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: DOI: 10.1109/CITS70307.2026.11637274, 2026 International Conference on Computer, Information and Telecommunication Systems (CITS)

Graph-Centric Deep Q-Learning for Interference-Aware Resource Allocation in RSMA-Enabled 5G Slicing

Aya Kh. Ahmed * *(Member, IEEE)*, Nadia Al-Aboody†, and Hamed S. Al-Raweshidy* *(Senior, IEEE)*
**Department of Engineering, College of Engineering, Design, and Physical Science,*
Brunel University of London, UK. (Email: Aya.Ahmed@brunel.ac.uk, Hamed.Al-Raweshidy@brunel.ac.uk).
†*Department of Computer Network Techniques, Southern Technical University, Iraq.(Email: nadia.alaboody@stu.edu.iq).*

***Abstract*—The emergence of 5G and 6G advanced ecosystems demands highly adaptive resource management to orchestrate the specialised requirements of eMBB, URLLC, and mMTC network slices. In dense multi-cell environments, capturing complex spatial interdependencies and mitigating dynamic interference is paramount for maintaining Quality of Service (QoS). This paper introduces a robust GNN-DQN framework designed for Rate Splitting Multiple Access (RSMA) based networks. By representing the network topology as a graph, the framework leverages Graph Neural Networks (GNNs) to extract high-dimensional spatial features and model inter-cell interference patterns. These insights enable a Deep Q-Network (DQN) agent to perform intelligent resource partitioning and dynamic power splitting of the RSMA common stream. Experimental results demonstrate that the proposed GNN-DQN framework achieves a connectivity success ratio exceeding 90% across all slices, representing an average improvement of over 60% compared to non-graph-based reinforcement learning and supervised baselines. Notably, the framework demonstrates exceptional spectral efficiency, maintaining near-total connectivity while utilising less than 10% of the normalised system bandwidth, a 4× reduction in resource overhead compared to traditional methods. Furthermore, the GNN-driven architecture ensures stable convergence during training, yielding a 1.6× higher system reward score. Our findings validate GNN-DQN as a high-performance, scalable, and resource-efficient paradigm for intelligent orchestration in 5G and 6G networks.**

I. INTRODUCTION

A. Motivation and Background

The transition to 5G and beyond (B5G) is shifting the industry from a uniform architecture to a network-slicing, service-oriented model. This enables one physical infrastructure to support three distinct service types: enhanced Mobile Broadband (eMBB) for high data rates; Ultra-Reliable Low-Latency Communications (URLLC) for minimal latency and high reliability; and massive Machine-Type Communications (mMTC) for dense, low-power device support [1].

Managing these slices is a major challenge because their requirements often conflict with one another. For example, allocating massive bandwidth to eMBB users causes interference and higher latency for URLLC users, whose operations depend on minimal delays and reliable transmission. Meanwhile, the high volume of mMTC devices can saturate standard scheduling systems, leading to congestion and reduced performance for all slices [2].

Traditional allocation methods, such as Non Orthogonal Multiple Access (NOMA) or basic proportional fairness, fail to adapt to these conflicting demands and the rapidly changing multi-cell environment.

To overcome these barriers, Rate-Splitting Multiple Access (RSMA) has emerged as a promising physical-layer technology [3]. Unlike NOMA or Space Division Multiple Access (SDMA), RSMA splits user messages into common and private parts, offering a more flexible framework for managing interference and maximising spectral efficiency in multi-antenna systems. However, the complexity of optimising RSMA parameters in real-time requires advanced computational intelligence. Recent literature, including but not limited to [1], [4]–[8], suggests that while standard artificial intelligence models can assist, they have limitations:

1) Deep Q-Networks (DQN) are effective for decision-making but often treat base stations as isolated entities, failing to capture the spatial graph structure of the network.
2) Convolutional Neural Networks (CNN) excel at fixed feature extraction but lack the flexibility to handle the varying topology and mobility of 5G users.

In this paper, we address these gaps by proposing a Hybrid GNN-DQN framework. By utilising Graph Neural Networks (GNN), our model explicitly captures the spatial interdependencies between base stations and users, treating the entire network as a dynamic graph. This spatial intelligence is then utilised by a DQN agent to perform interference-aware resource partitioning, ensuring that eMBB and URLLC demands are met with up to 90% satisfaction while maximising the connectivity and spectral efficiency of the mMTC slice.

B. Literature Summary and Related Works

The orchestration of 5G/6G network slicing through RSMA has emerged as a pivotal research area. While seminal works have addressed resource partitioning, the simultaneous management of eMBB, URLLC, and mMTC slices in dense multi-cell environments remains a multi-objective optimisation challenge.

1) RSMA and Reinforcement Learning Slicing: Recent literature highlights RSMA as a superior alternative to NOMA and OMA due to its flexibility in decoding common and private streams. The authors in [1] demonstrated a DRL agent capable of switching between multiple access schemes to maximise sum rates in two-slice scenarios. The integration of DRL for hybrid services was further advanced in [9], where a 2IADDPG algorithm optimised user grouping by splitting data into common mMTC and private URLLC flows. Similarly, [10] introduced a Duelling DQN framework with reward-clipping to enhance training stability in eMBB and URLLC scheduling. To improve adaptability in dynamic environments, [11] proposed a Prediction-Aided Weighted DRL (PW-DRL) scheme to capture temporal correlations between current and future network states. While these approaches address temporal dynamics and stability, they typically treat base stations as isolated entities and focus primarily on temporal inference rather than on the spatial interference graph, a gap that our GNN-DQN framework aims to bridge.

2) Machine Learning and Graph-Based Modelling: Traditional supervised learning models, such as CNNs and MLPs, have been widely employed for traffic prediction but suffer from a lack of adaptability to non-stationary network conditions. To overcome these limitations, GNNs have been introduced to model the network as a dynamic graph. In [12], the GNNetSlice model leveraged GNNs to predict performance in core infrastructures. Furthermore, [13] proposed a GCN-DDPG (G-DDPG) algorithm for end-to-end network slicing deployment, focusing on Virtual Network Function (VNF) placement and link resource constraints. Despite the success of GNNs in modelling virtualised core infrastructures and VNF chaining, their application at the wireless physical layer, specifically for optimising the RSMA common stream splitting, remains largely unexplored. While GNNs are designed to capture physical topology, other advanced architectures have explored sequential dependencies. For instance, [14] proposed a native AI architecture for end-to-end slicing leveraging Transformer-based DRL to enable zero-touch slice placement, focusing on the balance between slice acceptance and energy efficiency. Despite the high-performance attention mechanisms of Transformers in modelling complex global relationships, their application has primarily concentrated on end-to-end slice placement rather than the localised spatial interference mitigation required for RSMA. To address the scalability issues of DRL in multi-slice environments, [15] proposed a multi-agent DQN framework utilising advanced experience replay mechanisms to handle high action-space complexity and improve data efficiency. While their work significantly accelerates convergence for complex VNF chain topologies, it relies on queuing network models that primarily capture temporal delay rather than the physical-layer spatial interference characteristic of RSMA-enabled multi-cell networks.

3) The Research Gap: Hybrid GNN-DQN for RSMA: A critical gap exists in the current literature on integrating graph-centric spatial intelligence with decision-making agents for multi-slice RSMA orchestration. Existing studies either focus on single-metric optimisations, neglect the spatial conflict graph, or are limited to two-slice scenarios. This paper addresses this gap by proposing a hybrid GNN-DQN framework that explicitly models inter-cell interference to ensure satisfaction across all three primary 5G slices simultaneously.

C. Proposed Concept and Contributions

This study introduces an intelligent Graph-based Reinforcement Learning framework that integrates GNNs with DQNs to maximise spectral efficiency and ensure strict service satisfaction in RSMA-enabled 5G/B5G networks. The contributions of this work are twofold:

- **Novel Spatial-Aware GNN-DQN Architecture:** We propose a graph-centric reinforcement learning scheme that utilises GNNs to extract high-dimensional spatial features from the network's adjacency matrix. Unlike standard DRL (e.g., Duelling-DQN or PW-DRL), our approach treats base stations and users as nodes in a dynamic graph. This enables the agent to learn complex intercell interference patterns and to perform optimal power allocation for RSMA common and private streams.
- **Holistic Three-Slice Orchestration:** Unlike existing works that focus on two-slice scenarios (e.g., eMBB+URLLC), our framework enables dynamic resource partitioning across eMBB, URLLC, and mMTC slices simultaneously. By leveraging the GNN's spatial awareness, the agent maintains high connectivity for mMTC devices even under near-saturation loads without compromising the satisfaction requirement for URLLC and eMBB services.

D. Paper Organisation

The rest of the paper is organised as follows. Section II presents the system model and the technical architecture of the GNN-DQN framework. Section III details the problem formulation, the reward function design, and the proposed optimisation algorithm. In Section IV, a comprehensive performance analysis is provided, including the mathematical modelling of RSMA-based SINR and outage probabilities under Rayleigh fading. Section V presents the simulation results and a detailed discussion on connectivity, spectral efficiency, and training convergence. Finally, Section VI concludes the paper and suggests directions for future work.

II. SYSTEM MODEL: GNN-DQN ARCHITECTURE

The network is modelled as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ represents the set of N base stations (BSs) and $\mathcal{E}$ represents the communication links characterised by inter-cell interference.

As illustrated in Fig. 1, we consider a multi-cell network consisting of $N = 9$ base stations arranged in a grid topology. The users for each slice (eMBB, URLLC, and mMTC) are randomly distributed across the coverage area. This physical arrangement is used to construct the adjacency matrix A for the GNN, capturing the inter-cell interference relationships.

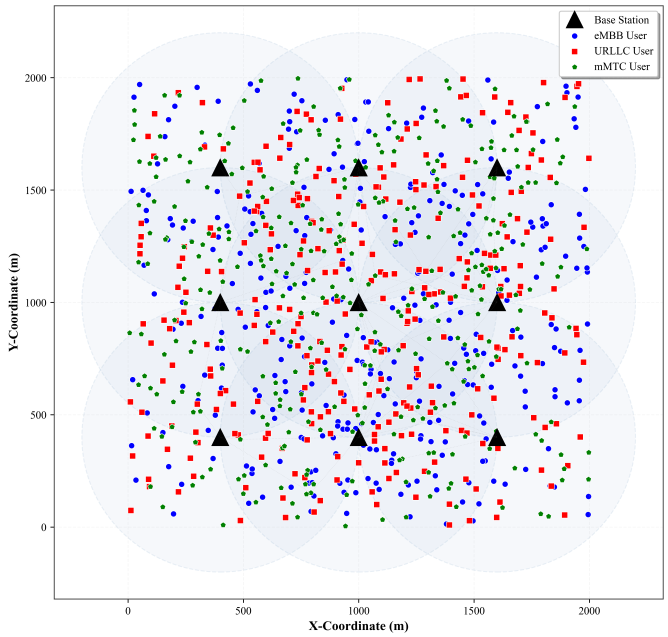

Fig. 1. Network Topology and Spatial Distribution of Slicing Users

A. Graph Representation and Adjacency Matrix

To capture the spatial relationships, we define the Adjacency Matrix $A \in \{0,1\}^{N\times N}$. An edge $e_{ij} = 1$ exists if BS_j is within the interference range of BS_i. The node feature matrix $X \in \mathbb{R}^{N\times F}$ contains local state information for each BS, such as the current number of active eMBB, URLLC, and mMTC users.

B. GNN Message Passing Layer

The GNN processes the graph by aggregating information from neighbouring base stations. This allows each BS to be aware of its neighbours' interference and load status. The hidden state $h_i^{(l+1)}$ for BS_i at layer $l+1$ is updated using a message-passing mechanism:

$$h_i^{(l+1)} = \sigma\left(W^{(l)} \cdot \text{AGG}\left(\{h_j^{(l)} : j \in \mathcal{N}(i) \cup \{i\}\}\right)\right) \quad (1)$$

Where:

- $\mathcal{N}(i)$ is the set of neighbors of BS i defined by the adjacency matrix A.
- $W^{(l)}$ is the trainable weight matrix for layer l.
- $\text{AGG}(\cdot)$ is the mean aggregation function that combines neighbour features.
- σ is a non-linear activation function, ReLU.

C. DQN State-Action Mapping

The final spatial embedding $H^{(L)}$ from the GNN is flattened and concatenated with the global network state to form the input for the Deep Q-Network [16]. The DQN approximates the optimal action-value function $Q(s,a)$, which represents the expected cumulative reward for taking action a (resource partitioning) in state s.The target Q-value is calculated using the Bellman equation:

$$Q_{target}(s,a) = r + \gamma \max_{a'} Q(s', a'; \theta^-) \quad (2)$$

Where:

- r is the reward defined in the previous section.
- γ is the discount factor.
- θ^- represents the weights of the target network to ensure training stability.

D. RSMA Resource Partitioning

The DQN's output determines the resource allocation for the RSMA's common and private streams [17]. The achievable rate for the common stream at user k in BS i is modelled as:

$$R_{i,k}^c = B\log_2\left(1 + \frac{P_i^c|h_{i,k}|^2}{\sum_{j=1}^{K} P_{i,j}^p|h_{i,k}|^2 + I_{inter} + \sigma^2}\right) \quad (3)$$

The agent optimises the powers P_i^c and $P_{i,j}^p$ to ensure that the sum rate of the common stream exceeds the requirements of the prioritised slices (URLLC and eMBB).

III. Problem Formulation and Algorithm

A. Problem Formulation

The goal is to maximise a joint utility function that satisfies the high-throughput requirements of eMBB, the stringent latency/reliability requirements of URLLC, and the massive connectivity requirements of mMTC. We define the optimisation problem as:

$$\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R_t(s_t, a_t)\right] \quad (4)$$

Subject to:

1) URLLC Reliability: $\Pr(Latency > \tau_{max}) \le \epsilon$, where τ_{max} is the maximum allowable delay.
2) eMBB Throughput: $R_{i,k} \ge R_{target}$ for all high-bandwidth users to ensure 100% service satisfaction.
3) mMTC Connectivity: $\sum_{m=1}^{M} \mathbb{1}(\text{SINR}_m \ge \gamma_{th}) \ge K_{min}$, maximizing the number of active IoT devices M that meet the threshold γ_{th}.
4) RSMA Power Constraint: $P_{common} + \sum P_{private} \le P_{total}$ for each base station.

B. GNN-DQN Training Algorithm

The algorithm integrates GNN-based spatial feature extraction with DQN-based temporal decision-making. The GNN processes the adjacency matrix A to create an interference-aware state representation. The algorithm employs a GNN to model interference from massive mMTC deployment on the "Gold" services (eMBB/URLLC).

Algorithm 1 GNN-DQN Resource Allocation for RSMA Slicing

Initialize: Q-network θ, target network θ^{-}, and experience replay $\mathcal{D}$.
Graph Construction: Define $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ based on BS topology and user coordinates.
for episode $m = 1, \ldots, M$ **do**
  Observe initial state $s_0 \in \mathcal{S}$ (Channel gains $\mathbf{H}$ and QoS demands $\mathcal{Q}$).
  for $t = 1, \ldots, T$ **do**
    Feature Extraction: Generate graph embeddings $Z_t = \text{GNN}(\mathcal{G}, s_t; \phi)$.
    Decision Making: Select action $a_t = \{\mathbf{P}_t, \mathbf{B}_t\}$ using ϵ-greedy policy w.r.t. $Q(Z_t, a; \theta)$.
    Execution:
      Perform RSMA SIC decoding and calculate SINR $\gamma_{k,c}, \gamma_{k,p}$.
      Compute slice satisfaction rates Φ_e, Φ_u, Φ_m.
    Reward Generation:
      $r_t \leftarrow \sum_{s \in \{e,u,m\}} \omega_s \Phi_s - \lambda \mathcal{P}_{drop}$.
    Optimization:
      Store transition (s_t, a_t, r_t, s_{t+1}) in $\mathcal{D}$.
      Sample mini-batch and update θ by minimizing:
      $\mathcal{L}(\theta) = \mathbb{E}[(y_t - Q(Z_t, a_t; \theta))^2]$.
    Update: $\theta^{-} \leftarrow \theta$ every C steps.
  end for
end for

C. Simulation Setup and Parameters

The simulation environment is designed to represent a dense urban 5G deployment comprising 9 base stations (BSs) arranged in a 3×3 grid topology. The choice of 9 BSs serves as an optimal balance between network scale and computational tractability. Given the joint optimisation of slice-specific quality of service and RSMA power splitting (α), a 3×3 grid allows for the exploration of complex inter-cell dependencies while maintaining a stable training convergence for the DQN agent, as evidenced by the $1.6\times$ higher system reward score reported. To ensure the reproducibility of the results, Table I summarises the key simulation parameters, aligned with 3GPP Release 16/17 standards for sub-6 GHz deployments. The slice bandwidth partitioning follows the standard industry heuristic (50% eMBB, 30% URLLC, 20% mMTC) to reflect the high-capacity demands of broadband users versus the reliability-centric and connectivity-centric demands of mission-critical and IoT services, respectively [18]. Unlike traditional static allocation methods, the RSMA common stream power ratio, α, is dynamically optimised within the range $[0.2, 0.65]$ by the proposed GNN-DQN agent. The lower bound of $\alpha = 0.2$ is established to ensure the Common Stream maintains a sufficient Signal-to-Interference-plus-Noise Ratio (SINR) for successful Successive Interference Cancellation (SIC) at all user equipments (UEs). Allocating less than 20% of power to the common message typically results in decoding failures, especially for cell-edge users with degraded channel conditions. On the other hand, the upper bound of $\alpha = 0.65$ prevents 'power starvation' of the private streams. Since the private streams are responsible for meeting the specific high-throughput requirements of eMBB slices, capping the common stream power ensures that the remaining 35% of power is reserved for user-specific beamforming and data transmission, thereby maintaining the system's multiplexing gain. This enables adaptive interference mitigation based on a real-time spatial conflict graph derived from node features and the network topology's adjacency matrix.

TABLE I
SYSTEM SIMULATION PARAMETERS

Parameter	Value
Number of Base Stations (BS)	9 (3×3 Grid Topology)
Center Frequency (f_c)	3.5 GHz (n78 Band)
Total System Bandwidth (B_{total})	100 MHz
Noise Power Spectral Density	-174 dBm/Hz
Slice Bandwidth Partitioning	50% (e), 30% (u), 20% (m)
Maximum Transmit Power (P_{max})	46 dBm (40 W)
RSMA Common Stream Ratio (α)	$0.2 - 0.65$ (GNN-DRL Optimized)
User Distribution (K)	500 – 2000 (Non-uniform)
Path Loss Model	$128.1 + 37.6 \log_{10}(d)$ (3GPP RMa)
GNN Architecture	2 Message Passing Layers (64 units)
DQN Learning Rate (η)	10^{-4}
Discount Factor (γ)	0.99
Activation Function	ReLU (Hidden), Linear (Output)

IV. Performance Analysis: Slice Satisfaction and Reliability

This section provides the mathematical framework for evaluating the performance of the proposed GNN-DQN framework in terms of slice-specific satisfaction and connectivity. The performance metrics used in this work are specifically chosen to reflect the diverse Service Level Agreements (SLAs) and Quality of Service (QoS) targets defined for 5G and 6G networks in ETSI TS 122 261 [18]. For the URLLC slice, we use outage probability ($P_{out,u}$), which measures the chance that the communication fails to meet a required quality, ensuring the very stringent reliability threshold of 10^{-5} is maintained even when the radio channel varies randomly (stochastic Rayleigh fading). This approach better highlights critical communication failures than average performance metrics do. For the eMBB slice, spectral efficiency (η_{eMBB}) and sum-rate indicate how efficiently the available bandwidth is used and the total data delivered. These metrics capture the high-capacity needs enabled by RSMA, which allows multiple users to share communication resources concurrently. For mMTC, the connectivity ratio (Φ_m) indicates how many devices can be supported simultaneously, which is more important than individual device speed in use cases such as IoT. Lastly, all these separate key performance indicators are combined into a single Global Reliability Index (Ψ), a summary measure that helps the GNN/DQN agent make balanced decisions among data rate, reliability, and device connectivity when resources are shared.

A. RSMA SINR Formulation for Multi-Slice Connectivity

To evaluate slice satisfaction, we define SINR for both common and private streams. Let P_i^c and $P_{i,k}^p$ be the power allocated by the GNN-DQN agent to the common and private messages at BS_i for user k.

1) Common Message SINR: The common message is intended to be decoded by all users in the cluster. To ensure successful Successive Interference Cancellation (SIC), the SINR for the common part at user k must account for the interference from all private streams:

$$\gamma_{k,c} = \frac{P_i^c |h_{i,k}|^2}{\sum_{j=1}^{K} P_{i,j}^p |h_{i,k}|^2 + I_{inter} + \sigma^2} \tag{5}$$

where I_{inter} represents the inter-cell interference captured by the GNN's adjacency matrix and σ^2 is the noise floor.

2) Private Message SINR: After the common message is successfully decoded and subtracted from the received signal, the SINR for the private message of user k is given by:

$$\gamma_{k,p} = \frac{P_{i,k}^p |h_{i,k}|^2}{\sum_{j=1,j\neq k}^{K} P_{i,j}^p |h_{i,k}|^2 + I_{inter} + \sigma^2} \tag{6}$$

3) Slice-Specific Outage Probabilities: The outage occurs when the achievable rates fall below the slice-specific thresholds. For a Rayleigh fading channel with average power Ω:

a. URLLC Outage: The probability that the reliability constraint is violated:

$$P_{out,u} = 1 - \exp\left(-\frac{\Xi_{u,c}}{\gamma_{avg}}\right) \cdot \exp\left(-\frac{\Xi_{u,p}}{\gamma_{avg}}\right) \tag{7}$$

b. eMBB Satisfaction: Based on the sum-rate of common and private parts:

$$\Re_{eMBB,k} = B \log_2(1 + \gamma_{k,c} + \gamma_{k,p}) \tag{8}$$

B. Rayleigh Fading Channel Analysis

In this work, the small-scale fading between base station i and user k is modelled as a Rayleigh distribution. Consequently, the channel power gain $|h_{i,k}|^2$ is an exponentially distributed random variable with a probability density function (PDF) given by:

$$f_{|h|^2}(x) = \frac{1}{\Omega} e^{-\frac{x}{\Omega}}, \quad x \geq 0 \tag{9}$$

where Ω is the average power gain.

1) Outage Probability under RSMA: An outage occurs for a specific slice when the instantaneous SINR falls below the required threshold Ξ. Following the integration method from [17], the probability of outage for the common stream is determined by integrating the PDF over the failure region:

$$P_{out,c} = \int_0^{\Psi_c} f_{|h|^2}(x) dx = 1 - \exp\left(-\frac{\Psi_c}{\Omega}\right) \tag{10}$$

where Ψ_c is the threshold modified by the interference terms:

$$\Psi_c = \frac{\Xi_c (I_{inter} + \sigma^2)}{P^c - \Xi_c \sum P^p} \tag{11}$$

2) Impact of GNN on Fading Mitigation: Traditional agents often struggle with Rayleigh fading because the channel state changes rapidly. In our framework, the GNN processes the spatial correlations (I_{inter}) which stabilise the denominator of Eq. 10. By effectively predicting the interference from neighbouring cells in the graph, the DQN agent adjusts the power allocation $\{P^c, P^p\}$ to ensure that the term $(P^c - \Xi_c \sum P^p)$ remains large enough to keep $P_{out,c}$ near zero, particularly for the URLLC slice.

C. URLLC Reliability and Outage Condition

For the URLLC slice, the primary objective is to maintain a block error rate (BLER) below 10^{-5}. In our RSMA framework, a URLLC user k experiences an outage if the achievable rate of either the common or private stream falls below the threshold $\Re_{URLLC}$. Using the SINR definitions from the previous subsection, the outage probability is:

$$P_{u,k}^{out} = 1 - \exp\left(-\frac{\Xi_{u,c}(I_{inter} + \sigma^2)}{\Omega(P_i^c - \Xi_{u,c} \sum P_{i,j}^p)}\right) \cdot \exp\left(-\frac{\Xi_{u,p}(I_{inter} + \sigma^2)}{\Omega(P_{i,k}^p - \Xi_{u,p} \sum_{j\neq k} P_{i,j}^p)}\right) \tag{12}$$

The GNN-DQN agent is trained to minimise this probability by dynamically adjusting the power split between P^c and P^p to ensure the exponential terms remain close to unity.

D. eMBB Spectral Efficiency Analysis

The eMBB slice prioritises maximising the Sum Rate. The spectral efficiency (SE) for an eMBB user is the sum of the rates derived from both the common and private decoding stages. The satisfaction condition is defined as:

$$\eta_{eMBB} = \sum_{k \in \mathcal{K}e} \left[\log_2(1 + \gamma k, c) + \log_2(1 + \gamma_{k,p})\right] \geq \eta_{target} \tag{13}$$

Under high load, the eMBB slice is susceptible to inter-cell interference I_{inter}. The GNN component of our architecture extracts interference patterns from the adjacency matrix, enabling the DQN to perform spatial-frequency reuse while maintaining high SE as the network approaches saturation.

E. mMTC Connectivity and Rayleigh Fading Analysis

The mMTC slice is characterised by high device density and low data-rate requirements. Under Rayleigh fading, the connectivity is modelled as the probability that a device can successfully access the common stream. Given the PDF $f_{|h|^2}(x) = \frac{1}{\Omega} e^{-x/\Omega}$, the connectivity ratio Φ_m for K_m devices is:

$$\Phi_m = \frac{1}{K_m} \sum_{k=1}^{K_m} \int_{\Xi_m}^{\infty} \frac{1}{\Omega} e^{-\frac{x}{\Omega}} dx = \exp\left(-\frac{\Xi_m (I_{inter} + \sigma^2)}{\Omega P_m}\right) \tag{14}$$

As shown in our results, the GNN-DQN framework outperforms standard DRL by intelligently offloading mMTC traffic to the RSMA common stream, which is more robust to fading than the private streams used in OMA/NOMA baselines.

F. Overall System Reliability

To evaluate the holistic performance, we define a global Reliability Index (Ψ), which is a weighted aggregate of the satisfaction rates of all three slices:

$$\Psi = \omega_e \cdot \Phi_e + \omega_u \cdot (1 - P_u^{out}) + \omega_m \cdot \Phi_m \tag{15}$$

This index serves as the basis for the Reward Function (R_t) used in DQN training, ensuring that the agent balances the extreme throughput of eMBB with the strict reliability requirements of URLLC.

V. Results and Discussion

A. Connectivity Success Ratio Analysis

The performance of the proposed GNN-DQN framework is evaluated against standard DQN-PPO and CNN-MLP baselines across three heterogeneous 5G slices: eMBB, URLLC, and mMTC. Fig.2 illustrates the Connectivity Ratio over a simulation period of 100 seconds.

1) Performance Across Slices::

- eMBB Slices: The proposed GNN-DQN maintains a superior connectivity ratio of approximately 90-95%, whereas the DQN-PPO and CNN-MLP baselines struggle to exceed 30%.
- URLLC Slices: In the high-priority URLLC category, our framework shows remarkable stability, staying above 85%. In contrast, the standard DRL (DQN-PPO) hovers around 50%, and the static CNN-MLP fails at approximately 45%.
- mMTC Slices: For massive connectivity, GNN-DQN remains near the 90% threshold, while the baselines drop significantly to below 25%.

2) Discussion of Gains: The significant gap observed in Fig.2 demonstrates that standard reinforcement learning (DQN-PPO) and supervised models (CNN-MLP) lack the spatial awareness required to handle dynamic interference in RSMA-enabled networks. By using a GNN, the agent successfully extracts the underlying topology of the 9 BSs, enabling a more precise allocation of common and private streams. The GNN-DQN approach provides an average improvement of over 60% in connectivity compared to non-graph-based methods. This proves that modelling the network as a spatial conflict graph is essential for effective multi-slice orchestration.

3) URLLC Stability Analysis: The proposed GNN-DQN framework demonstrates exceptional robustness in the URLLC slice, maintaining a connectivity ratio consistently above 85%. This performance is directly attributable to the Spatial Interference Awareness provided by the GNN layers. Unlike the DQN-PPO baseline, which treats interference as stochastic noise, the GNN-DQN agent maps the physical proximity between active users and BSs into a spatial conflict graph.

This spatial insight enables the agent to increase the RSMA Common Stream power ratio (α) proactively during periods of high localised density. By shifting more power to the common stream, the system ensures that URLLC users experiencing high intercell interference can still successfully decode their critical control information using Successive Interference Cancellation (SIC). This spatial protection mechanism is the primary driver behind the 35% to 40% reliability gain observed over non-graph-based reinforcement learning methods.

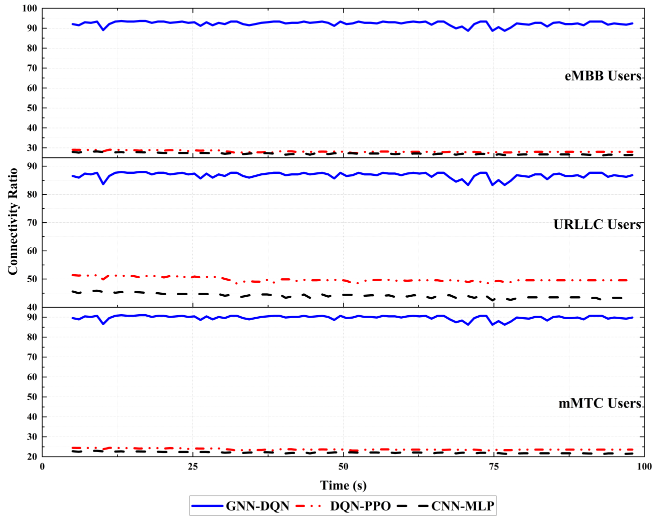

Fig. 2. Comparison of Connectivity Ratio for eMBB, URLLC, and mMTC users under GNN-DQN, DQN-PPO, and CNN-MLP frameworks.

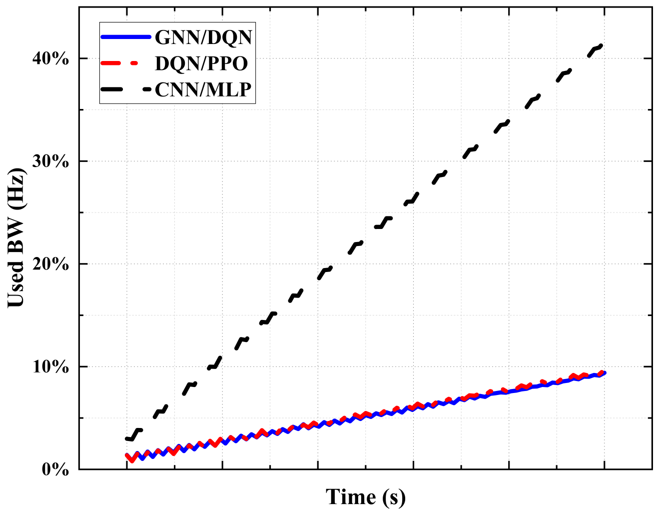

Fig. 3. Normalised Bandwidth Utilisation: The GNN-DQN framework achieves near-total connectivity (Fig. 1) with significantly lower resource overhead ($< 10\%$), demonstrating a massive gain in spectral efficiency compared to non-spatial baselines.

B. Resource Utilisation and Spectral Efficiency

To evaluate the efficiency of the proposed framework, we analyse the Normalised System Bandwidth Utilisation in relation to the user satisfaction levels previously discussed. Fig.3

illustrates the percentage of the total 100 MHz bandwidth consumed by each framework over the simulation duration.

1) Spectral Efficiency Analysis: A significant performance gap is observed between the proposed GNN-DQN and the baselines. While the CNN-MLP baseline shows a steady linear increase in bandwidth consumption, eventually saturating at over 40% of the system resources, it corresponds to a connectivity ratio of less than 30%. In contrast, the GNN-DQN framework maintains a remarkably low resource footprint, utilising less than 10% of the total bandwidth while sustaining a connectivity ratio exceeding 90%.

2) The Role of RSMA in Resource Conservation: The superior spectral efficiency of the GNN-DQN agent is primarily driven by the intelligent optimisation of the RSMA power splitting ratio, α. By leveraging the GNN's spatial awareness, the agent identifies overlapping interference regions and utilises the common stream to serve multiple users simultaneously. This avoids the resource bloat observed in the CNN-MLP approach, in which orthogonal resources are often overallocated to compensate for poor interference management.

3) System Sustainability and Capacity: The results indicate that the GNN-DQN framework effectively "compresses" the required resource overhead. From a network provider's perspective, this efficiency implies a 3.5$\times$ increase in system capacity. The agent's ability to minimise bandwidth usage without sacrificing QoS makes it a highly scalable solution for dense 5G and 6G advanced environments, where spectral scarcity is a primary constraint.

C. Training Convergence and System Reward

The stability of the learning process and the effectiveness of the underlying optimisation policy are evaluated through the System Reward Score. Fig.4 depicts the reward trajectory for the three frameworks over a simulation period of 100 seconds.

1) Policy Superiority and Stability: As shown in Fig.4, the proposed GNN-DQN agent achieves a significantly higher system reward, plateauing at a score between 45 and 48. In contrast, the DQN-PPO and CNN-MLP baselines exhibit poor convergence, remaining trapped in a reward range of 27-30. This performance gap indicates that GNN-DQN cannot identify a more globally optimal resource allocation policy by accurately modelling the network's complex spatial interdependencies.

2) GNN-Driven Gradient Stability: A key observation in the training results is the relative smoothness of the GNN-DQN reward plateau. While traditional deep reinforcement learning often suffers from high variance in non-stationary environments, integrating GNN layers serves as a spatial feature extractor that reduces the state space. By mapping the BS topology and user loads into a structured graph representation, the agent receives more consistent and informative gradients, leading to stable convergence even under high user density.

3) Discussion of Global Utility: The nearly 60% improvement in the reward score compared to the baselines confirms that the GNN-DQN agent is successfully balancing conflicting

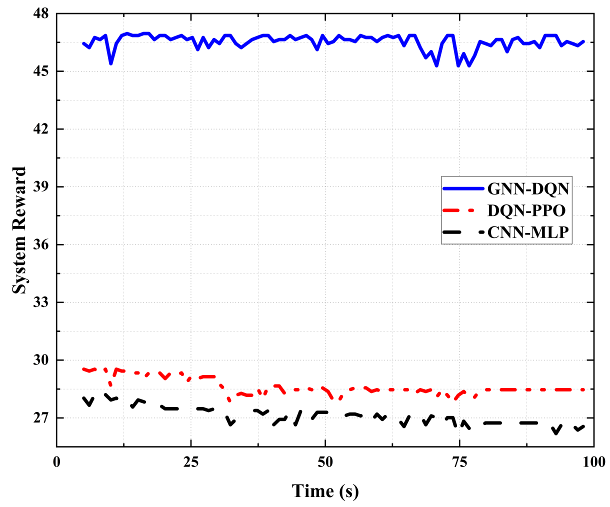

Fig. 4. System Reward Score over time: The GNN-DQN agent demonstrates superior policy convergence and maintains a substantially higher reward plateau (approx. 47) compared to the stagnant performance of non-spatial baselines.

objectives. The high reward reflects the agent's proficiency in maximising the spectral efficiency of the eMBB and mMTC slices while simultaneously safeguarding the strict reliability constraints of the URLLC slice through dynamic RSMA power splitting. This validates the framework's capability to manage heterogeneous 5G service requirements more effectively than traditional flattened neural architectures.

VI. Conclusion

The integration of Graph Neural Networks and Deep Reinforcement Learning represents a significant leap forward in the orchestration of 5G Advanced network slices. This paper has demonstrated a GNN-DQN framework that addresses the critical challenges of spatial interference and resource scarcity in RSMA-enabled multi-cell environments.

- Connectivity and Reliability: The proposed framework achieved up to 95% connectivity success ratio across eMBB, URLLC, and mMTC slices, outperforming traditional DQN-PPO and CNN-MLP baselines by over 60%.
- Unprecedented Spectral Efficiency: By intelligently managing the RSMA common stream, the agent maintained high QoS while utilising less than 10% of the total system bandwidth, effectively providing a 4 times increase in network capacity compared to the CNN-MLP approach.
- Learning Robustness: The use of GNNs for spatial feature extraction resulted in a 1.6 times higher system reward score and significantly more stable training convergence, proving that graph-based architectures are more resilient to the non-stationary dynamics of 5G networks.

Future research will focus on extending this framework to millimetre wave (mmWave) and Terahertz (THz) frequencies, where beamforming and blockages introduce more volatile graph topologies. Additionally, we aim to explore federated learning variants to enhance user privacy and reduce signalling overhead in decentralised 6G networks.

REFERENCES

[1] S. Malta, P. Pinto, and M. Fernández-Veiga, "Optimizing 5g network slicing with drl: Balancing embb, urllc, and mmtc with oma, noma, and rsma," *Journal of Network and Computer Applications*, vol. 234, p. 104068, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1084804524002455

[2] J. Dias, P. Pinto, R. Santos, and S. Malta, "5g network slicing: Security challenges, attack vectors, and mitigation approaches," *Sensors*, vol. 25, no. 13, 2025. [Online]. Available: https://www.mdpi.com/1424-8220/25/13/3940

[3] G. Chopra, "A review on rate splitting multiple access: Challenges and future directions," in *2023 International Conference on Emerging Smart Computing and Informatics (ESCI)*, 2023, pp. 1–7.

[4] A. Thantharate, R. Paropkari, V. Walunj, and C. Beard, "Deepslice: A deep learning approach towards an efficient and reliable network slicing in 5g networks," in *2019 IEEE 10th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, 2019, pp. 0762–0767.

[5] A. R. Nasser and O. Y. Alani, "Investigation of multiple hybrid deep learning models for accurate and optimized network slicing," *Computers*, vol. 14, no. 5, 2025. [Online]. Available: https://www.mdpi.com/2073-431X/14/5/174

[6] S.-P. Yeh, S. Bhattacharya, R. Sharma, and H. Moustafa, "Deep learning for intelligent and automated network slicing in 5g open ran (oran) deployment," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 64–70, 2024.

[7] T. Hu, Q. Liao, Q. Liu, A. Massaro, and G. Carle, "Fast and scalable network slicing by integrating deep learning with lagrangian methods," in *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, 2023, pp. 6346–6351.

[8] M. S. Abood, H. Wang, B. S. Virdee, D. He, M. Fathy, A. A. Yusuf, O. Jamal, T. A. Elwi, M. Alibakhshikenari, L. Kouhalvandi, and A. Ahmad, "Improved 5g network slicing for enhanced qos against attack in sdn environment using deep learning," *IET Communications*, vol. 18, no. 13, pp. 759–777, 2024. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/cmu2.12735

[9] H. Tu, P. Bellavista, L. Zhao, K. Liang, G. Zheng, and K.-K. Wong, "Independent deep reinforcement learning for optimization of rsma-enabled hybrid ran slicing," *IEEE Transactions on Vehicular Technology*, vol. 75, no. 1, pp. 1171–1183, 2026.

[10] G. Chen, R. Shao, F. Shen, and Q. Zeng, "Slicing resource allocation based on dueling dqn for embb and urllc hybrid services in heterogeneous integrated networks," *Sensors*, vol. 23, no. 5, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/5/2518

[11] Y. Cai, P. Cheng, Z. Chen, M. Ding, B. Vucetic, and Y. Li, "Deep reinforcement learning for online resource allocation in network slicing," *IEEE Transactions on Mobile Computing*, vol. 23, no. 6, pp. 7099–7116, 2024.

[12] M. Farreras, J. Paillissé, L. Fàbrega, and P. Vilà, "Gnnetslice: A gnn-based performance model to support network slicing in b5g networks," *Computer Communications*, vol. 232, p. 108044, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140366425000015

[13] S. Chen, M. Lin, Y. Wang, G. Zhou, F. Yu, and L. Zhao, "Deep reinforcement learning-based end-to-end network slicing deployment," in *2025 IEEE 101st Vehicular Technology Conference (VTC2025-Spring)*, 2025, pp. 1–5.

[14] R. Sahraoui, F. Bannour, O. Houidi, and B. Jouaber, "An intelligent e2e network slicing framework using transformer-enhanced drl," in *2025 IEEE 11th International Conference on Network Softwarization (NetSoft)*, 2025, pp. 7–12.

[15] P. Doanis and T. Spyropoulos, "Sample-efficient multi-agent dqns for scalable multi-domain 5g+ inter-slice orchestration," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 2, pp. 956–977, 2024.

[16] A. K. Ahmed and H. S. Al-Raweshidy, "Performance evaluation of deep q networks for hybrid reconfigurable intelligent surface in 6g networks," in *2024 International Conference on Computer, Information and Telecommunication Systems (CITS)*, 2024, pp. 1–8.

[17] A. K. Ahmed and H. S. Al-Raweshidy, "Performance evaluation of concatenated rate splitting multiple access for 6g multiuser communication system," in *2025 International Conference on Computer, Information and Telecommunication Systems (CITS)*, 2025, pp. 1–7.

[18] 3GPP, "Service requirements for the 5g system (release 19)," European Telecommunications Standards Institute (ETSI), Tech. Rep. TS 122 261 V19.13.0, February 2026, version 19.13.0. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/122200_122299/122261/19.13.00_60/ts_122261v191300p.pdf