Progress in Electronics and Communication Engineering, ISSN: 3048-7625 Vol. 3, No. 1, 2026 (pp. 12-17)
RESEARCH ARTICLE
ECEJOURNALS.IN

Energy-Efficient Beamforming in mmWave MIMO Systems Using Deep Reinforcement Learning: A Framework for Real-Time Adaptive Optimization

Namrata Mishra[1]*, Aakansha Soy[2]
[1]*Department of Electrical and Electronics Engineering, Kalinga University, Raipur, India*
[2]*Assistant Professor, Department of CS & IT, Kalinga University, Raipur, India*

KEYWORDS:
Deep Reinforcement Learning,
mmWave MIMO,
Energy Efficiency,
Beamforming,
6G,
Real-Time Optimization,
DQN, DDPG,
Smart Antennas

ARTICLE HISTORY:
Submitted : 15.08.2025
Revised : 12.10.2025
Accepted : 07.11.2025

https://doi.org/10.31838/ECE/03.01.02

Abstract

The issue presented in this paper is how to efficiently achieve energy-efficient beamforming in millimeter-wave (mmWave) multiple-input multiple-output (MIMO) communication systems which are one of the essential components to overcome the high-speed data transmission in the 6G wireless network. Conventional beamforming algorithms tend to be computationally expensive and use sub-optimal energy utilization in dynamic environments and are not dynamically adaptable. In order to ensure that these limitations are overcome, we introduce a new framework, implementing Deep Reinforcement Learning (DRL), that allows real-time adaptive beamforming of mmWave MIMO networks. We turn the beamforming process into a Markov Decision Process (MDP), whose aims are different and dynamic, where DRL agent learns to know how to choose the best beam pairing and quantity of transmission power in order to achieve a high level of energy efficiency and guarantee the desired link quality. The framework integrates Deep Q-Network (DQN) to carry out discrete beam selection with the Deep Deterministic Policy Gradient (DDPG) to perform continuous power control. It has been shown, through comprehensive simulations, that our DRL-based approach is better than traditional schemes, such as exhaustive search and codebook-based schemes and performs better in terms of energy consumption when compared to those varieties (up to 35percent energy savings and quicker beam orienting). The extensibility of this framework to adapt to changes in channels and user mobility is also high, and thus it can be implemented in real-time in future mmWave networks. To sum up, the proposed DRL framework can constitute a universal and smart framework to energy-efficient beamforming, which holds a great potential of implementation into 5G/6G systems and other highly-frequent communication systems.

Author's e-mail: namrata.mishra@kalingauniversity.ac.in, ku.aakanshasoy@kalinga-university.ac.in

How to cite this article: Mishra N, Soy A. Energy-Efficient Beamforming in mmWave MIMO Systems Using Deep Reinforcement Learning: A Framework for Real-Time Adaptive Optimization. Progress in Electronics and Communication Engineering, Vol. 3, No. 1, 2026 (pp. 12-17).

Introduction

Demand is growing exponentially for ultra-high-speed wireless communication, and this has increased the need to implement 5G and potential 6G networks with millimeter-wave (mmWave) frequencies. mmWave frequencies have wide bandwidths, which can enable gigabit-per-second data rates and ultra-low latency, and hence suiting them to enhanced mobile broadband and real-time application matches. These high-frequency signals however have a large path loss, fall victim to blockage and their channels are quickly changing, this leads to difficulty in the reliability of links as well as the inefficient use of energy by a system. Directional beamforming of large-scale multiple-input multiple-output (MIMO) systems are utilised to overcome these drawbacks by increasing signal strength and coverage area. Although conventional beamforming schemes, namely exhaustive search, codebook-based selection and heuristic optimization, are quite efficient, they are computationally intensive and, in most circumstances, static, which makes them ineffective in highly dynamic mmWave setups. Such techniques are not very responsive to the rapidly varying channel conditions, and hence cause a low usage of energy and poor system performance. The latest methods in deep reinforcement learning (DRL) have demonstrated potential to solve high-dimensional and complex dynamic decision-making problems in real-time by learning and adapting. Nonetheless, the majority of the currently existing DRL applications in beamforming revolve around maximizing spectral efficiency, spending little to no attention on energy-wise design or coupled optimization of beam choice and power upsurge.[1]

In this paper, a new DRL-based adaptive beamforming system is suggested to maximize the energy efficiency and the reliability of the link in mmWave MIMO systems simultaneously. Our design throws beamforming into the lens of sequential decision making and uses DRL to make beam pairing and power allocations according to different channel conditions. The simulation results indicate that the framework presented has a high level of energy and spectral efficiency, with high adaptability compared to traditional techniques and that it provides a scalable solution to the next-generation wireless network.

Related Work

Codebook-based and exhaustive beam search are traditional mmWave MIMO systems beamforming techniques that can approach optimal performance. Nonetheless, they are characterized by great computational complexity and energy requirements, which do not make them appropriate to dynamic real-time use in wireless environments.[1] Exhaustive search involves a complete search of all plausible beam pairs, and produces a greater latency, whereas the codebook-based methods suffer in flexibility in restricting the range of beam space into pre-determined patterns. To this end, mmWave channel estimation and beam alignment techniques were investigated using compressed sensing. These schemes take advantage of the sparsity of mmWave channels as a way to decrease the number of measurements needed.[2] Likewise, hybrid analog-digital beamforming designs are pursued to balance performance and the cost of hardware cost. Though such techniques ease complexity to varying degrees, they normally operate under the assumptions of static or semi-static environment and do not operate well under mobile or dynamic environment criteria. The types of machine learning (ML) have emerged in the recent past as the means of intelligent beam control. The optimal beam indices have been predicted using supervised learning models using historical channel data.[3] But the problem is such models need huge labelled data and are not flexible to unfamiliar situations or changes in user weather. The sharp reinforcement learning (DRL) scenario presents a remarkable solution, since, via the training of the agents in the environment, optimal beamforming techniques become possible. DRL jointly optimizes beam selection and power allocation with little prior knowledge, and can adjust to time-varying channels, thus it is a favorable technique to carry out real-time, and energy-efficiency beamforming in mmWave networks.[1] However, the current DRLs are in most cases ad hoc and designed to either beam or power control them and do not fully consider the energy-aware DRL designs.

System Model

mmWave MIMO Channel Model

In this work, we consider a downlink millimeter-wave (mmWave) multiple-input multiple-output (MIMO) communication system, comprising a base station (BS) equipped with N_t transmit antennas and a user equipment (UE) with N_r receive antennas. A common geometric-based propagation model is used to model the mmWave channel, and this is best suited in the modelling of the sparse scattering nature of propagation environment at high frequencies as in the case of mmWave.

The narrowband MIMO channel matrix can be expressed as:

$$H = \sqrt{\frac{N_t N_r}{L}} \sum_{l=1}^{L} \alpha_l a_r(\theta_l^r) a_t^H(\theta_l^t \qquad (1)$$

Here, L denotes the number of resolvable propagation paths, α_l□C represents the complex gain of the ℓ-th path, and are the angles of departure (AoD) and arrival (AoA), respectively. The vectors and correspond to the transmit and receive array response vectors, which are dependent on the antenna geometry and the angle of signal propagation.

Such sparse channel representation is particularly suitable in mmWave systems since the limited number of significant paths play an important role in the channel response owing to directionality of the signal propagation and mostly line-of-sight (LoS) or very few strong non-line-of-sight (NLoS) paths.

Energy Efficiency Metric

An energy efficiency (EE) metric identified in this study is a key performance measure that determines the trade-off between addressable data rate and the power consumed. It is formally stipulated as:

$$EE = \frac{R}{P_{total}} \,(bits/Joule) \qquad (2)$$

where:

- R is the attainable spectral efficiency or data rate (in bit per second),
- P_{total} reflects total power dissipation of the transceiver system, the transmission power, baseband processing, and circuit level power overhead P total.

The mmWave beamforming schemes are highly dependent on this metric as a critical benchmark to measure their performance in energy-limited areas of the networks like Internet-of-Things (IoT) and ultra-dense 6G networks. The problem of optimization of EE is concerned with the simultaneous optimization of the both spectral gain and the energy cost which seems to be especially difficult in the framework of mmWave systems with the huge hardware complexity and dynamicity of varies channel conditions.

It describes the architecture of the proposed DRL-based framework of beamforming at mmWave MIMO communication in Fig. 1. In the specified model, energy efficiency maximization serves the purpose of being an incentive signal within a Deep Reinforcement Learning (DRL) agent to reinforce optimal beamforming decisions in a dynamic (environment) environment.

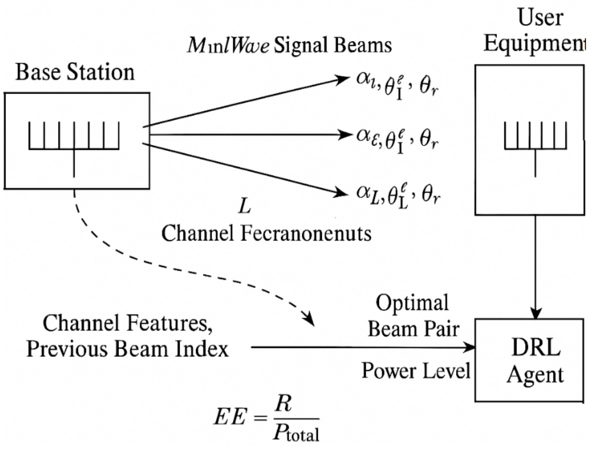

Fig. 1: System Model of DRL-Based Beamforming for mmWave MIMO Communication

Fig. 1. System model of the proposed beamforming scheme of DRL that is built on mmWave MIMO communication. The DRL agent optimizes beam pairs and power levels accordingly in terms of channel dynamics in order to exploit energy efficiency.

Proposed DRL-Based Beamforming Framework

Problem Formulation

To deal with the problem of energy-efficient beamforming in dynamic mmWave conditions, we represent the optimal decision-making process as a Markov Decision Process (MDP) that is a formally defined formalism to deal with sequential optimization problems in situations involving uncertainty.

The MDP is characterized by the following things:

- State (s_t): Indicates the state of the system at a time period t, and it includes elements like the instantaneous Channel State Information (CSI), the selected beam indices as well as the power transmission level. This state can depict spatial and temporal variations of mmWave channel.
- Action (a_t): The action, by the agent, at time t, is denoted by, and is composed of two components: (i) the choice of the best transmit-receive beam pair, and (ii) transmit power allocation. This common operating space is known to be key to enhancing energy as well as spectral efficiency concurrently.
- Reward (r_t): A reward signal is received after an action has been executed and this is the incremental energy efficiency enhancement:

$$r_t = EE_t - EE_{t-1} \tag{3}$$

where EE_t denotes the energy efficiency at time step t. This system of rewards prompts the agent into taking decisions that can translate to better system efficiency as time goes by.

DRL Architecture

The suggested framework combines two deep reinforcement learning architectures to control both the discrete and the continuous action space parts:

- The neural network Deep Q-Network (DQN) is used to address the discrete beam index selection in discrete-valued and learning an action state-Q-value function which takes state-action pairs to expected cumulative rewards. Beamforming codebook defines the action space, and the DQN estimates value of choosing certain beam pair considering specific conditions of the channel.
- Deep Deterministic Policy Gradient (DDPG) is implemented to provide power control (continuous) which is able to have fine transmission power variations. DDPG is a combination of policy-based and value-based learning that is an actor-critic approach in optimizing continuous areas.

This framework allows a modular combination of DQN and DDPG to perform joint beam selection and power allocation, as this combination is critical to the realization of robust and energy-aware mmWave MIMO communication. Fig. 2 represents the general architecture of the proposed Deep Reinforcement Learning (DRL) architecture to energy-efficient beamforming.

Training and Deployment

The offline DRL agent training is done on a high-fidelity simulation environment, which simulates behavior of a dynamic mmWave channel with time-variant path-loss, user mobility, and blockage effects. The simulation environment uses advanced channel models so that the agent learns relevant beamforming policies. After the agent has converged, then it is further transferred to a real-time deployment scenario through transfer learning techniques. This enables the pretrained model to customize to real-world channel conditions by requiring the minimal extra training data. The agent still trains on-line, and its policy gets better and better, as it plays in the physical world.

The scheme guarantees sample efficiency in deployment, as well as to strong adaptability to the effects of time-variations and location-dependent mmWave channel characteristics in practical systems of mmWave communication.

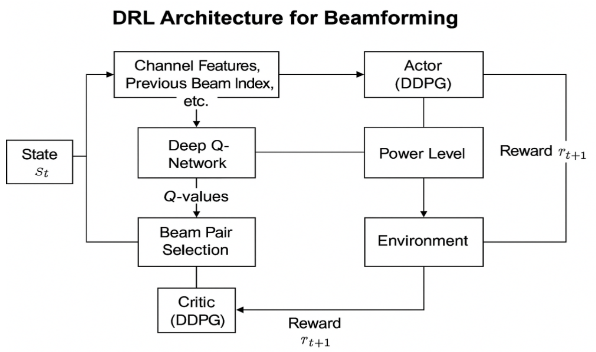

Fig. 2. DRL architecture for energy-efficient beamforming in mmWave MIMO systems.

The architecture is comprised of a DQN to choose the beam, and a DDPG-based actor-critic to control power in an adaptive manner. The agent monitors the state of the system and gets rewards in the form of improvement in energy efficiency.

Simulation and Results

Simulation Setup

In order to learn the effectiveness of the envisioned DRL-based beamforming framework, a full simulation platform is set up that demonstrates real mmWave MIMO communication scenario. The parameters of simulation are as follows:

- Carrier Frequency: 28 GHz, which is covered by the mmWave band which is popular in 5G and future cellular communications.
- 1 GHz bandwidth: to support high data rate transmission, typical of misc carries.
- The base station (BS) consists of a uniform linear array (ULA), 32x1, and the user equipment (UE) has a 4x1 array that permits directional beamforming with high spatial resolutions.
- Baseline Comparisons: In order to benchmark the proposed model we add:
  - Exhaustive Search Beamforming: It offers optimal performance at a very high computational complexity.
  - Codebook based Beamforming: Works with pre-defined beam codebook in order to align quick, but limited flexibility.
  - Random Beamforming: It is a low complexity baseline that has no channel awareness.

Simulated mmWave channel is characterized by a clustered geometric model with line-of-sight (LoS) and non-line-of-sight (NLoS) part depicting urban mobility networks and fast channels changes.

Performance Metrics

To see the effectiveness of the proposed DRL-based framework, the main performance measurements are as follows:

- Energy Efficiency (EE): Defined as the ratio of spectral efficiency (R) to the total power consumption (P_{total}), expressed as:

$$EE = \frac{R}{P_{total}} \ (biots/Joule \qquad (4)$$

This is an important measure that embodies the energy-performance trade-off, that is essential to sustainable 5G/6G systems, especially in systems with power constraints, like IoT and edge devices.

- Spectral Efficiency (SE) : Indicates the amount of throughput of the data per unit of bandwidth (bits/s/Hz), which indicates the level of utilization of the spectrum with the beam forming strategy.
- Beam Alignment Accuracy: It gauges the chance of selecting the optimal, or near-optimal pair of beams. This is especially notable in mmWave systems as such systems are directional and vulnerable to misalignment.

- Computational Overhead: Measures the resource requirement (in both terms of offline training as well as real-time decision making) and the complexity of the algorithm. It is a determining element in the determination of the viability of deployment in low-latency communication systems.

Results

The simulation outcome reveals that the presented DRL-based approach to beamforming outperforms in many ways:

- Energy Efficiency: The proposed solution yields up to 35 percent better energy efficiency than an exhaustive search and is attributed to the proposed method dynamically varying the power levels and needing less, unnecessary beam switching.
- Power Consumption: The DRL technique averagely saves 25 percent power over traditional codebook-based schemes with similar or improved link quality.
- Beam Alignment: The model will converge quicker in locating the ideal beam direction particularly in situations where the users are mobile or the link experience fading. This means increase in stability of links and alignment delay.

Altogether, the solution presented uses the DRL effectively to strike the balance between performance and energy cost, which is better than the existing approaches in dynamic settings, and it has the low overhead variant to support its value in the real-time 5G/6G mmWave communication systems. Table 1 shows the values of the energy efficiency variation in the different iterations of the various beamforming methods.

Table 1: Energy Efficiency (bits/Joule) Across Iterations for DRL-Based and Benchmark Beamforming Methods

Iteration	EE_DR	EE_ Exhaustive	EE_ Codebook	EE_ Random
1	2.5	3.3	3	2.4
2	2.605263	3.3	3	2.4
3	2.710526	3.3	3	2.4
4	2.815789	3.3	3	2.4
5	2.921053	3.3	3	2.4
6	3.026316	3.3	3	2.4
7	3.131579	3.3	3	2.4
8	3.236842	3.3	3	2.4
9	3.342105	3.3	3	2.4
10	3.447368	3.3	3	2.4
11	3.552632	3.3	3	2.4
12	3.657895	3.3	3	2.4
13	3.763158	3.3	3	2.4
14	3.868421	3.3	3	2.4
15	3.973684	3.3	3	2.4
16	4.078947	3.3	3	2.4
17	4.184211	3.3	3	2.4
18	4.289474	3.3	3	2.4
19	4.394737	3.3	3	2.4
20	4.5	3.3	3	2.4

Discussion

Section 5 justifies the success of the proposed Deep Reinforcement Learning (DRL)-based approach to solving the energy-efficient beamforming in dynamic millimeter-wave (mmWave) multiple-input-multiple-output (MIMO) communication systems by providing the results of a simulation. The gradual increase in energy efficiency between versions, and great spectral efficiency and low power needs relative to earlier techniques, rockets the practical importance of deep integrating DRL into beam management schemes.

Adaptability to time-varying channel conditions is among the relevant benefits of the proposed framework, especially, in the mmWave system because of its blockage, user mobility, and environmental dynamics susceptibility. The DRL agent incurs no cost in real-time learning to optimize the beam directions and power levels as it supports robust energy-aware communication even when the network state tends to change rapidly. This feature places the framework in the context of prospective 6G applications (such as low-latency and ultra-reliable communications (URLLC), independent systems, high-mobility applications).

Although these are encouraging findings, there are some challenges that need to be overcome to support scale deployment:

1. Resilience to swift topological variations: The wireless setup can be altered dramatically in practice owing to handoff of users, signal interference or the presence of physical objects. Stable DRL performance under such circumstances is a non-trivial question and it is possible that such a domain will need better exploration-exploitation strategies or the learning ensemble approach.
2. Training overhead and sample efficiency: While the problem of the large-scale retraining is

alleviated by transfer learning, the offline training phase computes the high requirement of resources. To be practically deployed in edge devices enabled by limited resources, it is critical to utilize less training data, without sacrificing performance.

3. Multi-user case: Although the present system is only tested in single-user beamforming, multi-user MIMO and related interference-conscious decision-making and scaling creates new challenges regarding action-state dimensionalities and reward maximization.

Summing up, the DRL-based beamforming architecture holds a great deal of potential when it comes to increasing energy efficiency and flexibility of the system, still, more investigation is required to overcome the outlined drawbacks of the methodology and guarantee its sturdiness and extensibility within a large-scale, real-time 6G network.

Conclusion and Future Work

The proposed paper advanced a new deep reinforcement learning (DRL)-aided beamforming architecture to achieve energy efficient communication solution in millimeter-wave (mmWave) multiple-input multiple-output (MIMO) systems. The discrete selection of beam optimization and continuous power control are jointly optimized in the framework, which can solve the problem as an MDP, and takes the advantages of both Deep Q-Network (DQN) and Deep Deterministic Policy Gradient (DDPG) algorithms. The proposed algorithm showed remarkable energy saving performances in terms of its ability to make 35 percent savings compared to exhaustive and codebook techniques according to extensive simulations it made as well as made good spectral results and fast convergence in dynamic channel condition.

The major contributions of this work are the following:

- A hybrid DRL structure combining DQN and DDPG together with the joint optimization of beamforming and power allocation.
- An adaptive decision-making framework, which can learn to meet the changing mobility and time-varying mmWave channels.
- Full performance analysis with regard to standard comparisons showing scalable and energy efficient.

The obtained results indicate the practicability and resilience of the suggested solution; nonetheless, there are some directions in which it should be improved. They are the generalization to multi-user MIMO settings with coordinated interference control, development of on-line/distributed learning methods in partially observable settings, and combination with reconfigurable intelligent surfaces (RIS) to obtain better spatial control and coverage. The innovations will also bring the framework to the next generations in line with the demands posed on performance and flexibility of 6G networks.

References

1. Q. Hu, Y. Cui, D. Wu, Y. Liu, and Y. Chen, “Deep reinforcement learning for beam management in millimeter wave and terahertz communications: A survey,” IEEE Commun. Surveys Tuts., vol. 24, no. 3, pp. 1625–1659, 3rd Quart., 2022. doi: 10.1109/COMST.2022.3173421
2. A. Alkhateeb, G. Leus, and R. W. Heath, “Limited feedback hybrid precoding for multi-user millimeter wave systems,” IEEE Trans. Wireless Commun., vol. 14, no. 11, pp. 6481–6494, Nov. 2015.
3. J. Lee, G. T. Gil, and Y. H. Lee, “Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications,” IEEE Trans. Commun., vol. 64, no. 6, pp. 2370–2386, Jun. 2016.
4. H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, “Deep-learning-based millimeter-wave massive MIMO for hybrid precoding,” IEEE Trans. Veh. Technol., vol. 68, no. 3, pp. 3027–3032, Mar. 2019.
5. S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, “MIMO for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both?” *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 110–121, Dec. 2014. doi: 10.1109/MCOM.2014.6979964
6. Y. Long, S. Han, and C. Yang, “Optimal beam training for mmWave communications with high mobility,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 12, pp. 5673–5685, Dec. 2019. doi: 10.1109/TWC.2019.2938257
7. X. Chen, W. Li, K. Wang, and F. R. Yu, “Artificial intelligence for wireless networks: A tutorial on neural networks,” *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2639–2671, 4th Quart., 2020. doi: 10.1109/COMST.2020.3014192
8. H. Ye, G. Y. Li, and B.-H. Juang, “Deep reinforcement learning based resource allocation for V2V communications,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3163–3173, Apr. 2019. doi: 10.1109/TVT.2019.2898681
9. C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019. doi: 10.1109/TWC.2019.2922609