\Delta x, \pm\Delta y$
reward	achievable rate change

강화학습 환경은 표 1과 같이 agent를 AIRS로 설정했고 reward를 최대화하도록 action을 수행한다. action은 x, y축 방향으로의 이동과 움직이지 않았을 경우를 고려하여 총 9가지로 설정했다.

C. 시뮬레이션 결과

그림 2는 위상 선택 범위에 따른 achievable rate 비교 결과이다. AIRS-사용자가 이루는 위상이 $\pm 90^\circ$ 일 때까지는 반사 요소 선택 범위와 동일 위상으로 신호를 받아 성능이 향상되는 것을 볼 수 있다. 그러나 이후 선택 범위가 확장될수록 반대 위상을 가진 신호들도 포함되어 오히려 통신 성능이 감소하는 것을 확인했다.

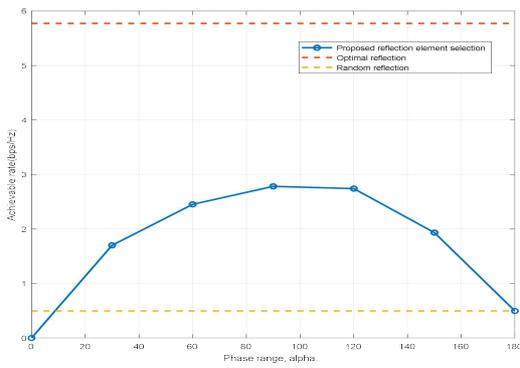
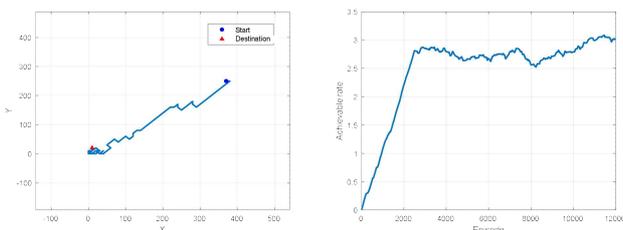


그림 2. 위상 범위 선택에 따른 achievable rate.

그림 3은 $\alpha = 90$ 인 경우 강화학습을 통해 최적 위치를 탐색한 결과를 나타낸다. (a)는 단일 episode 내에서 각 step에 따른 진행이며 초기 생성 지점에서 achievable rate이 향상되는 방향으로 이동하는 과정을 보여준다. (b)는 강화학습 진행 결과로 학습이 지속될수록 통신 성능이 특정 지점으로 수렴하는 것을 볼 수 있다.



(a) AIRS 최적 위치 탐색 결과.

(b) 에피소드 진행 과정.

그림 3. 강화학습 결과.

III. 결론

본 논문에서는 AIRS를 활용한 통신 시스템에서 복잡도 완화를 위한 반사 요소 선택 및 최적 위치 선정에 고려한 통신 성능에 관한 연구를 진행했다. 반사 요소 선택 범위에 따른 achievable rate을 비교하여 반사 요소와 AIRS-사용자 간에 형성되는 위상이 $\pm 90^\circ$ 범위 내일 때 성능이 향상되는 것을 확인할 수 있었다. 이후 사전에 설정된 범위에 대해 최적 위치 탐색을 위한 강화학습을 진행했다. 강화학습은 Q-learning을 이용하여 achievable rate을 극대화하는 위치로 이동하도록 진행했다. 향후 본 논문에서 다루지 않은 에너지 효율 측면에서의 반사 요소 선택 범위 설정 및 최적화 등에 관해서 연구를 발전시킬 가능성이 있다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT혁신인재 4.0 사업과 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(IITP-2024-RS-2022-00156326, 2021R1A2C1013150)

참고 문헌

- [1] C. Pan et al., "Reconfigurable Intelligent Surfaces for 6G Systems: Principles, Applications, and Research Directions," IEEE Communications Magazine, vol. 59, no. 6, pp. 14-20, June 2021, doi: 10.1109/MCOM.001.2001076.
- [2] S. Li, B. Duo, X. Yuan, Y. -C. Liang and M. Di Renzo, "Reconfigurable Intelligent Surface Assisted UAV Communication: Joint Trajectory Design and Passive Beamforming," IEEE Wireless Communications Letters, vol. 9, no. 5, pp. 716-720, May 2020, doi: 10.1109/LWC.2020.2966705.
- [3] X. Liu, Y. Yu, F. Li and T. S. Durrani, "Throughput Maximization for RIS-UAV Relaying Communications," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 10, pp. 19569-19574, Oct. 2022, doi: 10.1109/TITS.2022.3161698.