METERACOM

Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation

FOR 2863 Meteracom Metrology for THz Communications

UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR TECHNISCHE INFORMATIK

C1: Reinforcement Learning for Adaptive Beam Codebook Design in mmWave/THz Systems

Anouar Nechi, Zakaria Narjis, Rainer Buchty, Mladen Berekovic, Saleh Mulhem

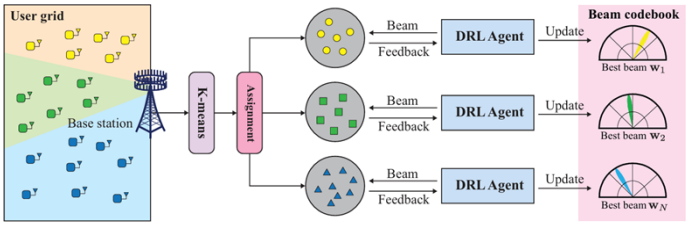
Introduction

Traditional predefined beamforming codebooks for mmWave and THz systems often lead to large sizes, high training overhead, and are unoptimized for challenging conditions like Non-Line-of-Sight (NLoS) propagation and hardware impairments. This work introduces an adaptive reinforcement learning (RL) framework that optimizes beam patterns using only receive power measurements, thereby eliminating the need for explicit channel knowledge.

We present a comparative analysis of three advanced RL algorithms, Deep Deterministic Policy Gradient (DDPG), Twin Delayed Deep Deterministic Policy Gradient (TD3), and Soft Actor-Critic (SAC), to design robust and adaptive beam codebooks. This study provides the first in-depth evaluation of these algorithms under the combined challenges of NLoS conditions and hardware imperfections, to identify a solution that offers superior beamforming gain, faster convergence, and enhanced robustness.

Adaptive RL Framework

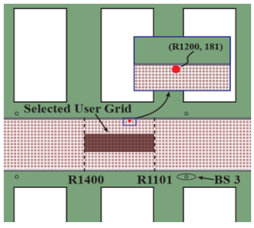
We propose an RL framework that redefines the complex codebook design problem as an optimization task aimed at maximizing user SNR. The approach employs a multi-agent RL paradigm to decompose this task into more manageable sub-problems.



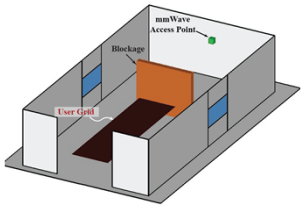
Architecture of the beam codebook framework with the proposed RL algorithms

Experimental Setup

Communication Scenarios: The performance evaluation was conducted across two distinct communication scenarios using the DeepMIMO dataset: an outdoor LoS scenario operating at 60 GHz and an indoor NLoS scenario at 28 GHz.



LoS scenario 60 GHz

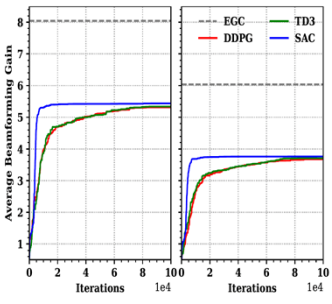


NLoS scenario 28 GHz

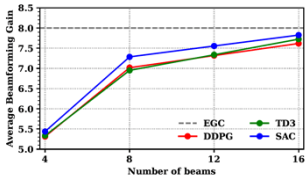
The system model employed a massive MIMO base station with a single RF chain and analog-only beamforming using quantized phase shifters. A primary focus was on evaluating the robustness of the RL solutions against hardware imperfections, such as antenna spacing and phase mismatches, which were unknown to the learning agent.

Results & key Findings

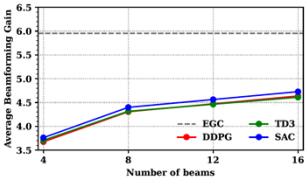
Performance & Convergence: Simulation results demonstrate the superiority of the SAC algorithm, which consistently achieved higher beamforming gain and faster convergence compared to DDPG and TD3 across various scenarios. SAC reached its maximum gain within 10^4 iterations, whereas DDPG and TD3 required $7 \cdot 10^4$ iterations.



Learned 4-beam patterns for a 32-antenna (LoS / NLoS)



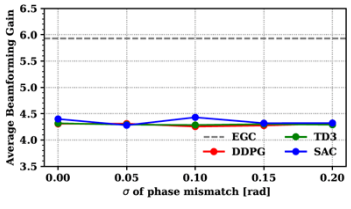
Beamforming gain vs. number of beams in a LoS scenario



Beamforming gain vs. number of beams in a LoS scenario

Robustness to Hardware Impairments: The evaluation showed that all proposed RL solutions are robust against hardware impairments. Specifically, under NLoS conditions with varying degrees of phase mismatch, the SAC-based solution consistently maintained higher beamforming gain.

Impact of hardware impairments on 8-beam codebook learning



Conclusion

Our work successfully applied an RL framework to optimize beam codebooks for mmWave and THz massive MIMO systems. The results demonstrated that the SAC algorithm consistently outperformed DDPG and TD3, achieving superior beamforming gain and faster convergence. SAC was particularly robust in a challenging NLoS scenario with hardware impairments. The findings highlight the potential of this RL-based framework for enabling efficient and adaptable communication in future wireless systems.