

The logo for RADemics, featuring the text "RADemics" in white on a blue arrow-shaped background pointing to the right. The arrow is part of a larger blue horizontal bar that is attached to a dark blue vertical bar on the left side of the slide.

RADemics

# Reinforcement Learning in Antenna Array Control and Resource Allocation for 6G Networks

A decorative graphic consisting of several thin, curved lines in shades of blue and grey, originating from the bottom left and extending upwards and to the right, resembling stylized grass or reeds.

Patil Vaishnavi Pradip, Makane Jayashree Nilkanth  
M.S.BIDVE ENGINEERING COLLEGE

# Reinforcement Learning in Antenna Array Control and Resource Allocation for 6G Networks

<sup>1</sup>Patil Vaishnavi Pradip, Assistant Professor, Department of Electronics & Telecommunication Engineering M. S. BIDVE Engineering College Barshi Road Latur, Maharashtra -413512. [vaishu9921@gmail.com](mailto:vaishu9921@gmail.com)

<sup>2</sup>Makane Jayashree Nilkanth, [makanejayashree659@gmail.com](mailto:makanejayashree659@gmail.com)

## Abstract

The rapid evolution of sixth-generation (6G) wireless networks introduces unprecedented demands on throughput, latency, reliability, and connectivity, driven by applications such as ultra-reliable low-latency communication (URLLC), autonomous systems, and immersive services. High-frequency bands, including millimeter-wave and terahertz spectra, coupled with extremely large-scale antenna arrays (ELAAs) and dense network deployments, present significant challenges in beamforming, user association, and resource allocation. Conventional optimization and heuristic techniques struggle to address the dynamic and high-dimensional nature of these networks. Reinforcement Learning (RL), particularly Deep and Multi-Agent RL, provides a model-free framework capable of adaptive and intelligent decision-making, enabling real-time optimization of antenna array control and distributed resource allocation. By jointly optimizing beamforming, user association, power control, and spectrum management, RL frameworks enhance spectral efficiency, energy efficiency, and network reliability while meeting stringent latency requirements. This chapter presents an in-depth exploration of RL methodologies applied to antenna array and resource management, emphasizing URLLC scenarios, highlighting system models, learning algorithms, performance metrics, and future research directions. The chapter demonstrates the transformative potential of RL for autonomous, self-optimizing 6G networks.

Keywords: 6G Networks, Reinforcement Learning, Antenna Array Control, Resource Allocation, Ultra-Reliable Low-Latency Communication (URLLC), Multi-Agent Systems.

## Introduction

The advent of sixth-generation (6G) wireless networks marks a transformative phase in communication technologies, driven by the need for ultra-high data rates, ultra-low latency, and ubiquitous connectivity [1]. Emerging applications such as autonomous vehicular networks, remote robotic surgery, holographic telepresence, and industrial automation impose stringent performance requirements that extend beyond the capabilities of fifth-generation (5G) systems [2]. Achieving reliable communication in 6G necessitates leveraging high-frequency bands, including millimeter-wave (mmWave) and terahertz (THz) spectra, which offer wide bandwidths for high data throughput but also present severe propagation challenges [3]. Signal attenuation, susceptibility to blockage, and limited diffraction in these frequency ranges demand the

deployment of highly directional and adaptive antenna systems. Extremely large-scale antenna arrays (ELAAs), massive MIMO, and intelligent reflecting surfaces (IRS) have emerged as key enabling technologies, providing spatial diversity, beamforming gains, and controllable propagation environments [4]. These technologies, however, introduce substantial complexity in system design and control, particularly in dynamically allocating resources and maintaining optimal connectivity across dense user populations. Traditional rule-based or optimization-driven solutions often fail to meet the real-time and adaptive requirements of 6G, necessitating intelligent frameworks capable of learning and self-optimization [5].

Beamforming and antenna array management play a central role in addressing the challenges associated with high-frequency 6G communications [6]. Precise beam steering, alignment, and tracking are required to maintain robust links, especially in mobile and high-mobility scenarios such as vehicular-to-everything (V2X) and drone-assisted communications [7]. Beam misalignment or delayed adaptation can cause substantial signal degradation, leading to intermittent connectivity or increased latency. Conventional beam management techniques, including exhaustive search, codebook-based methods, and iterative optimization, are computationally intensive and often unable to respond to rapid channel variations in real time [8]. The interdependence between beamforming and user association introduces additional complexity; the selection of a beam directly influences which users can be served efficiently, while user mobility patterns affect the optimal beam configuration [9]. Addressing this coupling in dense, dynamic environments requires adaptive algorithms capable of learning optimal strategies from ongoing network interactions. Reinforcement Learning (RL) has demonstrated the potential to fulfill this requirement by providing a model-free approach that continuously optimizes beam selection and user association policies based on network feedback, enabling autonomous and real-time antenna control [10].

Resource allocation constitutes another critical aspect of 6G networks, where the dynamic assignment of spectrum, power, and scheduling resources determines overall network efficiency, reliability, and user quality of service [11]. The high dimensionality of 6G resource allocation problems, compounded by the heterogeneity of services such as enhanced mobile broadband (eMBB), massive machine-type communication (mMTC), and ultra-reliable low-latency communication (URLLC), creates a complex, non-convex optimization landscape [12]. Conventional optimization methods, such as convex programming, iterative algorithms, and heuristic-based approaches, struggle to accommodate dynamic channel conditions, high user density, and rapid mobility [13]. Reinforcement Learning offers a scalable and adaptive solution by enabling agents to interact with the environment, evaluate reward functions aligned with network objectives, and iteratively refine policies for spectrum assignment, power control, and user scheduling [14]. Multi-Agent Reinforcement Learning (MARL) further enhances scalability and decentralization by allowing multiple network entities, including base stations and access points, to independently optimize local resources while coordinating with other agents, effectively balancing local and global performance objectives in distributed deployments [15].