This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/CITS61189.2024.10608029, 2024 International Conference on Computer, Information and Telecommunication Systems (CITS)

Performance Evaluation of Deep Q Networks for Hybrid Reconfigurable Intelligent Surface in 6G Networks

Aya Kh. Ahmed
Department of Electronic and Electrical Engineering
College of Engineering, Design, and Physical Science
Brunel University London
London, United Kingdom
https://orcid.org/0000-0002-3902-1760
aya.ahmed@brunel.ac.uk

Hamed S. Al-Raweshidy
Department of Electronic and Electrical Engineering
College of Engineering, Design, and Physical Science
Brunel University London
London, United Kingdom
hamed.al-raweshidy@brunel.ac.uk

***Abstract*—The emergence of 6G wireless communication introduces a new era of connectivity demands, marked by high data rates and varying network conditions. To address these challenges, we propose HRISDQN, a framework that combines Hybrid Reconfigurable Intelligent Surfaces (HRIS) with Deep Q-Network (DQN)-based reinforcement learning. HRISDQN represents a significant advancement in optimising communication in 6G networks, enabling transformative improvements. In our work, we compare HRISDQN with conventional Semi Definite Relaxation (SDR), Maximum Ratio Transmission (MRT), and Minimum Mean Square Error (MMSE) as traditional beamforming techniques. We demonstrate HRISDQN's adaptability to dynamic scenarios through extensive simulations and evaluations, including varying Signal-to-Noise Ratios (SNR) and changing user densities. Our results show that HRISDQN consistently outperforms its counterparts; HRISDQN's resource allocation capability ensures 40% better fairness, lower delay by 80%, and three times higher spectral efficiency, even in high-density user environments. The designed HRISDQN excels under diverse SNR conditions, providing robust and reliable connectivity. HRISDQN's exceptional performance holds great promise for the future of 6G communication. HRISDQN offers ultra-efficient, low-latency, and adaptive communication networks for augmented reality and autonomous vehicles using HRIS and DQN.**

***Index Terms*—Beamforming, deep Q networks, delay, fairness, hybrid reconfigurable intelligent surfaces, 6G.**

I. Introduction

The field of wireless communication has experienced significant progress and has demonstrated notable advancements. With each new technological generation, there is a growing demand for unlimited global interconnectivity, which has become a valued necessity and a central focus. The fundamental essence of the new wireless generation 6G integrated capacities lies in its ability to provide exceptionally high data rates, very imperceptible latency, and uninterrupted connectivity, thereby propelling humanity into an era when the limits of digitalisation are limitless [1]. Given the higher expected capabilities of the 6G, it is crucial to address the current complex issues of high network deployment costs and high energy consumption to develop future sustainable and environmentally friendly wireless strategies. Hybrid Reconfigurable Intelligent Surfaces (HRIS) have received significant attention due to their ability to enhance wireless network capacity and coverage by intelligently modifying the wireless propagation environment [2] As network capacity, coverage, and ultra-low latency communications are the main concerns in 6G wireless communications and a critical requirement for applications such as augmented reality, self-driving vehicles, and the massive Internet of Things (IoT), the HRIS can address these challenges effectively.

HRIS is essential in 6G wireless communication because it can independently change the wireless communication environment [3]. Moreover, HRIS provides high-precision sensing to create a suitable environment for signal propagation and make the interference between wireless signals from two nearby objects less significant, thus improving overall sensing resolution and system performance [4]–[6]. Its low manufacturing, simple hardware adjustment, and low energy consumption also make it a good choice. The core of HRIS revolutionary technology is centred around signal redirection and beamforming. Beamforming is a fundamental technique in wireless communication that is used to enhance the signal-to-noise ratio (SNR) of received signals, eliminate undesirable interference sources, and focus transmitted signals on specific locations [7].

Conventional Semi Definite Relaxation (SDR), Maximum Ratio Transmission (MRT), and Minimum Mean Square Error (MMSE) have traditionally served as fundamental techniques in the field of beamforming [8]. SDR, renowned for its inherent simplicity, endeavours to mitigate interference by orthogonal use of the transmission signals. The MRT technique, in contrast, utilises the concept of optimising the SNR at the receiver, hence offering resilience in environments with high noise levels [9]. The Minimum Mean Square Error (MMSE) technique is more intricate and aims to minimise the average

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/CITS61189.2024.10608029, 2024 International Conference on Computer, Information and Telecommunication Systems (CITS)

squared difference between the desired and received signals. This technique provides enhanced reliability compared to other methods [10].

Over recent years, deep learning has shown an exceptional ability to address irregular network communication issues and facilitate rapid computational processes compared to conventional iterative methods [11]. Additionally, Deep Q-Network (DQN) has the potential to completely change the way beamforming works in 6G communication because it can learn from past mistakes and dynamically improve the distribution of resources in real-time situations [12]. As we explore the continuously expanding realm of 6G communication, a significant development emerges in the combination of HRIS and DQN, which holds the potential for transformational impact. The phenomenon of our new dynamic fusion, referred to as HRISDQN, is evidence of the limitless innovation in wireless communication. The fundamental principle of HRISDQN is rooted in DQN reinforcement learning as an artificial intelligence framework designed to emulate the cognitive capacities of the human brain in beamforming decision-making processes.

Currently, there have been several works investigating the use of DQN with RIS that only reflect the incident signal, as in [11], [13]–[24].

Even though reflective RIS and DQN have been combined, there is still a gap in operating DQN for HRIS. This is because HRIS, which can send and receive the incident signal simultaneously, is being used with DQN for the first time, and its performance will depend on the user and the BS channel. Furthermore, this achievement is justified by different system metric parameters.

Our novel HRISDQN contributes to the 6G communication system with terahertz data rates by:

- We propose an HRIS model to expand the scope of RISs that only reflect signals. Particularly, we consider the practical electromagnetic characteristics of HRIS elements, leading to a combined phase shift in transmission and reflection. Using the suggested model, we formulate a problem solution of joint active and passive beamforming that necessitates hybrid control over phase shift and amplitude. This solution aims to enhance the fairness between users in the network and minimise long-term delays.
- We develop a collaborative HRIS-DQN algorithm as a high-beamforming solution. The joint HRISDQN scheme can handle hybrid control by using two combined net matrices for the BS and user channel for each HRIS element. The HRIS controls whether signals are sent or reflected, and the DQN handles beamforming control.
- A unidirectional neural network is utilised for the prediction of the phase shift matrix and beamforming matrix simultaneously; the proposition of the neural network design is intended to diminish the computational complexity of the optimisation process, which is typically computationally complex in most previous works that use iterative optimisation algorithms to obtain suboptimal solutions.
- The suggested HRISDQN design considers how the best transmit beamforming matrix depends on the useful channel. It does this by finding the best phase shift and beamforming matrices for each HRIS element simultaneously to get the highest sum rate.

The rest of this paper is organised as follows: Section II shows the system model and interpretations of our HRISDQN-assisted wireless network. This section includes the coupled phase-shift, channel, and signal models. The HRIS dual functionality algorithm and the DQN beamforming method are introduced in Section III as solutions that are both simple and effective. In Section IV, the network architecture, training, and validation are explained. Section V provides a performance analysis of the proposed framework and algorithms. Section VI concludes the paper.

II. System Model and Interpretation

We consider a three-point system consisting of one HRIS with M horizontal and L vertical patches forming H reflecting and transmitting surface, BS with A antennas to serve N users, as in Fig.1. To demonstrate the ability of our DQN with the HRIS, we assume that there is no direct link between the BS and users; as a result, DQN beamforming will perfectly serve users on both HRIS sides.

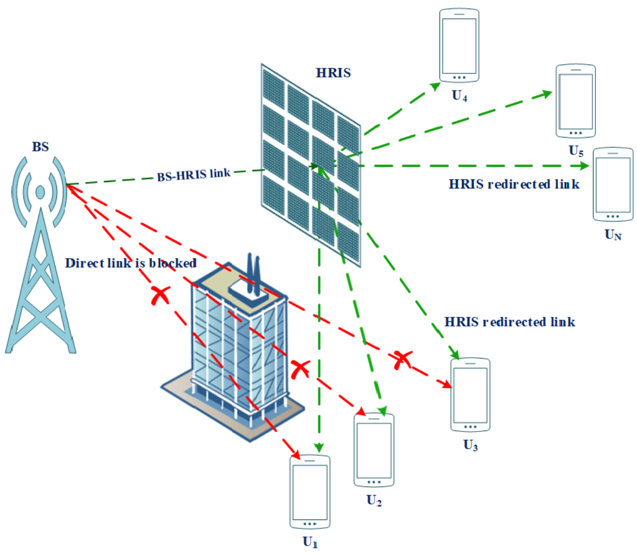

Fig. 1. System model.

The received signal at the N_{th} user is a superposition between the signal from BS to HRIS and the signal from HRIS to the user, which is equal to:

$$Y_n = b_n^H \Theta G_A^H X + w_n \quad (1)$$

In our work, we refer to the complex space vector as S,$b_n^H \in S^{H\times 1}$ denotes the beamforming link between HRIS and user n,$G_A^H \in S^{H\times A}$ denotes the link between HRIS and BS, and

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/CITS61189.2024.10608029, 2024 International Conference on Computer, Information and Telecommunication Systems (CITS)

the hybrid transmission or reflection beamforming matrix is denoted as Θ, where $\Theta = diag\left(\left[e^{j\theta_1}, e^{j\theta_2}, \ldots, e^{j\theta_H}\right]\right)$ and $, \theta_H \in (0, 2\pi)$,$X = O_n s_n$, is the transmitted signal at the BS. $O_n \in S^{A\times 1}$ is the beamforming vector from BS, s_n is the information symbol for the n_{th} user, and w_n is the additive white Gaussian noise (AWGN) at the user n.

In order to determine the n_{th} user received SNR, the vector of beamforming weights for all users is denoted as $O = [o_1, \ldots, o_n]$ where O_n is the weight of user n. The channel matrix for all users in the HRIS channel vector is defined as $B_n^H = \left[b_1^H, b_2^H, \ldots, b_n^H\right]$ and $B = B^H \Theta G_A^H \in S^{N\times A}$, which is the overall channel matrix as obtained by multiplying the HRIS beamforming matrix with HRIS and BS link channel matrix. Then the SNR is defined as:

$$SNR_n = \frac{(o_n B_n)^T (o_n B_n)}{\mathrm{I}_n}, n = 1, 2, \ldots, N \tag{2}$$

where $(.)^T$ denotes the conjugate transpose of the matrix and I_n is the total interference and noise at user n. Maximising the transmit power with the total data rate of users requires considering two important constraints at first, the total transmit power of all users should not exceed a maximum value of P_{max} as in (3), where the second constraint states that the phase shift values should stay between 0 and 2π as in (4):

$$\sum_{n=1}^{N} (O_n)^T O_n \leq \mathrm{P}_{\max} \tag{3}$$

$$0 \leq \theta_n \leq 2\pi \tag{4}$$

From (2), (3) and (4) the beamforming maximisation is calculated as in (5):

$$\max_{\Theta, Q} RT = \log(1 + SNR_n) \tag{5}$$

As designated in (2) and (5), maximising data rate directly relates to beamforming and phase shift matrices. This means that both matrices must have an optimal design to maximise the sum rate. The maximisation involves dividing the main maximisation approach into sub-approaches and optimizing them iteratively. The approaches are then applied through randomisation and normalization quadratically constrained approach, where (5) can be rewritten as:

$$\max_{\bar{\Theta}} \left(\bar{\Theta}\right)^T RT \bar{\Theta} \tag{6}$$

where, $\bar{\Theta} = \begin{bmatrix} \Theta \\ r \end{bmatrix}, RT = \begin{bmatrix} B_n^T G^T B G & B_n^T G^T b_n \\ b_n^T G B & 0 \end{bmatrix}$.

Then, to find the suboptimal solution for (6), it must be normalised as a complex scalar modulus as below:

$$\Theta = Norm\left(\bar{\Theta}\left[1 : H\right] / \bar{\Theta}_{H+1}\right) \tag{7}$$

III. HRISDQN NEW APPROACH

A. HRIS dual functionality and power optimisation

HRIS was designed and validated using a new power optimisation system model for the signal destination at the HRIS controllers. For each user signal, this is determined and calculated using the new general formulation for power optimisation:

$$\mathrm{P}_{r,t} = SP \times \mathrm{A}_e = \frac{|E_r|}{2\mathrm{Z}_{air}}^2 \times U_{AG}^n \frac{\lambda^2}{4\pi} \tag{8}$$

where, E_r is computed using (9):

$$\begin{aligned} &\mathrm{E}_r = \sum_{M=1}^{M} \sum_{L=1}^{L} \sqrt{2\mathrm{Z}_{air} \mathrm{P}_s U_{AG}^n \mathrm{A}_{HRIS} G_{HRIS}} \times NP \times R_{M,L} \\ &and, NP = \frac{\sqrt{NP^{BS} NP^{HRIS,BS} NP^{HRIS,U} NP^U}}{4\pi l_{HRIS}^{BS} l_{HRIS}^{U}} \end{aligned} \tag{9}$$

From (9), the HRIS functionally determination at the HRIS controller for transmission or reflection considers the user location from the HRIS by introducing the parameters D_{ptrans}, D_{pref}, where these newly defined parameters take values 0 or 1 so that equation (8) can be rewritten and calculated using (10):

$$\begin{aligned} &\mathrm{P}_{\max} = \mathrm{P}_t + \mathrm{P}_r \\ &= (SP \times \mathrm{A}_e \times D_{ptrans} \times \mathrm{A}_t) + (SP \times \mathrm{A}_e \times D_{pref} \times \mathrm{A}_r) \\ &where: SP \times \mathrm{A}_e = \frac{|E_r|}{2\mathrm{Z}_{air}}^2 \times U_{AG}^n \frac{\lambda^2}{4\pi} \end{aligned} \tag{10}$$

This power optimisation takes the system to the next step of signal transmission and redirection towards the user location. The suggested implementation determines signal direction and characteristics according to HRIS space wave impedance η_{HRIS}, which is reconfigured and calculated using (11):

$$\eta_{HRIS} = j\frac{\eta_0}{\cos\theta_i} \cot\left(\frac{(\sin\theta_i - \sin\theta_{r,t})}{2}\right) .d_{ue} \tag{11}$$

where θ_i, θ_r and θ_t are, the incident, reflected and transmitted signal phase shifts, respectively. Based on the new power optimisation and decision-making illustrated above, the transmitted or reflected signal from HRIS is calculated using (12) and (13), respectively, where Y_n is the U_n intended signal. Table I summarises HRIS design parameters.

$$Y_t = \mathrm{PA}_t \cdot \eta_{RIS} e^{j\theta_t} Y_n \tag{12}$$

$$Y_r = \mathrm{PA}_r \cdot \eta_{RIS} e^{j\theta_r} Y_n \tag{13}$$

B. DQN beamforming

Motivated by the fact that conventional optimisation methods often result in suboptimal solutions with high complexity, we present a novel reinforcement deep learning-based framework as an alternative approach, providing a more efficient and effective solution to the optimisation problem.

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/CITS61189.2024.10608029, 2024 International Conference on Computer, Information and Telecommunication Systems (CITS)

TABLE I
MODEL DESIGN PARAMETERS

Parameters	Definitions
U_{AG}^{n}	User antenna gain
SP	Power density
Z_{air}	Air impedance
E_r	Electric field
A_e	Effective aperture
A_{HRIS}	HRIS Effective aperture
P_s	Signal power
l_{HRIS}^{U}	Distance between HRIS and user
l_{HRIS}^{BS}	Distance between HRIS and BS
$R_{M,L}$	Reflection/transmission coefficient for which equals $\lvert R_{M,L}\rvert \exp\left(-j\phi_{M,L}\right)$ $\exp\left(-j\frac{2\pi}{\lambda}\left(l_{HRIS}^{BS}+l_{HRIS}^{U}\right)\right)$
NP^{BS}	BS normalised power radiation
NP^{U}	User normalised power radiation
$NP^{HRIS,BS}$	HRIS normalised power radiation from BS
$NP^{HRIS,U}$	RIS normalised power radiation to Un
$A_{r,t}$	Signal amplitude response

This novel approach improves performance and efficiency in HRIS-aided 6G wireless communication systems. Instead of relying solely on the BS or user channels, combining the two is utilised to establish unique characteristics for each transmitting and reflecting element of the HRIS. Thus, using B_n^H and G_A^H as a two-dimensional feature vector allows more streamlined processing and data analysis, in contrast to utilising each vector separately, as a result this will leads to enhanced efficiency. Precisely, the new combined feature for i_{th} HRIS element is described in (14), where $V \in S^{N\times A}$:

$$V_i = B_n^H G_A^H \tag{14}$$

The focal point of the new DQN network is to tackle the problem of the optimal transmit beamforming matrix reliant on the effective channel. In wireless communication systems, the effective channel considers the impact of the environment, for instance, reflections and scattering, between the transmitter and receiver. The optimal transmit beamforming matrix is a matrix that adjusts the phase and amplitude of the transmitted signals to amplify the received signal power at the receiver.

Our novel HRISDQN aims to determine the optimal action by selecting the highest Q-value while factoring in the critical parameters associated with Q-learning. It is designed to repeatedly iterate the Q-value for every observation to identify the maximum value. This process enables the algorithm to make informed decisions that are efficient and optimised for maximum performance.

The designed HRISDQN depends on three main stages as illustrated in Fig.2; in the state stage indicated as (A), the network observes and collects information on the BS channel and user channel, which is identified as a collection of the wireless environment characteristics E; where, $e_n \in E$ is the combined BS and user channel characteristics for each user n, in which E is the current situation for user n and sent to the HRIS controller. The next stage (B) is where the HRIS controller takes the decision e_{n+1} that is made based on the current channel characteristics e_n. The last stage (C) is wireless environment feedback f_n and update on the taken decision, in which the HRISDQN reinforce good behaviour and discourages bad behaviour based on the highest Q-value achieved as in (15), where the second part of the equation can be rewritten as in (16).

$$Q\left(S_n, S_{n+1}\right) \leftarrow Q\left(S_n, S_{n+1}\right) + \alpha Q^*\left(S_n, S_{n+1}\right) \tag{15}$$

$$\begin{aligned} Q\left(S_n, S_{n+1}\right) &\leftarrow Q\left(S_n, S_{n+1}\right) \\ &+\alpha\left\{r+\gamma \max_{S'_{n+1}} Q^*\left(S'_n, S'_{n+1}\right) - Q\left(S_n, S_{n+1}\right)\right\} \end{aligned} \tag{16}$$

where r is the immediate decision obtained by the HRIS controller, α is the learning rate, which determines the extent to which new information overrides old information, Q^* is the optimal Q-value for the next (e_n, e_{n+1}) pair, and γ is the buffer used to reduce the correlation between the training data and enhance the stability of convergence, which determines the importance of future decisions in the update.

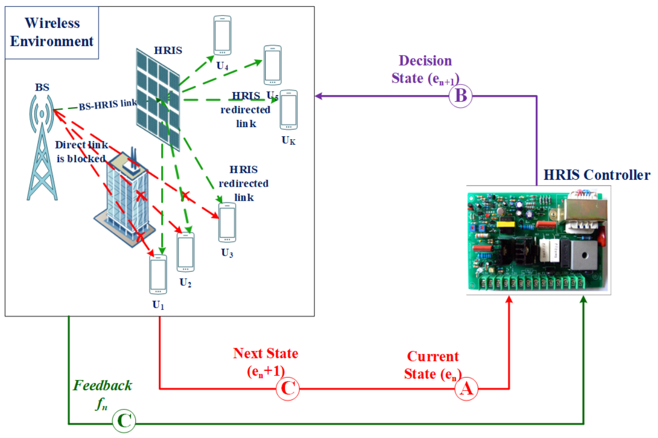

Fig. 2. DQN System model.

IV. HRISDQN ARCHITECTURE AND VALIDATION

A. Network Architecture and Training

The adopted network architecture, as illustrated in Fig. 3, consists of one convolutional layer (CV) followed by six fully connected layers (FC). Each FC layer neuron is proportional to the number of reflecting elements H to ensure that the HRISDQN network has enough capacity to learn from larger datasets as the wireless system scales up. Therefore, the FC layers are made up of $64H, 32H, 16H, 8H, 4H$ and H neurons, respectively. To prevent network overfitting and improve the training process, a Batch Normalization (BN) layer is placed between each FC layer and after the first CV

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/CITS61189.2024.10608029, 2024 International Conference on Computer, Information and Telecommunication Systems (CITS)

layer. All FC layers use a rectified linear unit (ReLU) as an activation function to prevent the vanishing gradient problem.

Adam optimiser is used to train the network with an initial training rate of 0.0001, as the maximum batch size Z is set to 6000, 80% of the generated data samples are used for training, and the remaining 20% are used for validation.

B. Loss Function

The loss function is a crucial metric that determines a given model's performance regarding its training data. It serves as an indicator of the disparity between the predicted values and the actual values. Minimising the loss function is imperative to improving a model's accuracy since this implies that the model's predictions align more with the target values.

The loss function for single beamforming through HRIS-DQN is calculated from (15) and (16) as in (17).

$$Loss = \left(Q\left(S_n, S_{n+1}\right) - \left(r + \gamma \max_{S'_{n+1}} Q^*\left(S'_n, S'_{n+1}\right)\right)\right)^2 \tag{17}$$

Given the anticipated expansion of wireless communication in the 6G THz new generation, which will ultimately support ten times the number of users compared to the current generation, it becomes imperative to establish a comprehensive characterization of the loss function for this particular goal. The overall loss function for HRISDQN is expressed in (18),

$$Loss = \frac{\left(Q\left(S_n, S_{n+1}\right) - \left(r + \gamma \max_{S'_{n+1}} Q^*\left(S'_n, S'_{n+1}\right)\right)\right)^2}{Z} \tag{18}$$

C. Simulation Parameters

The HRIS used in this work is designed to transmit and reflect the incident signal to their intended users at the same time, 200 hybrid reflection and transmission elements H is applied in which their element response equals 0.9, transmission power 1000 mW, wavelength λ of 0.1, HRIS antenna gain is 1 and power radiation parameters are 0.99. The tested and configured area experiences high user density of 3×10^6, 6×10^6, 9×10^6, 12×10^6, and 16×10^6 $user/km^2$ in which they are randomly distributed, with one BS which has eight transmitting antennas A, in which wireless communication system with high user density indicates higher transmission data rates.

To elucidate the performance of HRISDQN, considering system stability, training efficiency, and testing performance, we generate 12×10^6 samples for training, 10×10^6 samples for validation, and 8×10^6 samples for testing. This new HRIS-DQN is designed to support the latest wireless communication generation, 6G; the operating frequencies for THz transmission used to justify HRISDQN performance with the 6G THz new wireless generation are 300, 400, and 500 GHz.

D. Working Methodology

MATLAB is selected as the primary tool for implementing the proposed methodology in designing a novel HRISDQN that will enhance the performance of a wireless multi-user system. This system is characterised by its ability to accommodate multiple users simultaneously. Incorporating HRIS technology is intended to improve the overall efficiency and reliability of the system in question.

The primary stage of the proposed methodology is developing an HRIS-enhanced wireless network. This is achieved by attaining optimal configuration of varied parameters, including wavelength, radiation power, antenna gain, bandwidth, noise power, transmit power, source and destination locations, the number of HRIS components, the number of users, and the number of antennas in the BS. The next stage is to train and validate the suggested HRISDQN network, as defined in the previous subsections, considering data generation, loss function application, and the data standardisation process by dividing the data average by the standard deviation.

The third stage of the work methodology involves implementing a real-time scenario that entails randomly allocating users for signal beamforming through HRIS elements. This approach is designed to provide a comprehensive elucidation of the new beamforming method's achievement in contrast to conventional methods. As such, the results are carefully plotted and compared to determine the effectiveness of the new approach.

V. SIMULATION RESULTS AND DISCUSSION

In this section, we show numerical results that are performed to evaluate the proposed approach. Our new HRISDQN is validated against different beamforming techniques in single and multi-user scenarios. Conventional Semi Definite Relaxation (SDR), Maximum Ratio Transmission (MRT), and Minimum Mean Square Error (MMSE) are methods used to optimise the transmission and reception of signals in multiple-antenna communication systems. Each has advantages and disadvantages regarding difficulty, interference, channel estimation error, data rates, and noise reduction.

A benchmark assessment measure is system fairness. Our research shows that as the number of users increases, HRIS-DQN displays exceptional fairness results. The adaptive resource allocation and learning characteristics of HRISDQN enable it to dynamically alter signal beamforming to ensure equal resource distribution across users.

At the beginning of the simulation, where the number of users is 3×10^6, HRISDQN elucidates better fairness performance than the other approaches, which is higher than SDR, MRT, and MMSE by 30% to 40%. On the other hand, SDR, MRT, and MMSE demonstrated low adaptability for signal beamforming. As the number of users increased, their performance would not exceed 50% at the maximum user density reaches 16×10^6 for 500 GHz, while HRISDQN shows 90% system fairness. We discovered that fairness became a more significant concern as the number of users in the wireless communication scenario increased. Jain's Fairness Index (JFI)

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/CITS61189.2024.10608029, 2024 International Conference on Computer, Information and Telecommunication Systems (CITS)

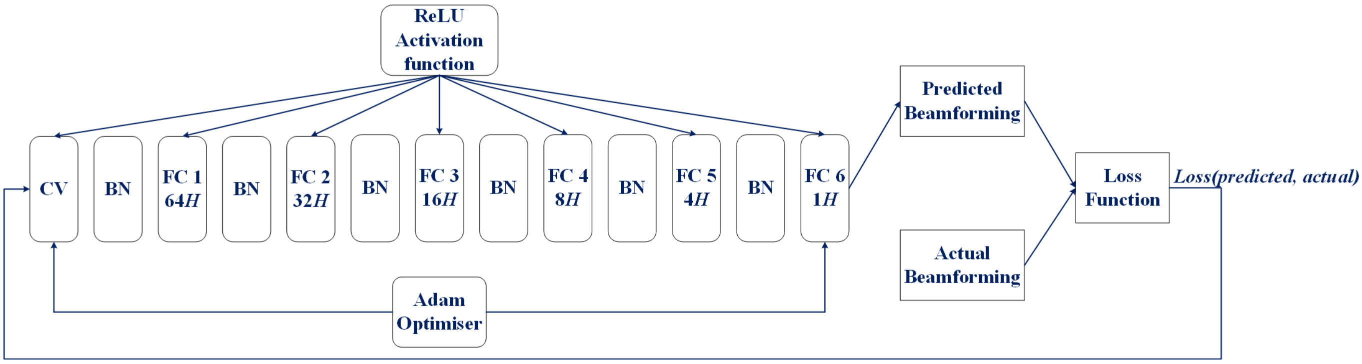

Fig. 3. DQN Network architecture.

measures fairness by determining how users allocate resources. We observed a consistent trend, as shown in Fig.4, 5 and 6, where fairness indices tended to decrease for SDR, MRT, and MMSE as the number of users increased. This trend is typical in scenarios involving a shared wireless medium where resources must be divided among many users.

As we employed JFI as fairness measures, which range between 0 and 1, HRISDQN preserved its performance against network circumstances changes and user dynamics and proved to achieve a substantial advantage, maintaining user fairness of 85% considering different user densities.

HRISDQN's flexibility allowed it to dynamically adjust its resource allocation strategy, which leads to a more even distribution of resources. Its adaptability was particularly useful when the number of users was changing rapidly. In contrast to traditional beamforming methods, which may achieve initial fairness but struggle to maintain it over time, HRISDQN showed the ability to learn and adapt continuously, resulting in more sustainable levels of fairness. These findings are significant for developing 6G wireless communication scenarios, where the demand for equitable resource allocation is expected to be even higher than in current 5G networks.

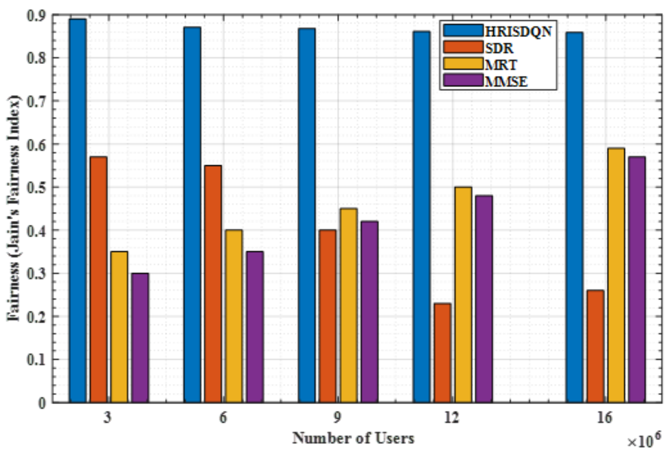

Fig. 4. Fairness Comparison for 500GHz.

An important measure for a communication system is a delay, as illustrated in Fig.7; the delay experienced by users during communication is heavily influenced by the density of users in the network.

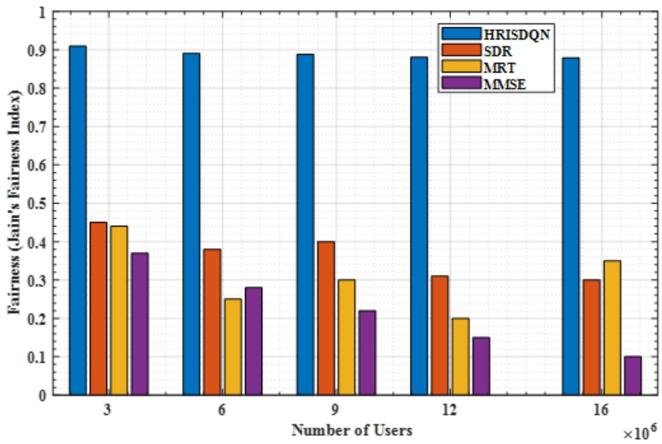

Fig. 5. Fairness Comparison for 400GHz.

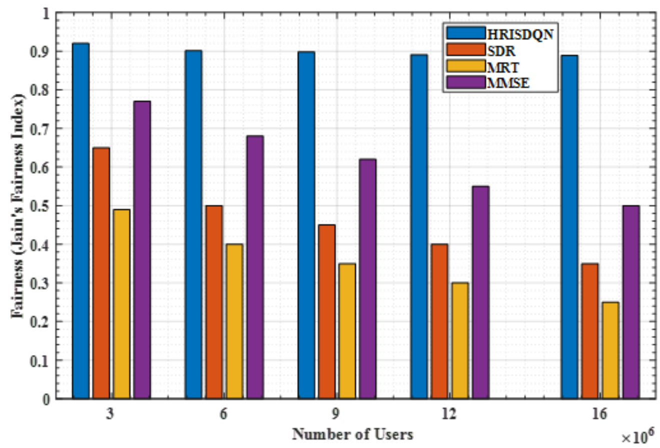

Fig. 6. Fairness Comparison for 300GHz.

All techniques exhibit lower delays in a low-density network with 3×10^6 users. However, as the density increases, a noticeable different trend emerges. HRISDQN consistently

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/CITS61189.2024.10608029, 2024 International Conference on Computer, Information and Telecommunication Systems (CITS)

outperforms traditional beamforming techniques in terms of reducing latency. Its dynamic resource allocation and learning-based decision-making enable HRISDQN to quickly adapt to changing conditions and priorities for low-latency communication. A comparison between HRISDQN and traditional beamforming techniques highlights a significant advantage of HRISDQN in minimising communication delay. In situations with moderate to high user density levels in which the user's density is equal to 12×10^6 users and 16×10^6 users, HRISDQN achieves considerably lower delays than SDR, MRT, and MMSE.

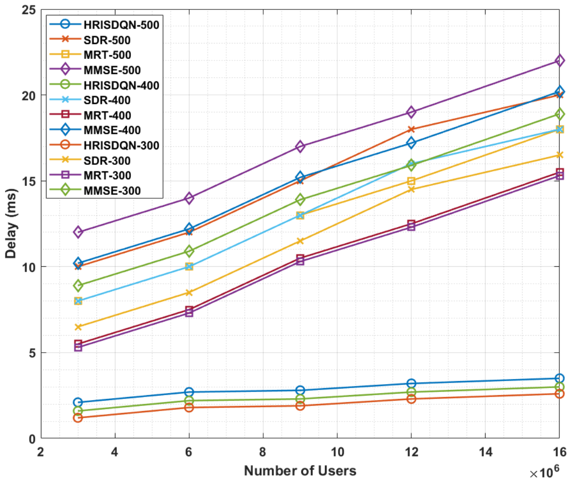

Fig. 7. Delay Comparison for different frequencies.

This difference is particularly pronounced when the SNR level is high. Traditional techniques use more time to manage the network resources efficiently, leading to more delays as user numbers grow. HRISDQN, however, excels in reducing latency even in congested environments.

The SE analysis findings, as in Figs.8, 9, and 10 expose captivating patterns at different SNR stages and beamforming methods for 6G frequencies. SE, quantified by terabits per second per hertz (Tbps/Hz), is a crucial gauge for a system's capacity to optimise data throughput by efficiently employing the available spectrum resources. SE is a significant measure of the best possible spectrum used for data transmission. HRISDQN outperforms other approaches and maintains high-performance levels as user density increases. As different sets of 6G frequencies are tested, 14 Tbps/Hz gain is achieved at higher SNR levels, with a maximum of 16×10^6 users compared to SDR. This achievement reaches 11 Tbps/Hz compared to MRT and MMSE.

HRISDQN's ability to optimise resource allocation, priorities error resilience in low SNR conditions, and maximise data rates in high SNR environments are the reasons for its adaptability. This adaptability is particularly valuable for 6G applications requiring reliable communication in low SNR scenarios, such as IoT devices and critical infrastructure. HRISDQN's adaptability can lead to a more efficient allocation of resources in multi-user systems, improving overall network performance. Although widely used, SDR shows lower SE than HRISDQN, particularly in challenging SNR conditions. MRT and MMSE have competitive performance but lack HRISDQN's dynamic adaptation, resulting in considerably lower SE. HRISDQN's superior SE across diverse SNR conditions highlights the value of adaptive resource allocation in 6G wireless communication. HRISDQN can learn and adapt its beamforming strategies, maximising data rates while maintaining reliability.

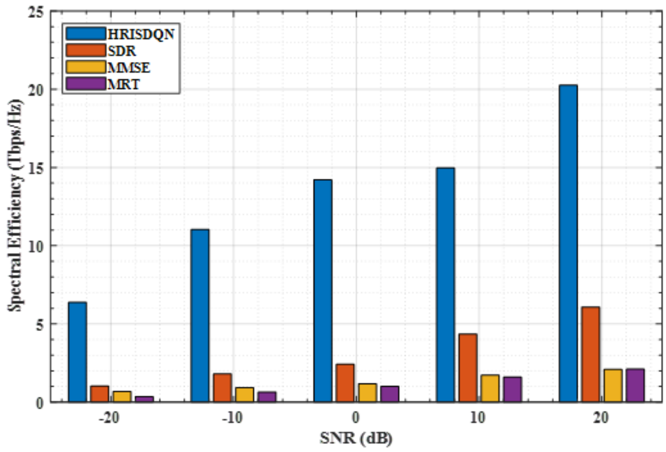

Fig. 8. Spectral Efficiency Comparison for (16×10^6 users, 500 GHz).

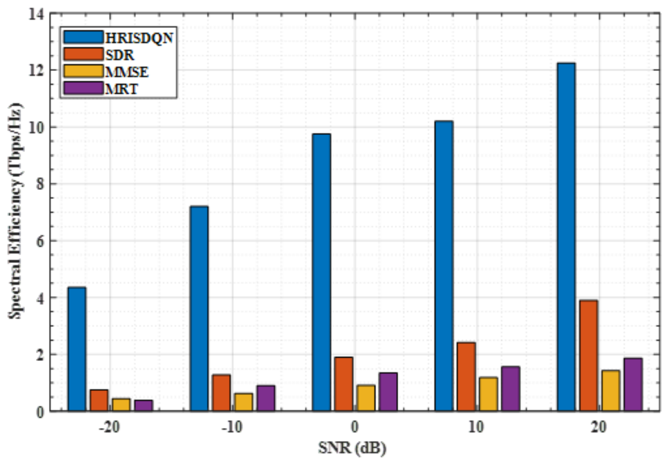

Fig. 9. Spectral Efficiency Comparison for (16×10^6 users, 400 GHz).

VI. Conclusion

HRISDQN's adaptive capabilities position it as a promising tool for improving spectral efficiency in wireless communication systems. Its ability to excel across various SNR conditions suggests potential benefits for multiple IoT applications and 6G. The findings emphasize the importance of adaptive solutions in addressing the complexities of modern wireless networks, where dynamic resource allocation can significantly enhance performance and spectral efficiency. The

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/CITS61189.2024.10608029, 2024 International Conference on Computer, Information and Telecommunication Systems (CITS)

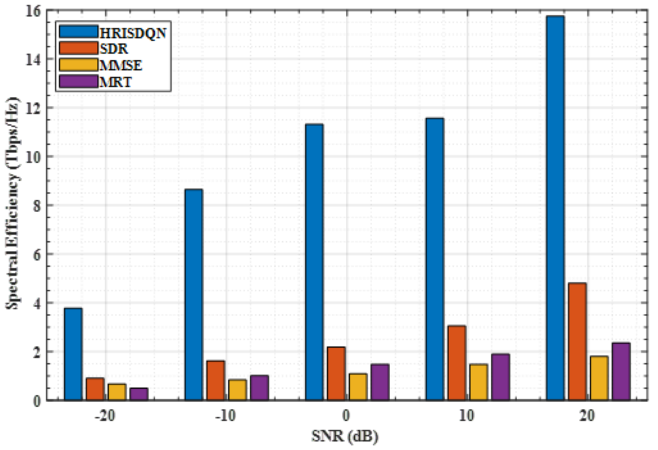

Fig. 10. Spectral Efficiency Comparison for (16×10^6 users, 300 GHz).

adaptability, resilience, and knowledge acquisition capabilities of HRISDQN make it a promising approach for ensuring equitable distribution of resources in future networks. While traditional beamforming methods have their advantages, the effectiveness of HRISDQN strengthens its suitability for the dynamic requirements of 6G wireless communication systems.

The findings emphasise HRISDQN's potential as a powerful tool for addressing spectral efficiency, fairness, and delay concerns in 6G wireless communication. This makes it a promising contender for resource distribution in high-user density environments and complex interference patterns. Further research and testing can better understand the entire scope of HRISDQN's capabilities in promoting fairness across various communication contexts. These findings highlight the potential of HRISDQN as a pivotal technology in augmenting the effectiveness and dependability of wireless communication in forthcoming 6G networks. These networks are anticipated to encounter intricate hurdles in the form of user density and dynamic channel circumstances.

References

[1] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6g," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 118–125, 2020.

[2] M. A. ElMossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, "Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 990–1002, 2020.

[3] J. Ye, S. Guo, S. Dang, B. Shihada, and M.-S. Alouini, "On the capacity of reconfigurable intelligent surface assisted mimo symbiotic communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1943–1959, 2022.

[4] Z. Zhang, C. Zhang, C. Jiang, F. Jia, J. Ge, and F. Gong, "Improving physical layer security for reconfigurable intelligent surface aided noma 6g networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 4451–4463, 2021.

[5] S. Basharat, S. A. Hassan, H. Pervaiz, A. Mahmood, Z. Ding, and M. Gidlund, "Reconfigurable intelligent surfaces: Potentials, applications, and challenges for 6g wireless networks," *IEEE Wireless Communications*, vol. 28, no. 6, pp. 184–191, 2021.

[6] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 421–439, 2021.

[7] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3313–3351, 2021.

[8] J. Fink, R. L. Cavalcante, and S. Stanczak, "Multicast beamforming using semidefinite relaxation and bounded perturbation resilience," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 4749–4753.

[9] H. Huang, X. Wang, C. Zhang, K. Qiu, and Z. Han, "Reward-maximization-based passive beamforming for multi-ris-aided multi-user miso systems," in *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE Press, 2021, p. 01–06. [Online]. Available: https://doi.org/10.1109/GLOBECOM46510.2021.9685372

[10] X. Cao, Y. Zhang, and P. Du, "Robust mmse beamforming of downlink multi-user miso systems in the presence of beam pointing error," in *2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2021, pp. 1–5.

[11] I. Ahmad, R. Narmeen, Z. Becvar, and I. Guvenc, "Machine learning-based beamforming for unmanned aerial vehicles equipped with reconfigurable intelligent surfaces," *IEEE Wireless Communications*, vol. 29, no. 4, pp. 32–38, 2022.

[12] P. Susarla, Y. Deng, M. Juntti, and O. Sílven, "Hierarchial-dqn position-aided beamforming for uplink mmwave cellular-connected uavs," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 2022, pp. 1308–1313.

[13] A. Faisal, I. Al-Nahhal, O. A. Dobre, and T. M. N. Ngatched, "Distributed ris-assisted fd systems with discrete phase shifts: A reinforcement learning approach," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 2022, pp. 5862–5867.

[14] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "Deep reinforcement learning based beamforming codebook design for ris-aided mmwave systems," in *2023 IEEE 20th Consumer Communications and Networking Conference (CCNC)*, 2023, pp. 1020–1026.

[15] C. Huang, Z. Yang, G. C. Alexandropoulos, K. Xiong, L. Wei, C. Yuen, Z. Zhang, and M. Debbah, "Multi-hop ris-empowered terahertz communications: A drl-based hybrid beamforming design," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1663–1677, 2021.

[16] Y. Ge and J. Fan, "Beamforming optimization for intelligent reflecting surface assisted miso: A deep transfer learning approach," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3902–3907, 2021.

[17] M. Liu, R. Wang, Z. Xing, and J. Yu, "Complex-valued reinforcement learning based dynamic beamforming design for irs aided time-varying downlink channel," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–5.

[18] R. Zhong, Y. Liu, X. Mu, Y. Chen, X. Wang, and L. Hanzo, "Hybrid reinforcement learning for star-riss: A coupled phase-shift model based beamformer," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 9, pp. 2556–2569, 2022.

[19] X. Liu, Y. Liu, and Y. Chen, "Machine learning empowered trajectory and passive beamforming design in uav-ris wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 7, pp. 2042–2055, 2021.

[20] K. Li, C. Huang, Y. Gong, and G. Chen, "Double deep learning for joint phase-shift and beamforming based on cascaded channels in ris-assisted mimo networks," *IEEE Wireless Communications Letters*, vol. 12, no. 4, pp. 659–663, 2023.

[21] H. Song, M. Zhang, J. Gao, and C. Zhong, "Unsupervised learning-based joint active and passive beamforming design for reconfigurable intelligent surfaces aided wireless networks," *IEEE Communications Letters*, vol. 25, no. 3, pp. 892–896, 2021.

[22] A. Abbasi and H. Liu, "Oneai - novel multipurpose deep learning algorithms for uwb wireless networks," in *2021 IEEE Workshop on Signal Processing Systems (SiPS)*, 2021, pp. 129–133.

[23] V. D. Potu, V. S. Sunku, M. M., and P. S. B. M., "A deep learning scheme for integrated active and passive beamforming in reconfigurable intelligent surface aided wireless miso networks," in *2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, 2022, pp. 226–230.

[24] A. I.AL-Shaeli, S.Hburi., "Reconfigurable intelligent surface passive beamforming enhancement using unsupervised learningintisar," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 1, pp. 493–501, 2023.