

## 강화학습을 이용한 안테나 개구면 튜닝

노상미, 김민기, 김형철, 최두석  
삼성전자 System LSI 사업부

sangmi.noh@samsung.com, mingi12.kim@samsung.com,  
hcheol8.kim@samsung.com, doo.choi@samsung.com

## DQN-based Antenna Aperture Tuning

Sangmi Noh, Min-Gi Kim, Hyungcheol Kim, Dooseok Choi  
Samsung Electronics, System LSI Business

## 요약

본 논문은 무선통신 단말 안테나의 주파수 튜닝 문제에 강화학습을 적용하여 실시간 주파수 튜닝 방안을 제안하였다. 뉴럴 네트워크 기반 강화학습 에이전트를 설계하고 실제 단말 상황을 모사한 강화학습 환경을 구성하여 훈련을 진행하였다. 성능 검증 결과 강화학습 에이전트는 5 가지 개구면 튜닝용 코드 중 현재 주파수 및 단말 상황에 맞는 최적의 코드를 89.54%의 정확도로 예측하여 강화학습을 이용한 실시간 단말 개구면 튜닝 가능성을 확인하였다.

## I. 서론

무선통신 기술의 발전에 따라 한정된 단말 폼팩터 내에서 다양한 주파수를 지원하는 안테나를 구현해야 한다. 주파수별 안테나를 모두 실장하기 어려우므로 주파수 튜닝 기술을 활용한다.

안테나의 전기적 길이를 조정하여 방사 성능을 제어하는 기술을 안테나 개구면 튜닝이라고 하며 안테나와 접지부 사이에 캐패시터와 인덕터를 삽입하는 방식으로 구현된다. 실시간 개구면 튜닝을 위해서 Programmable Array Capacitor(PAC) 및 sp4t(single-pole four throw switch) 등 재구성 가능한 수동 소자들을 활용하여 안테나와 접지부 사이에 연결되는 lumped element를 실시간 변화시킬 수 있다[1]. 단말의 실시간 동작 상황에서 안테나 및 주변 환경에 따라 최적의 방사 성능을 유지할 수 있는 튜닝 상태를 유지하는 것은 통신 성능을 결정짓는 매우 중요한 요소이다.

한편, 강화학습은 일련의 의사결정이 당장에 효과는 알 수 없지만 먼 미래에 효과를 나타내는 문제를 해결할 때에 최적화된 학습 방식으로 로봇제어, 차량 자율주행 등의 분야에서 많이 연구되고 있다 [2]. 안테나 개구면 튜닝의 경우 튜너의 설정값 변화에 따라 즉각적인 통신 품질 변화가 있으므로 강화학습의 강점을 활용하는 것은 어려우나 주어진 환경에 대해 최적의 의사결정을 하는 에이전트를 훈련시킨다는 점에서 실시간 단말 안테나 튜닝에 적용될 가능성이 있다고 판단하였다.

본 연구에서는 강화학습을 이용하여 현재 단말이 사용중인 주파수에서 최적의 개구면 튜닝 조합을 찾아내는 방법을 제안하고 시뮬레이션을 통해 성능을 검증하였다.

## II. 본론

실시간 단말 안테나 튜닝은 일정 주기마다 안테나 임피던스를 관찰하고 주파수 및 튜너 설정값 등의 정보와 함께 현재의 무선통신 환경을 유추하여 최적의

튜너 설정값을 도출하는 문제이다. 따라서 해당 문제를 해결하기 위해 강화학습 구조를 차용하였다.

그림 1 은 본 연구에서 가정한 시스템 모델이다. 안테나와 접지면 사이에 실시간 조작 가능한 개구면 튜너를 실장하고 일정 주기마다 튜너를 포함한 안테나의 반사계수를 측정할 수 있는 무선통신 단말 시스템이다.

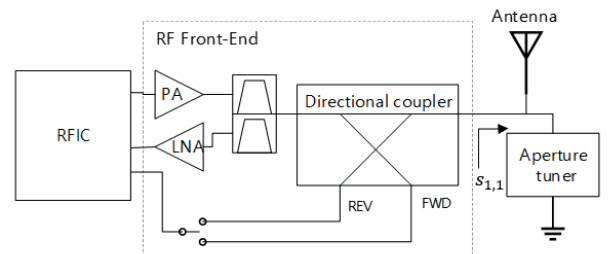


그림 1 개구면 튜너를 포함한 무선 통신 단말

단말은 매 주기마다 안테나 반사계수를 실시간 측정하여 현재 안테나의 상태를 판별하고 그에 맞는 최적 튜너 셋팅을 적용한다.

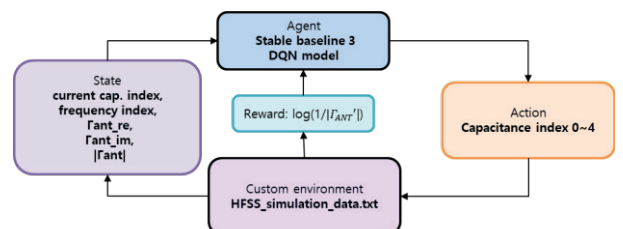


그림 2 실시간 개구면 튜닝을 위한 강화학습 구조

그림 2 는 본 연구에서 사용한 강화학습 구조이다. 강화학습 에이전트는 현재 단말의 상태, 즉 개구면 튜너 셋팅값, 주파수 정보, 안테나 반사계수 정보를 입력으로 받아 현 상태에서 안테나 반사계수의 크기를 최소화할 수 있는 튜너 셋팅을 최적 동작으로 출력한다. 강화학습 환경은 출력된 최적 동작, 즉 최적 튜너 세팅값을 적용한 후 안테나 반사계수와 튜너 세팅값, 그리고 현재

사용하는 주파수 정보 등 현재 상태를 업데이트하여 에이전트에게 넘겨주도록 설계하였다. 강화학습 환경은 또한 에이전트가 출력한 동작에 대한 리워드를 에이전트에게 피드백한다. 안테나 반사계수를 최소화해야 하므로 리워드는 수식 (1)과 같이 안테나 반사계수에 반비례하도록 설계하였다.

$$r(s,a) = \log\left(\frac{1}{|s_{1,1}|}\right) \cdots (1)$$

강화학습 에이전트는 Deep-Q Network 로 구성하였으며 그림 3 과 같이 5 개의 입력을 받아 2 개의 컨벌루션 레이어를 거쳐 각 개구면 튜너 셋팅값에 대한 Q-value 5 개를 출력하는 구조[3]이며 PyTorch 에서 Stable Baseline 3 라이브러리[4]를 이용하여 구현하였다. 강화학습 에이전트는 출력된 Q-value 를 기반으로 주어진 상태에 대하여 가장 Q-value 가 높은 동작을 선택한다.

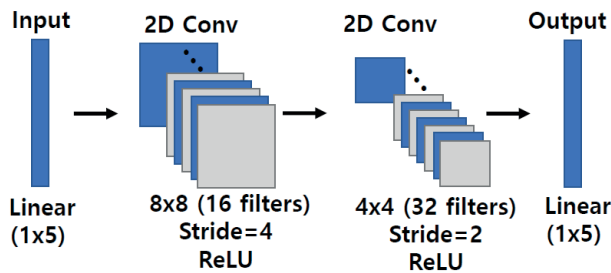


그림 3 Deep Q-Network 구조

HFSS 시뮬레이션을 통해 Inverted-F 안테나 구조에  $L_f=1nH$ ,  $C_f=1.5pF$ ,  $C_r=0.5pF \sim 2.5pF$  을 적용하여 5 가지 안테나 임피던스 커브를 데이터셋으로 사용하였다.

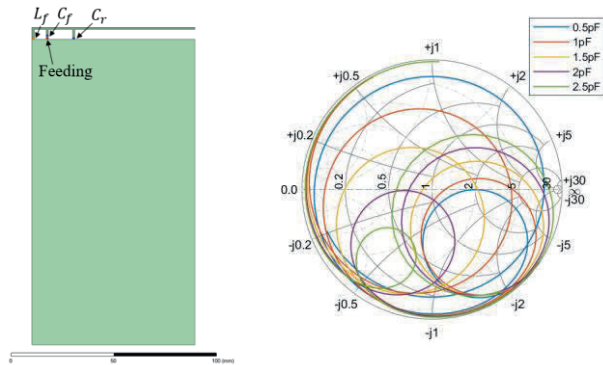


그림 4 Inverted-F 안테나 시뮬레이션 데이터

강화학습 환경은 주어진 데이터셋을 이용하여 단말 실시간 동작을 모사하였으며 현재 사용하는 주파수 및 현재 설정된 캐패시터 값에 따른 안테나 반사계수를 현재 상태에 업데이트 한다. Gymnasium 라이브러리[5]를 활용하여 커스텀 환경을 구현하였다.

에이전트 훈련 시 experience replay,  $\epsilon$  (엡실론)-greedy, gradient clipping 기법 등을 적용하였다. 엡실론  $\epsilon$  값의 경우 훈련이 진행됨에 따라 선형적으로 감소하도록 설계하였다. 에이전트는 다양한 경험을 하기 위해 훈련 초기에는 높은 확률로 랜덤 동작을 선정하며 훈련이 마무리되어 갈수록 낮은 확률로 랜덤 동작을 선정하며 높은 확률로 Q-value 가 높은 최적 동작은 선정한다. 또한 DQN 의 learning rate 도 훈련이 진행될수록 선형적으로 감소하도록 설계하였다.

표 1 강화학습 에이전트 훈련 파라미터

Hyper parameter	Value
Episode 개수	8000
훈련 시작 에피소드	256
시작 Learning rate	0.0004
최종 Learning rate	0.00005
시작 $\epsilon$ (엡실론)	0.8
최종 $\epsilon$ (엡실론)	0.5
Gradient clipping 레벨	8

제안한 강화학습 에이전트는 무선단말 실시간 상황을 모사한 시뮬레이션에서 89.54%의 정확도로 안테나 반사계수의 크기가 최소가 되는 캐패시터 인덱스를 출력하였다. 본 연구를 통해 실시간 단말 안테나 개구면 튜닝에 강화학습을 적용할 수 있는 가능성을 보였다.

### III. 결론

본 연구에서는 고정된 단말 상황에서 다양한 병렬 캐패시터값 및 주파수에 따른 반사계수 시뮬레이션 데이터를 이용하여 주파수별 최적 개구면 튜닝용 캐패시터값을 출력하는 강화학습 에이전트를 설계하고 성능을 검증하였다. 향후 다양한 단말 상황에서 동작할 수 있도록 모델 확장이 필요하며 시뮬레이션이 아닌 실제 단말 상황에서의 성능 검증이 필요하다.

### 참 고 문 헌

- [1] Qizheng Gu, and Arthur S. Morris, "A new method for matching network adaptive control," IEEE transactions on microwave theory and techniques vol. 61, no.1, pp. 587-595, Jan. 2013.
- [2] L. Brunke, et al., "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annual Review of Control, Robotics, and Autonomous Systems* vol. 5, no.1, pp. 411-444, (2022)
- [3] Volodymyr Mnih, et al., "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602 (2013).
- [4] A. Raffin, et al., "Stable-baselines3: Reliable reinforcement learning implementations," *Journal of machine learning research(JMLR)*, vol. 22, no.268, pp. 1-8, (2021)
- [5] M. Towers, et al., "Gymnasium: A standard interface for reinforcement learning environments," arXiv preprint arXiv:2407.17032 (2024)