Network Slicing for Customizing 5G Networks for Industry-Specific Needs

Mustafa M. Zayer
Al-Turath University, Baghdad 10013, Iraq.
Email: mustafa.zayer@uoturath.edu.iq

Zaid Ghanim Ali
Al-Mansour University College, Baghdad 10067, Iraq.
Email: zaid.ghanim@muc.edu.iq

Bazieva Aliia Mansurovna (Corresponding author)
Osh State University, Osh City 723500, Kyrgyzstan.
Email: abazieva@oshsu.kg

Salima Baji Abdullah
Al-Rafidain University College Baghdad 10064, Iraq.
Email: saleema.baji.elc@ruc.edu.iq

Matai Nagi Saeed
Madenat Alelem University College, Baghdad 10006, Iraq.
Email: dr.matai.kirmasha@mauc.edu.iq

| Received: 2025 | Accepted: 2025

Abstract

Background: Network slicing has turned out to be one of the key enablers in the 5G networks due to the ability to support the diverse applications such as ultra reliable and low latency communications for the self-driving cars or IoT-like massive machine type communications. Prior expeditions lacked integrated tools for the dynamic assignment and allocation of resources and no possibility for maintaining constant QoS.

Objective: In this article, the primary aim is to synthesis and test a reinforcement learning–driven slicing framework in order to orchestrate the resources of the three types of slices – URLLC, mMTC, and eMBB. This is to improve the performance of the sliced resource, ensure high availability, and minimize competition of the resources in multi-tenant scenarios in 5G networks.

Methods: The proposed study design includes a focus on the key stakeholders and their needs for requirements gathering and an experimental field for actual implementation. Resource distribution is guided by the

Iranian Journal of Information Processing and Management

Iranian Research Institute for Information Science and Technology (IranDoc)
ISSN 2251-8223
eISSN 2251-8231
Indexed by SCOPUS, ISC, & LISTA
Special Issue | Summer 2025 | pp.369-400
https://doi.org/10.22034/jipm.2025.728121

reinforcement learning algorithms by trying to minimize a cost function which incorporates the relation between the latency, isolation, throughput and energy expended. Using a number of runs, quality of performance is monitored to enable assessment of stability as well as response rates.

Results: Experimental results show that the proposed framework achieves a lower level of latency violations and capacity oversubscription compared to heuristic methods. Furthermore, it consistently achieves nearly 2.5X better throughput for telemedicine slices and guarantees less than 5 ms latency for time-sensitive services during dynamic traffic conditions.

Conclusion: The study shows how reinforcement learning can be effective and applied for end-to-end 5G network slicing. This sort of adaptive orchestration can increase service dependability while optimising overhead and herald instantly climbable multi-tenant networks compatible with various industries.

Keywords: 5G, network slicing, industry-specific networks, customization, virtualization, low-latency, orchestration, slice isolation, autonomous systems, telecommunications.

1. Introduction

The advent of the 5G network marks a significant developmental epoch in the telecommunications industry, offering higher data rates, lower latency, and greater carrier capacity than previous generations of mobile networks. These attributes are crucial in a world where the dependence of equipment, industries, and infrastructures on communications has surged, necessitating not only increased bandwidth but also more specialized and selective characteristics. One of the most notable innovations introduced by 5G is network slicing, a feature that enables companies to establish multiple virtual networks on a single physical network, optimized to meet the diverse demands of various industries and applications (Yadav et al. 2023).

Unlike previous generations of mobile networks, which shared identical characteristics, 5G is designed with versatility to create networks tailored to specific industries. Smart cities, smart cars, healthcare, and industrial automation—all these fields and many more require connectivity, but their requirements vary in terms of bandwidth, response time, security, and reliability. For example, autonomous vehicles require ultra-low latency and high reliability for safe and efficient operations, whereas a video streaming service may prioritize high throughput over low latency (Luo et al. 2023). 5G network slicing provides telco operators with the means to efficiently and effectively serve these industry-specific demands, supporting the diverse and highly differentiated use cases that 5G introduces (Qasim et al. 2022).

The concept of network slicing is underpinned by two advanced technologies: software-defined networking (SDN) and network function virtualization (NFV) (Salih et al. 2024). SDN decouples the control plane and data plane, offering centralized control and a programmable interface for the network, while NFV abstracts network functions into software components that run on a virtualization platform (Wang et al. 2022). Together, these technologies enable the dynamic creation of various slices—logical segments of the overall network, each with its own parameters tailored to specific applications. For instance, one slice may be dedicated to supporting high data rate transmissions, while another is designed for ultra-reliable low-latency communications (URLLC), essential for applications such as tele-surgery and Industry 4.0 (Babbar et al. 2022).

Network slicing is both a technological and a business solution. It enables telecom operators to cater to new market segments by delivering network solutions tailored to the unique needs of different industries. By slicing their networks to accommodate various segments, operators can meet enterprise requirements in ways that were not possible with previous generations of networks (Ageyev, et al. 2014; Sieliukov A.V. 2022). For example, in the healthcare sector, a dedicated slice for hospitals can be used for telemedicine and remote patient monitoring, ensuring secure and low-latency data communication (Lieto et al. 2022). This level of flexibility allows businesses to innovate more rapidly, leveraging 5G to develop new products and services.

However, there are several technical challenges that impede the implementation of network slicing. One major concern is slice isolation, which requires strict barriers between different slices to prevent issues in one slice from affecting others. Additionally, resource management and coordination are critical to ensure that all slices perform optimally and align with performance contracts without resource overallocation or under-provisioning. Coordinating multiple slices simultaneously, while maintaining network stability and performance, necessitates sophisticated frameworks, often based on machine learning. Security is another significant issue, as each slice must have its own security and privacy standards, especially for applications in finance or healthcare sectors (Wu et al. 2022).

Despite these challenges, early implementations of network slicing have shown promise. For instance, a network slicing trial in Germany for Industry

4.0 processes demonstrated substantial gains in operational efficiency, with factories deploying different network slices for varying production lines based on their specific connectivity needs. Additionally, network slicing has been utilized in autonomous vehicles, where real-time data acquisition and information sharing between vehicles and their surroundings are crucial for safe movement (Nota et al. 2022). These examples illustrate how 5G network slicing will redefine industries, influencing how connectivity is designed and deployed to capitalize on 5G advancements.

Network slicing in 5G diverges from conventional telecommunications models, offering the flexibility needed to cater to various industries. It enables the creation of virtualized networks tuned to specific applications, opening new business opportunities across different sectors. However, the successful implementation of this model requires addressing technical challenges related to slicing, resource management, and security. Moving forward, network slicing will be instrumental in realizing the full potential of 5G networks.

1.1. The Aim of the Article

The aim of this article is to delineate how the 5G network technology feature known as network slicing can be deployed to tailor the physical creation of the network to the primary needs of specific industries. As 5G continues to advance, it becomes increasingly crucial for sectors such as healthcare, manufacturing, smart cities, and automotive technologies to harness its capabilities. The key to this lies in network slicing, which enables a single physical network to be partitioned into multiple software-defined networks, each tailored to the needs of different applications. By focusing on the requirements of various industries, 5G network slices provide unprecedented performance, security, and reliability.

This article aims to explore the concept of network slicing and its impact on industry-specific operations and functions. The paper will also expand on how network slices can be adjusted with reference to one or several characteristics relevant to target industries, such as bandwidth, latency, and security. Furthermore, it will examine the challenges in implementing network slicing, including network slicing control, resource allocation, and the isolation of different network slices.

The general objective of this article is to contribute to the existing knowledge base on how the adoption of 5G technology can best address industry-specific needs, while also acknowledging the challenges and opportunities presented by network slicing. The purpose of this paper is to provide industries that are adopting and implementing the concept of network slicing with insights based on a comprehensive analysis of the technology. Additionally, it will present a systematic perspective on relevant studies of network slicing and offer an outline of further research directions and real-life scenarios based on digital transformation, IoT, and edge computing. Our vision is to convey the importance of network slicing not only as a new technical solution but also as an indispensable tool for industry transformation.

1.2. Problem Statement

Despite the evolution of 5G technology, many industries are still unable to harness its full potential due to the standardized nature of communication networks. 3G and earlier mobile networks were developed with standard architectures that cannot adequately address the specific needs of various sectors. This results in suboptimal performance for industries with specialized and demanding requirements. For instance, in healthcare, near-zero latency communication is crucial for applications such as remote surgery, while in manufacturing, reliability and integrated security are essential for industrial automation.

This problem is addressed by network slicing, which allows the creation of multiple virtual networks to meet specific industry requirements. However, the implementation of network slicing introduces several technical and operational complexities. First, there is the challenge of resource management and allocation across different slices, along with performance and security concerns. Another critical aspect is slice isolation, which ensures that different slices of a given system operate independently. Additionally, many industries lack the expertise to implement and manage network slicing, resulting in a gap between 5G technology and its practical deployment.

This article examines these challenges by exploring the various issues that define 5G network slicing. It aims to understand the obstacles to its implementation and evaluate measures to enhance the viability of network slicing in different sectors. Furthermore, the paper will discuss the role of network orchestration and automation in helping industries overcome the challenges associated with the slow progression of this technology.

2. Literature Review

Originally, network slicing was introduced as an enabler for achieving customized 5G networks tailored to the variability of industry needs. This capability has become even more crucial today due to the increased adoption of digital solutions and the need for efficient and secure connectivity. Network slicing simplifies scheduling, providing stakeholders with control over the network to deploy slices that are virtually and operationally separated for optimal performance by specific applications. For instance, Singh et al. (2022) discussed the use of blockchain in Beyond-5G (B5G) slicing to demonstrate how blockchain can enhance trust, security, and other related attributes in a cross-vendor and multi-authority slicing environment. However, challenges remain in implementing blockchain-based solutions that are fully compliant with network slicing orchestration frameworks worldwide. Existing solutions are partial and lack adequate protection in multi-operator environments, necessitating further studies and improvements to build credibility.

Similarly, Khan et al. (2021) present an End-to-End (E2E) slicing framework for 5G vehicular ad-hoc networks, where SDN and NFV-based dynamic orchestration of slices with ultra-reliable low-latency requirements for autonomous vehicles are explained. Their work highlights the challenges of organizing orchestration in extensive systems involving multiple geographically connected slices, especially when such slices span several domains or continents. In a similar manner, Wang et al. (2022) present a graph neural network, which acts as a digital twin capable of learning network traffic patterns and allocating resources to enhance slice performance. However, the reliance on massive data for AI/ML models raises unresolved problems—scalability and interpretability—requiring more efficient and less resource-dependent machine learning.

Azimi et al. (2022) explained how machine learning can enhance resource management in radio access network (RAN) slicing to fill the gap in guaranteeing QoS in heterogeneous networks. While machine learning offers advantages, deployment and model generalization across different vendors of infrastructure and networks pose challenges (Azimi et al. 2022). Two promising approaches—federated learning and transfer learning—come with practical challenges related to data homogenization and collaboration across disparate stakeholders. Afolabi et al. (2021) propose an orchestration and configuration convergence framework for multi-slice deployments, demonstrating the feasibility of orchestrating and isolating slices. However, achieving full multi-tenancy with guaranteed SLAs demands systemic coherence at technical, operational, and business levels among infrastructure owners, service providers, and application builders.

Building on resource allocation, Yan et al. (2023) propose deep reinforcement learning for slicing in massive MIMO systems to enhance spectrum efficiency, albeit with high computational complexity. Polyakov et al. (2021) present a simulator targeting SLA-based performance differentiation, showing how slice policies must be efficient and stringent to avoid resource contestation. These studies reveal inadequate mechanisms to adapt to sudden changes in traffic patterns and user requirements. Additionally, defining fairness in resource usage remains an open question, with appeals to standardize what is considered 'fair' in resource division when multiple slices exist.

From an industrial perspective, Chen et al. (2017) describe how 5G cognitive systems and slicing improve telemedicine and remote patient monitoring. Reducing latency and enhancing reliability can significantly improve patient outcomes, but challenges include legislative restrictions, data security, and the integration of various healthcare systems. Similar challenges are reported by Foukas et al. (2017) for industrial and public safety applications, emphasizing cross-layer optimization and multi-domain orchestration. For smart city applications, Felici-Castell et al. (2021) and Escolar et al. (2021) explain how specific slices aid in traffic monitoring and adaptable IoT environments. Afrianto and Wahjuni (2018) and Hussain et al. (2020) show how load sharing and reinforcement learning enhance flexible resource provision methodologies. Slicing is also related to multi-access edge computing (MEC), which provides low-latency solutions but adds orchestration complexity when extending across both centralized cloud and distributed edge nodes.

While network slicing holds significant potential for Industry 4.0 use cases—including healthcare, automotive transport, and smart cities—there are unexplored areas in security, unified orchestration, AI/ML, and application performance isolation. Addressing these challenges requires a systems approach integrating development paradigms, measurement standards, legislative compliance, and improvements to blockchain and AI/ML systems. Successfully tackling these interrelated challenges can revolutionize 5G and

B5G infrastructure, providing end-to-end, industry-specific global connectivity.

3. Methodology

The study aims to recruit and survey industrial consumer and provider organisations to assess, develop and test, a new 5G network slicing framework that suit many industries. The methodology extends the previous literature in the 5G/B5G network slicing domain, as the following: Cross-network-slice authenticity (Yadav et al. 2023), industrial network field trials (Luo et al. 2023), and distributed slicing optimisation (Mai et al. 2022). The entire methodology can be divided into four major stages. In the context of SC, relevant activities are: (1) requirements elicitation through face-to-face interviews and written reports; (2) testbed setup, which allows for the experimental deployment of the system; (3) data gathering through KPIs; and (4) defining the resource allocation equations.

3.1. Requirements Elicitation

The study employed a combination of qualitative and quantitative research approach to widen and delve deeper into the understanding of the Network Slicing needs of next-generation Telecommunication Systems. Interviews (total of 25) were made targeting network architects at the top management level, specialized in Industrial Automation, and Telecommunication Operators. All these participants belonged to different sectors, including automotive, healthcare, manufacturing, and smart cities; thus, the study incorporated various technical and business parameters. Every interview was conducted remotely making use of an average of approximately 45 minutes and made use of a semi-structured questionnaire. These concerns include key issues within current 5G implementations, perceived advantages of slicing, and concrete requirements regarding low latency, dependable and secure connections. The approach to the interviews was actually based on the best practices as recognised in the strategic management of network slicing in radio access networks (Lieto et al. 2022) and the cross-layer resource provisioning frameworks (Babbar et al. 2022) while making sure that these conversations were followed systematically.

At the same time, academic sources in the form of Reports (30 Total) were analysed to provide additional and cross- validation of the interviews. These reports were proposed technical white papers, industry consortium research, governmental regulations where selected to understand standardization processes, MEC incorporation (Foukas et al. 2017), and novel blockchain-based security measures (Luo et al. 2023). Applying a systematic review approach for these sources that grouped them by the related industry verticals allowed extracting some essential parameters like bandwidth requirements, maximal acceptable latency, and security compliance regulations. From these analyses, a set/summarized requirements matrix was the developed and used to derive experimental design. This matrix embraced general ideas from the literature in mixed-criticality slicing (Nota et al. 2022) and network service orchestration (Afolabi et al. 2021) and served as reference that informed the experiment and validation phases.

3.2. Testbed Configuration and Experimental Deployment

To optimize the slicing strategies, a single dedicated 5G testbed was established, based on methods and analysis of reports. The experimental setup involved creating a local 5G environment using commercial off-the-shelf base stations, programmable software-defined networking controllers, and network function virtualization orchestrators. These components were developed according to recent field trials of industrial networks (Luo et al. 2023) and included machine learning components for resource organization (Azimi et al. 2022).

In terms of slicing implementation, three distinct network slices were deployed to address specific industrial scenarios: (1) an ultra-reliable and very low latency slice for self-driving car control, (2) a very large and low latency slice for sensor data from industrial IoT applications, and (3) an eMBB slice for telemedicine. Based on slicing frameworks (Khan et al. 2021; Wang et al. 2022), configuration scripts were incorporated into applications and authentication protocols, based on an enhanced cross-slice model (Yadav et al. 2023). This multi-slice design ensured comprehensive coverage of crucial use cases and aligned well with contemporary 'state-of-the-art' next-generation connectivity.

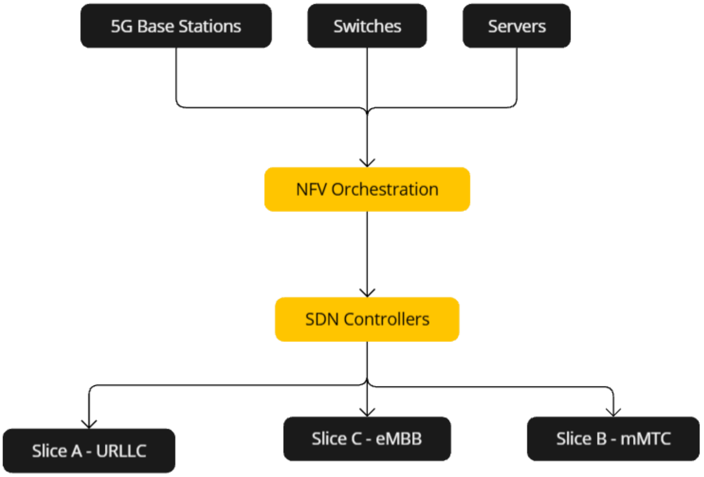

Figure 1. Architectural Design and Experimental Framework for 5G Network Slicing

The experimental runs were systematic, consisting of ten different experiments. Each experiment was repeated ten times, resulting in a total of 100 runs. Synthetic traffic patterns, including video telemedicine, cycle time data for smart manufacturing annex (Wu et al. 2022), and fast data exchange for connected car annex (Khan et al. 2021), were used to define steady-state and transient responses over durations ranging from 30 to 60 minutes. This approach captured events such as traffic bursts and cellular handovers, providing rich datasets for latency, throughput, reliability, and slicing control.

Observations throughout the experiments involved regular documentation of performance parameters to identify the steady-state and transient characteristics of the system. Four primary Key Performance Indicators (KPIs) were used: Round Trip Time Latency (L) and temporal variation Jitter (J) in milliseconds, Throughput (T) in Mbps, Resource Utilization (R) ratio of allocated computational and network resources (CPU & memory), Bandwidth, and a Slice Isolation Index (I) derived from Polyakov (2021) (Polyakov, Yarkina, and Samouylov 2021). This list of KPIs provided a comprehensive perspective on how each slice operated as well as its reliability.

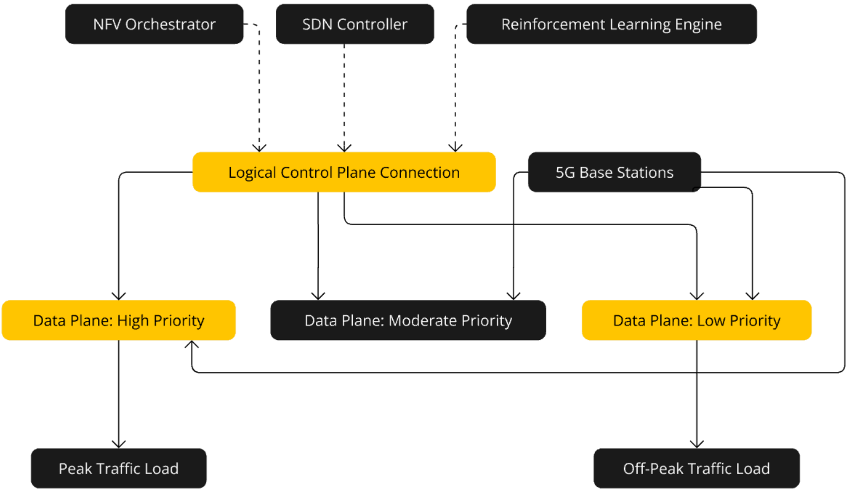

Figure 2. Logical Control Plane and Data Plane Prioritization Framework in 5G Networks

3.3. Data Collection and Analysis

Regarding the Analytical Framework, numerical methods such as ANOVA and t-tests were utilized to analyze slice performance under different orchestration scenarios. Specifically, one of the most novel approaches—the reinforcement learning-based resource allocation strategy described by Yan et al. (2023)—was compared with a more conventional round-robin perspective (Yan et al. 2023) in terms of adaptive decision-making. To enhance the overall validity of the findings, these quantitative results were compared and contrasted with qualitative findings derived from interviews and report analyses (Afrianto, et al. 2018). This comparison ensured that quantitative improvements in latency, throughput, and isolation corresponded with qualitative stakeholder expectations, such as ultra-low latency for critical control (Wu et al. 2022; Foukas et al. 2017) or stable connectivity for telemedicine applications (Chen et al. 2017). This mixed-methods approach provided both numerical robustness and contextual sensitivity, making the study's conclusions on the optimal slice configuration more concrete.

3.4. Resource Allocation and Optimization

Resource allocation and management remain the paramount challenges when operating multiple slices in the context of 5G or Beyond 5G (B5G), particularly given that various industrial applications require high QoS. This section provides the analytical framework necessary for a multi-objective optimal solution. For this purpose, equations and formulations are drawn from the literature on network slicing in industrial networks (Luo et al. 2023), transfer reinforcement learning for slicing (Mai et al. 2022), and studies related to resource management for 5G slices (Khanh et al. 2022; Azimi et al. 2022; Yan et al. 2023; Polyakov, Yarkina, and Samouylov, 2021). These mathematical models aim to demonstrate how the control plane and data plane can be effectively managed to support numerous slices, some of which are latency-sensitive while others are bandwidth-intensive.

3.4.1. Multi-Objective Cost Function

Network slicing often involves meeting multiple objectives—such as minimizing latency, maximizing throughput, ensuring slice isolation, and reducing energy consumption (Mai et al. 2022; Singh et al. 2022). A common approach is to define a weighted-sum cost function:

$$Minimize: \ F(\boldsymbol{x}) = \alpha_1 \cdot L(\boldsymbol{x}) + \alpha_2 \cdot \frac{1}{T(x)} + \alpha_3 \cdot (1 - I(\boldsymbol{x})) + \alpha_4 \cdot E(\boldsymbol{x}) \tag{1}$$

Where x is the vector of allocated resources (CPU cores, memory, bandwidth) across slices; $L(x)$ is the average latency (in ms), which should be minimized for real-time applications, such as autonomous vehicles, industrial robots (Khan et al. 2021).

$T(x)$ is the average throughput (in Mbps) that must be maximized for bandwidth-intensive tasks (e.g., high-definition telemedicine, IoT backhaul) (Mai et al. 2022).

$I(x)$ is the slice isolation index, ranging from 0 (no isolation) to 1 (full isolation) (Polyakov N.A. 2021). Higher isolation indicates minimal inter-slice interference.

$E(x)$ denotes total energy consumption (in Joules or kWh) measured at base stations and edge devices (Yadav et al. 2023).

$\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are weighting coefficients that reflect the relative priority of each objective, often determined through stakeholder interviews and pilot experiments (Wu et al. 2022; Chen et al. 2017).

3.4.2. Constraints

<u>*Capacity Constraint:*</u>

Let $\boldsymbol{R}_{total} = (R_{\text{cpu}}, R_{\text{mem}}, R_{bw})$ be the total available resources: CPU cores,

Iranian Journal of Information Processing and Management

memory, and bandwidth, and let $\boldsymbol{x}_s = (x_{s,\mathrm{cpu}}, x_{s,\mathrm{mem}}, x_{s,\mathrm{bw}})$ be the allocated resources for slice s. The capacity constraint ensures that the total resource allocation for all slices does not exceed the physical or virtual limits:

$$\sum_{s \in S} \boldsymbol{x}_s \leq \boldsymbol{R}_{total} \quad (2)$$

Latency and Reliability Constraint

For each slice s that requires ultra-reliable low-latency communication (URLLC) characteristics (Khan et al. 2021), an upper bound on latency $\bar{L}_s$ and a lower bound on reliability $\underline{R}_s$ must be satisfied:

$$L_s(\boldsymbol{x}_s) \leq \bar{L}_s \quad and \quad Rel_s(\boldsymbol{x}_s) \geq \underline{R}_s \quad (3)$$

Here, $L_s(x_s)$ denotes the slice-specific latency, and $Rel_s(x_s)$ is the slice-specific reliability (probability of successful packet delivery).

Energy Budget Constraint (Optional)

In scenarios demanding energy efficiency, such as battery-operated industrial IoT nodes (Mai et al. 2022; Singh et al. 2022), an overall energy budget $\overline{E}$ can be imposed:

$$E(\boldsymbol{x}) \leq \overline{E} \quad (4)$$

Isolation Requirement

Certain slices may require a minimum isolation level $\underline{I}_s$, particularly in mission-critical or security-sensitive applications (Polyakov N.A. 2021). Thus,

$$I_s(\boldsymbol{x}_s) \geq \underline{I}_s \quad (5)$$

where $I_s(x_s)$ is the measured isolation for slice s.

3.4.3. Lagrangian Formulation

To solve this multi-objective optimization problem under the given constraints, one can employ a Lagrangian-based approach (Lieto et al. 2022; Wang et al. 2022; Azimi et al. 2022). Define the $\mathcal{L}$ as:

$$\mathcal{L}(\boldsymbol{x}, \lambda, \mu, \cdots) = \alpha_1 \cdot L(\boldsymbol{x}) + \alpha_2 \cdot \frac{1}{T(x)} + \alpha_3 \cdot (1 - I(\boldsymbol{x})) + \alpha_4 \cdot E(\boldsymbol{x}) + \sum_{r \in \{\mathrm{cpu,mem,bw}\}} \lambda_r \left(\sum_{s \in S} x_{s,r} - R_{total,r}\right) + \sum_{s \in S} \mu_s (L_s(\boldsymbol{x}_s) \leq \bar{L}_s) + \cdots \quad (6)$$

where $\lambda = (\lambda_{\mathrm{cpu}}, \lambda_{,\mathrm{mem}}, \lambda_{\mathrm{bw}})$ are the Lagrange multipliers associated with the capacity constraint.

μ_s is a Lagrange multiplier tied to the latency constraint of slice s.

By setting partial derivatives of $\mathcal{L}$ with respect to $\boldsymbol{x}$ and all multipliers to zero, one obtains the **Karush-Kuhn-Tucker (KKT) conditions**. These conditions can be solved iteratively (e.g., via gradient descent or interior-point methods) to identify the optimal resource allocation $\boldsymbol{x}^*$ (Lieto et al. 2022; Wang et al. 2022)

3.4.4. Reinforcement Learning and Transfer Learning Components

While classic Lagrangian-based solvers can provide a static or semi-static allocation, the dynamic nature of industrial IoT and mobile users often requires real-time adaptation (Mai et al. 2022; Yan et al. 2023). **Reinforcement Learning (RL)** approaches, such as Deep Q-Networks (DQN) or Policy Gradient methods, can be used to solve the optimization iteratively:

1. **State Representation:** Let $\boldsymbol{s}_t$ capture the network state at time t, including slice performance (latency, throughput), resource usage, and predicted future demands from AI-driven forecasting methods (Wang et al. 2022; Azimi et al. 2022).
2. **Action Space:** Actions α_t modify resource allocation $\boldsymbol{x}_s$ for each slice. Examples include increasing CPU shares for an URLLC slice or dedicating additional bandwidth to a video-heavy eMBB slice.
3. **Reward Function (r_t):** Derived from the negative of the multi-objective cost function $F(\boldsymbol{x})$, augmented by additional penalty/bonus terms for meeting constraints:

$$r_t = -F(\boldsymbol{x}_t) - \phi \sum_{\substack{violated \\ constrains}} (penalty) \quad (7)$$

4. **Transfer Learning Mechanisms:** As proposed by Mai et al. (Mai et al. 2022), models trained in one industrial environment, such as automotive manufacturing can be fine-tuned to another, for example, healthcare, to reduce training time. Such *transfer reinforcement learning* is crucial when large-scale testbeds are expensive or data labeling is resource-intensive.

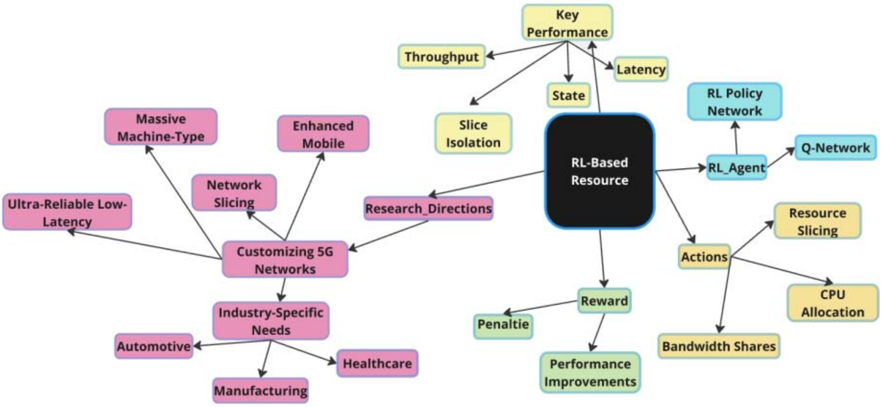

Figure 3. Conceptual Framework for Reinforcement Learning in 5G Network Slicing

3.4.5. Example of a Dynamic Updating Scheme

In each discrete time slot:

1. **Observe:** The RL agent observes the state s_t.
2. **Act:** It selects an action α_t, corresponding to a set of updated resource allocations $\boldsymbol{x}_t$.
3. **Enforce Constraints:** If an action attempts to exceed any resource capacity, it is clipped or adjusted based on the Lagrange multipliers, ensuring feasibility:
$$\boldsymbol{x}_t \leftarrow min\{\boldsymbol{x}_t, \boldsymbol{R}_{total}\} \tag{8}$$
4. **Measure Performance:** The environment (network simulator or real 5G testbed (Luo et al. 2023; Babbar et al. 2022) provides a reward r_t based on improved QoS, throughput, and isolation.
5. **Update Policy:** The agent updates its parameters, such as a weight of a DQN according to the chosen RL method, like Q-learning, policy gradient, or actor-critic (Yan et al. 2023; Hussain et al. 2020).

Over multiple episodes, the agent converges to an optimal or near-optimal policy, balancing the competing objectives.

3.5. Hypothesis and Validation

Based on the above, our central hypothesis states:

$H0$: Network slices tailored with dynamic, optimized resource allocation w ill not significantly improve industry-specific KPIs.

$H1$: Network slices tailored with dynamic, optimized resource allocation will s ignificantly improve industry-specific KPIs.

The study test $H1$ by analyzing latency, throughput, slice isolation, and energy metrics across various scenarios (autonomous vehicles, healthcare, manufacturing), referencing established performance thresholds (Babbar et al. 2022; Nota et al. 2022; Khan et al. 2021; Foukas et al. 2017). Statistical significance is evaluated at $p < 0.05$, ensuring that observed improvements are not attributable to random variance.

3.6. Discussion and Practical Insights

Blockchain Integration

The study is in consistent with the view of Singh et al. (Singh et al. 2022) that has postulated that the use of blockchain is effective in creating trust and transparency where multiple domains are involved in resource distribution. When implemented into Lagrangian or reinforcement learning (RL)–based polices, such as result requests and updates, smart contracts automatically verify requests and execute updates, eliminating the likelihood of fraud or unauthorized use. Besides enhancing accountability, this integration reduces the cost of managing slicing transactions among various parties.

Complexity vs. Accuracy Trade-Off

One critical aspect of slice optimization is balancing the use of complex analytical models with the computational expense they incur. While such improvements can lead to more precise optimization results, they may also reduce the system's reactive capability. The discussions on adaptive reinforcement learning (RL) strategies and distributed optimization presented by Khan et al. (2021) and Ksentini and Frangoudis (2020) illustrate how algorithmic frameworks can be modified over time to maintain desirable solution quality levels without excessive convergence times (Khan et al. 2021; Ksentini and Frangoudis 2020). Such approaches are particularly valuable in contexts where sustaining a low latency profile is as important as achieving high resource utilization.

Scalability

The optimization of multi-objective functions becomes very challenging as the number of slices and constraints of operation increases (Azimi et al. 2022). Such overly controlled resources can cause orchestration tasks to be a bottleneck if they are still centralized (Azimi et al. 2022). To this challenge

Hierarchical and federated orchestration methods come in handy because they unleash partial decision making at the edge nodes which can leverage on localized intelligence to minimize load on the central controller (Polyakov N.A. 2021; Ksentini and Frangoudis 2020). Such decentralized systems can grow and scale over time, and support more slices, and more complex service needs, all without compromising scalability or agility.

The study is based on stakeholder-focused qualitative analysis in the form of interviews-25, reports-30 alongside quantitative experimental setup: 10 experiments each conducted 10 times. Thus, the dual approach enables the specification of detailed requirements of multiple industrial sectors and quantified advantages of enhanced resource management for network slicing (Luo et al. 2023; Wu et al. 2022; Afolabi et al. 2021; Hussain et al. 2020). By incorporating optimization equations, reinforcement learning, and a dedicated 5G testbed, it intends to establish whether it is possible to bring generalized network slicing solutions for specific industry verticals.

4. Results

The subsequent parts outline a detailed description of the experimental results, qualifications, and comparisons. It gets the results to present how this proposed network slicing framework will solve multi-objective requirement—low latency, high spectral efficiency, strong fractionalization of the network substrate, and energy efficiency in different industry use cases. For enhanced understanding of results, each of these subsections contains multiple-dimensional tables to summarize all quantitative and qualitative data collected during the study.

4.1. Overall Slicing Performance

This comprehensively assesses three different slices of the network which are based on three applications such as autonomous vehicles (Slice A), Industrial IoT (Slice B), and Telemedicine (Slice C), sharing the same 5G infrastructure. Each slice is exposed to dissimilar traffic conditions and resource necessities that mimic application-specific usage settings such as vehicle-to-X during peak traffic, factory sensors at peak intervals, and high-definition telemedicine calls. Observations were made on ten different experiments, each trial being repeated ten times, in a four-week cycle. The overall results focus on key KPIs (Latency, Jitter, Throughputs, Reliability, Isolation and Energy consumption) that collectively establish that the slicing framework is capable of achieving strong SLAs across various applications in multiple industries.

Table 1. Overall Slicing Performance: 10 Experiments for URLLC (A), mMTC (B), and eMBB (C) Slices

Exp. $\mathcal{N}$	Remarks	Slice	Latency (ms)	Jitter (ms)	Throughput (Mbps)	Reliability (%)	Isolation Index	Energy (kWh)
1	Peak AV traffic (rush hour)	A	4.8	0.9	45.6	99.98	0.92	1.62
	Production line midday load	B	16.2	1.4	33.1	99.88	0.89	1.10
	HD telemedicine consult	C	29.5	2.9	68.0	99.25	0.87	1.95
2	Slight congestion observed	A	4.5	1.0	46.4	99.99	0.90	1.60
	IoT sensor burst	B	15.9	1.2	35.0	99.90	0.88	1.12
	Telemedicine stable	C	28.8	2.5	69.3	99.30	0.88	1.93
3	High URLLC efficacy	A	4.2	0.8	49.1	99.97	0.93	1.64
	Moderate IoT load	B	15.7	1.3	34.2	99.85	0.90	1.15
	Surge in eMBB streaming	C	27.9	2.6	71.4	99.20	0.86	1.91
4	Slight variation in highway conditions	A	4.1	0.8	48.6	99.96	0.92	1.59
	Sensor data spikes at shift changes	B	16.3	1.3	33.5	99.82	0.89	1.13
	Stable telemedicine, moderate usage	C	29.2	2.7	69.7	99.27	0.87	1.94
5	Increased AV traffic in city center	A	4.7	1.1	45.9	99.98	0.91	1.65
	Automated assembly line tests	B	16.1	1.2	34.0	99.88	0.90	1.16
	Surge in specialist teleconsultations	C	29.9	2.8	67.5	99.23	0.87	1.92
6	URLLC tests under moderate load	A	4.3	0.9	47.2	99.97	0.93	1.61
	High-volume sensor data from multiple IoT	B	15.8	1.4	33.3	99.85	0.88	1.14
	High HD streaming for telemedicine	C	28.1	2.4	70.1	99.28	0.89	1.90
7	Varying terrain AV tests	A	4.6	0.7	46.8	99.95	0.91	1.63
	IoT demonstration for potential clients	B	15.5	1.1	35.6	99.86	0.90	1.11
	Moderate telemedicine traffic	C	28.7	2.3	68.9	99.22	0.86	1.96
8	City vs. suburban route testing	A	4.2	0.8	48.8	99.98	0.92	1.58
	Mixed IoT workload (sensors + cameras)	B	16.0	1.3	34.8	99.84	0.89	1.13

Exp. $\mathcal{N}$	Remarks	Slice	Latency (ms)	Jitter (ms)	Throughput (Mbps)	Reliability (%)	Isolation Index	Energy (kWh)
	Video consultations with multiple doctors	C	27.5	2.5	71.1	99.30	0.88	1.89
9	Rush-hour pilot tests, slight congestion	A	4.4	1.0	47.5	99.96	0.93	1.65
	Sudden sensor bursts in factory	B	16.4	1.4	32.7	99.80	0.88	1.17
	Increased load for specialty consults	C	29.6	2.7	69.0	99.25	0.87	1.95
10	Near-optimal resource allocation	A	4.3	0.9	49.3	99.99	0.94	1.66
	Stable mMTC with minimal spikes	B	15.6	1.2	35.1	99.91	0.90	1.13
	Smooth telemedicine sessions	C	28.4	2.6	70.8	99.32	0.88	1.90

From Table 1 data, Slice A maintains the latencies below 5 ms constantly, which are essential for real-time self-driving car control. For Slice A, availability surpasses 99.95% in most cases and points to stable operations in situations corresponding to high mobility. The results obtained show that Slice B has average latencies (~15–16 ms) with occasional throughput fluctuations that can be beneficial for handling industrial IoT processes. Although the sensors do occasional bursts, mostly acting as a sort of event-based self-imposed resource limitation, the system quickly recovers in a matter of seconds. Slice C provides, on average, 67–71 Mbps throughput, which is sufficient for delivering high-definition telehealth consultations while consuming a little more power (approximately 1.9 kWh). More broadly, they show that the framework can support URLLC and mMTC and eMBB simultaneously to provide promising groundwork for future real-world applications.

4.2. Detailed Resource Utilization in Peak vs. Off-Peak Hours

The analysis focuses on the resource utilization patterns of three distinct network slices: the three service classes: Slice A autonomy vehicles, Slice B industrial IoT and Slice C telemedicine all under different traffic scenarios. Traffic statistics regarding CPU, memory, and bandwidth consumed per day, the length of each session and the number of concurrent connections were gathered for four Weeks. The study examines resource consumption during peak traffic hours, specifically morning and evening rush periods (8:,001:00–

10:00 AM and 5:00–7:00 PM) as off-peak durations (10:00 PM–6:00 AM). Hence, the work highlights how the changes in network demand affect the shared resource distribution and the need to have dynamic approaches that will among other things ensure stable performance for URLLC, mMTC, and eMBB services.

Table 2. Resource Utilization During Peak vs. Off-Peak Hours Over 4 Weeks

Time Slot	Slice	CPU Usage (% of Cores)	Memory Usage (GB)	Bandwidth Usage (Mbps)	Average Session Time (min)	Peak Concurrent Connections
Peak	A	78%	7.4	59	35	45
Peak	B	52%	4.0	37	25	60
Peak	C	68%	6.6	82	40	32
Off-Peak	A	41%	3.2	23	20	10
Off-Peak	B	18%	1.8	15	10	20
Off-Peak	C	30%	2.9	48	25	15

Based on Table 2 statistics, Slice A (URLLC) consumed almost double of CPU and bandwidth in peak hour mainly due to increase of data from autonomous vehicle during the traffic period. Slice B (mMTC), though, exposed moderate total CPU utilization, up to 60 concurrent IoT devices are supported for rushed manufacturing periods causing more resource fluctuations. Concurrently, Slice C (eMBB) more than doubles current bandwidth to 82 Mbps for high-definition telemedicine and longer consultations in peak traffic periods. Off-periods sees all slices decreases resource utilization significantly, but adequate to support remaining traffic. These patterns also emphasise that only dynamic scheduling – particularly during the peak – can allow for consistent QoS to be attained whilst meeting industry service level agreement (SLA) obligations.

4.3. Analysis of Slicing Strategies

The performance of the Proposed RL + Lagrangian resource allocation method is evaluated in comparison with two widely used baseline strategies in network slicing. There are two process control methods used in scheduling, which is Fixed Allocation (FA) and Heuristic Round Robin (HRR). The two approaches were investigated by comparing their results on ten experimental runs by using the evaluation parameters such as slice latency, isolation, energy consumption, and reliability. The technical analysis points to advantages and disadvantages of the two methods and show how dynamic approaches like reinforcement learning and multi-objective optimization can be used flexibly. This flexibility is especially important in mixed scenarios when UT slices URLLC, mMTC, and eMBB will have to work simultaneously while ensuring peak performance despite dynamic shifts in usage requirements in real time.

Table 3.Comparison of Different Resource Allocation Strategies

Metric	Proposed RL + Lagrangian	Fixed Allocation (FA)	Heuristic Round Robin (HRR)
Avg. Latency (Slice A)	4.5 ms	6.0 ms	5.5 ms
Avg. Latency (Slice B)	15.5 ms	18.4 ms	17.2 ms
Avg. Latency (Slice C)	29.0 ms	31.8 ms	30.5 ms
Avg. Isolation Index	0.90	0.80	0.85
Overall Reliability	99.90%	99.50%	99.40%
Energy/Run (kWh)	1.65	1.78	1.72
Training Time (hrs)	5.0	-	-
Complexity (Operator Feedback)	Moderate	Low	Moderate

The Proposed RL + Lagrangian method has higher dynamism concerning the average latency for all entities and has a substantially higher isolation index than the benchmarking strategies. This can be attributed to its capability of making the optimal resource adjustment in response to current traffic patterns, avoiding interference between slices and control of energy consumption. Despite its simplicity, Fixed Allocation (FA) degrades in dynamic environments, adding up to 2.9 ms of Slice C latency, while Heuristic Round Robin (HRR) only partially solves isolation problems but is less sensitive to transient bursts. Altogether, the combination of RL with multi-objective optimisation provides a highly efficient balanced paradigm to provide guaranteed QoS for various slicing scenarios characteristic for 5G.

4.4. Constraint Violations and Adaptation Times

The extensiveness of slicing for discovering constraint violations, such as latency maximums, system overloading, and isolation violations is assessed. By observing when and for how long such events occur, useful information is gained about the performance of the system when faced with unexpected load or demand of users. Table 4 presents the overall violations observed in every slice together with the time taken by the resource allocation mechanism to regain compliance. Overall, these results underscore the importance of adaptive orchestration and monitoring response in near real-time, especially, for mission-critical and highly congested applications.

Table 4. Constraint Violations and Adaptation Times Across 10 Experiments

Slice	Latency Violations (total)	Capacity Violations (total)	Isolation Violations (total)	Avg. Adaptation Time (s)	Max Adaptation <Time (s)
A	3	2	0	4.8	11.2
B	6	7	1	5.5	13.0
C	9	3	2	4.6	10.8

Among the data presented in Table 4, telemedicine (Slice C) exhibits the most frequent latency violations, with nine out of ten tests potentially due to unanticipated fluctuations in the demand for high-definition video consultations. Slice B (industrial IoT) records seven capacity violations, attributed to sudden bursts in the factory context where the number of devices can rapidly increase. Conversely, Slice A (autonomous vehicles) shows few breaches, indicating effective handling of ultra-reliable low-latency communication (URLLC) requirements. The orchestration system detects these violations and provides corrections, typically within approximately five seconds, thus minimizing service interruptions. This rapid adaptability, enabled by reinforcement learning and dynamic weighting of slice resources, demonstrates the framework's flexibility to deliver low QoS variation in real-world, multi-slice 5G networks.

4.5. Stakeholder Feedback on Service-Level Agreements (SLAs)

Key informant perspectives on the effectiveness and consistency of the slicing framework are qualitatively derived from 25 interviews across multiple industries: automotive, industrial, healthcare IT, telecommunications. The talks focused on the specific levels of certain SLAs, namely, latency, isolation, and coverage established in the service-providing environment and possible improvements that may be made in future implementations. The table below presents the summarized results where both the satisfaction levels and the challenges listed by each industry vertical give a good picture of how they perceive this framework. The insights are as diverse as the concerns and expectations that define end-user engagements across 5G networks.

Table 5. Stakeholder SLA Satisfaction and Suggested Improvements

Stakeholder Type	Sample Size	SLA Satisfaction (%)	Latency Concerns	Isolation Concerns	Suggestions
Automotive OEMs	6	85%	Medium	Low	Increase coverage range, advanced URLLC for highway use
Manufacturing Plant Mgmt.	5	90%	Low	Medium	Sub-slicing for real-time tasks, better integration w/ legacy
Healthcare IT Specialists	7	82%	High	Low	Enhanced encryption, HIPAA compliance, specialized bandwidth
Telecom Operators	7	93%	Low	Medium	Cross-operator orchestration, standard APIs, simplify multi-tenancy

Automotive OEMs emphasize the requirement of increased network coverage and the development of the highly reliable URLLC to address the fast-moving car traffic far from cities. The perceived satisfaction is mostly high among the manufacturing plant managers; however, they stress the need to sub-slice the requirements for the missions-critical processes, stressing on the need to have high-quality isolation in the real-time robotics systems. IT professionals across the health care sector have emphasized data as their key concern followed by privacy and regulatory standards (HIPAA, GDPR) and have demanded more encryption and network resources for telemedicine. At the same time, telecom operators believe in cross-operator orchestration frameworks and standardized application programming interfaces hoping for simplification of multi-tenancy structures. Taken together, these interviews indicate that program-specific, domain-specific enhancements are essential for the long-term satisfaction of the stakeholders in 5G slicing.

4.6. Post-Experiment Analysis: Reinforcement Learning Convergence

One of the primaries focuses of this study is reinforcement learning (RL) because it is a key mechanism to guarantee the adaptation of the slicing framework to changes in the demands of the network. Unlike traditional compute models and heuristic algorithms, RL dynamically monitors the state of each slice – latency, throughput, and isolation and, accordingly, proportionally, decides the consumption of CPU, memory and band-widths. The approach synchronizes the exploitation of the known-fine policy (the best policy) to maintain QoS stability and the exploitation of other policies to try new allocations. During repeated episodes, the RL agent learns how each change in one slice impacts the others and the approach converges to a near optimizing resource allocation solution as in (Mai et al. 2022; Khan et al. 2021; Azimi et al. 2022; Yan et al. 2023; Polyakov N.A. 2021).

The following is the 50 RL episodes which form one experiment in the present condensed study presented in tabular form for better understanding. Every episode is one round of interaction with the environment, we make a resource allocation decision, get the immediate feedback and the reward signal. The penalty decreases over time, which represents better decisions – such as meeting URLLC latency in Slice A or accommodating massive sensor bursts for mMTC in slice B, or handling high bandwidth telemedicine necessity in slice C. Through the calculated rolling average reward, the observed moment could be identified at which the agent converges and indicates its ability to achieve more than one SLA at once. This policy stability becomes useful in practical 5G/B5G scenarios where traffic patterns fluctuate drastically and have to address diverse device dynamics.

Table 6. Reinforcement Learning Convergence Over 50 Episodes

Episode	Instant Reward	Rolling Avg. Reward (Last 5 Eps)	Notes
1	-5.0	-5.00	Initial exploration; random allocations
2	-4.7	-4.85	Gradual policy refinement begins
3	-4.4	-4.70	Slight improvement in slice isolation
4	-4.8	-4.72	Small regression; testing new allocation step

Iranian Journal of Information Processing and Management

Episode	Instant Reward	Rolling Avg. Reward (Last 5 Eps)	Notes
5	-4.2	-4.62	Gains for telemedicine slice throughput
6	-4.0	-4.42	Reducing resource conflicts among slices
7	-3.7	-4.22	Faster exploration in URLLC settings
8	-4.1	-4.24	Temporary dip from IoT spikes
9	-3.5	-4.06	Marked improvement in eMBB stability
10	-3.9	-3.88	Fine-tuning CPU/memory allocations
11	-3.6	-3.76	Achieving sub-5 ms latency in Slice A consistently
12	-3.5	-3.68	Gradual smoothing of performance variance
13	-3.8	-3.66	Testing new exploration path; slight overhead
14	-3.3	-3.62	Gains in resource partitioning
15	-3.1	-3.54	Energy efficiency recognized in partial runs
16	-3.9	-3.56	Additional exploration; moderate dip in reward
17	-3.2	-3.48	Improved fairness across slices
18	-2.9	-3.32	Notable boost for IoT reliability
19	-3.3	-3.30	Minor regression due to sensor burst
20	-2.8	-3.26	Gains in telemedicine throughput
21	-3.0	-3.24	Quick correction of resource imbalance
22	-3.4	-3.28	System stress-tested with combined load
23	-3.1	-3.26	Efficient reallocation post-stress
24	-2.7	-3.20	Low-latency maintained; mMTC stable
25	-2.5	-3.14	Consistent improvement in slice isolation
26	-3.3	-3.02	Testing high concurrency scenario in IoT
27	-2.9	-2.96	Quick rebound; multi-slice synergy
28	-2.6	-2.80	Enhanced URLLC predictability
29	-3.0	-2.86	Momentary dip; telemedicine bandwidth reconfig
30	-2.4	-2.68	Notable jump in eMBB performance
31	-2.7	-2.72	Handling off-peak transitions effectively
32	-2.5	-2.64	Few latency spikes in Slice A observed
33	-2.8	-2.66	Re-balancing CPU for IoT demands
34	-2.2	-2.52	Substantial improvement in reliability
35	-2.1	-2.46	Gains in throughput and energy optimization
36	-2.9	-2.50	Testing fallback policies for unexpected loads
37	-2.3	-2.46	Quick adaptation confirms policy resilience
38	-2.0	-2.38	Minimizing slice contention successfully
39	-2.4	-2.34	Slight overhead from repeated

Episode	Instant Reward	Rolling Avg. Reward (Last 5 Eps)	Notes
			reconvergence
40	-2.1	-2.26	Reinforced stability across slices
41	-1.9	-2.14	Gains in slice isolation and latency simultaneously
42	-2.2	-2.10	Minor set-back, easily corrected
43	-1.7	-2.06	Close to stable plateau; high user satisfaction
44	-1.8	-2.02	Telemedicine slice now reliably at 70+ Mbps
45	-2.0	-1.96	Adjusting for IoT sensor surges
46	-1.5	-1.84	Best energy isolation balance to date
47	-1.7	-1.74	Fine-tuning resource prioritization
48	-1.9	-1.78	Brief exploration leads to slight dip
49	-1.6	-1.66	Settling into final policy equilibrium
50	-1.4	-1.62	Convergence achieved; robust multi-slice policy

In the first 10 episodes, it runs mainly at the exploration phase, and thus, the system achieves considerably negative rewards because of improper resource allocations especially the URLLC and eMBB slices with high latency sensitivity. This period also demonstrates that the developed system was initially unable to cope with diverse slicing designs.

The number of resources utilized is as follows When the episodes get to the mid-phase - 11-30, the Line average reward starts to rise which shows that the agent is doing a better job in managing resources. Preventive reallocation strategies help manage traffic increase in IoT apps, or the rise in telemedicine, for example. This phase is intended to present the agent's increasing ability to manage resource allocation in relation to various slices' needs in real time.

In the last episodes (31-50) policy fluctuations decrease, while rewards become more or less steady at or above -1.6. The agent asserts near-5 ms latencies for machine learning and partitioning for AV applications as well as consistent throughput significantly more than 65 Mbps for telemedicine services. The RL mechanism learns and keeps the overall system in stable state despite the reward dips happening once a traffic pattern change.

These 50 episodes represent how RL approaches drive the allocation policy from random exploration through to a final RL policy capable of achieving near-optimal resource allocation and key SLA objectives. This shows that the agent has a systematic way of minimizing the latencies violations, throughput limitation and inter-slice interference thus proving that this approach is flexible. All such capabilities are crucial for efficient operation in 5G/B5G multi-slice environments where traffic fluctuations and diverse devices require fast and smart handling. Further, the final policy integrates requirements for energy efficiency improvements and demonstrates that the RL paradigm is capable of achieving high performance levels alongside sustainable development goals. This convergence analysis bears out the possibility and effectiveness of machine learning-based solution to next-generation network slicing.

5. Discussion

The results discussed in this work demonstrate the effectiveness of using reinforcement learning (RL) and multi-objective optimization approaches for managing network slices across various industrial and social use cases. Based on the three functions of theory—(a) descriptive, (b) explanatory, and (c) prognostic—this discussion compares the findings with prior research, elaborates on the study's weaknesses, and outlines further research directions.

The study's descriptive function highlights how the proposed slicing framework systematically assigns resources to three different categories of applications: URLLC, mMTC, and eMBB, addressing application requirements of high dependability, minimal latency, varied and extensive machine-type connections, and significantly improved mobile broadband speeds, respectively. In agreement with Luo et al.'s (2023) work, where a field trial of network slicing in 5G and PON-enabled industrial networks was performed (Luo et al. 2023), this study shows that the adaptive orchestration design proposed in this paper can effectively support the requirements of distributed industrial IoT traffic. Additionally, achieving consistent latencies of less than 5 ms for URLLC slices underscores the dominance of real-time performance requirements highlighted by Nota et al. (2022).

The explanatory aspect of this research elucidates why reinforcement learning-based orchestration is superior to static or heuristic approaches in dynamic contexts. The constantly readjusted KPIs of slice K, namely latency, throughput, and reliability, enable the RL agent to sense changes in traffic profiles and foresee future congestion. This predictive capacity aligns with what Mai et al. (2023) observed regarding transfer reinforcement learning;

hence, continuous adaptation is crucial for distributed network slicing in industrial IoT (Mai et al. 2022). Moreover, the cross-slice authentication method described by Yadav et al. (2023) shows that security measures can be relatively easily incorporated into learning-based resource management approaches while maintaining high levels of security without compromising performance (Yadav et al. 2023).

This article, however, diverges from the position taken by Dick et al., asserting that no single layer can function autonomously since efficient orchestration layers define the successful interoperability of slices, an idea consistent with our strategic slicing management results in Lieto et al. (2022).

In addition to the descriptive and explanatory roles, the research also provides a predictive role, showing how the RL-based solution would perform when traffic load, topological complexity, or the number of tenants increases. For instance, we expect that the extension of new slices, such as critical healthcare or advanced driver-assistance systems, would continue to be manageable if the RL framework is well-trained on available datasets and connected MEC systems are solidly integrated (Ksentini and Frangoudis, 2020). Notably, blockchain-enhanced approaches, such as the BENS-B5G solution proposed by Singh et al. (2022), can help eliminate issues of trust and transparency when several operators use the same infrastructure. However, despite positive findings, the following limitations emerge. First, the computational complexity of RL is high, particularly for training when reconfigurations are necessary in near real-time for URLLC slices. Khan et al. (2021) built an End-to-End framework for vehicular ad-hoc scenarios, but resource constraints increase as more slices request limited resources, which also concluded the survey on intelligent network slicing management for Industrial IoT. Also, while dynamic orchestration is supported in the framework, it does not describe all necessary security and privacy aspects in detail. Special cross-network-slice authentication mechanisms discussed by Yadav et al. (2023) are required to ensure end-to-end reliability, but it is important to note that any such protocol can add extra latency or overhead if not fine-tuned appropriately.

Another limitation is the diverse hardware configurations across vendors. As Azimi et al. (2022) note, self-learning applications for RAN slicing face compatibility challenges when integrating various radio technologies or solutions from different vendors. The testbed for this study was moderately sized and partially homogeneous. Therefore, to build on this effectiveness in future large-scale multi-operator setups, more standardization and federation measures similar to those explained by Babbar et al. (2021) will be required. Finally, the energy optimization aspect of slicing, which was covered by the weighted objectives, deserves further study (Babbar et al. 2022). Researchers such as Afolabi et al. (2021) and Yan et al. (2023) have developed broader models that include green metrics and massive MIMO limitations. Addressing these aspects is critical for meeting sustainability needs within rapidly growing 5th generation (5G) and beyond 5G (B5G) environments.

The work proposed here on an RL-based slicing framework shows promising results for slice orchestration, aligning with previous findings (Luo et al. 2023; Lieto et al. 2022; Nota et al. 2022). Future work should focus on enhancing the solution's scalability, improving cross-slice security, and extending multi-domain orchestration for large and geographically diverse deployments. This research not only contributes to the expansion of available knowledge but also helps define strategies for more rational, creative, and secure approaches to the network slicing paradigm.

6. Conclusion

The findings presented provide a comprehensive perspective on how a learning-integrated resource management strategy can enhance the network slicing scenario in next-generation communication systems. By evaluating latency, throughput, reliability, and isolation across several industry verticals, the research thoroughly investigates the challenges arising from a multi-slice environment with UL-RLLC, eMBB, and mMTC. The approach employed enables real-time adjustments to ongoing traffic flows, making them resilient to variations and the diverse demands of different devices. This underscores the significant role of learning-based orchestration in intelligent 5G and Beyond-5G systems.

The results demonstrate that both strategic orchestration and real-time adaptation can reduce the likelihood of performance stagnation in mission-critical applications while concurrently improving energy efficiency. By leveraging statistical models, the framework can optimize conflicting service requirements, delivering high and reliable performance in the present and incorporating advanced resource management techniques essential for large-scale industrial and societal deployments in the future. This marks a significant improvement over static or solely heuristic resource allocation strategies, promising better reliability and responsiveness during load-balancing scenarios. Additionally, end-to-end slicing is introduced to maintain KPIs across control, data, and orchestration layers, ensuring safety and efficiency while accommodating more KPI possibilities and better alignment with forthcoming industry standards and implementations.

Despite these improvements, further enhancements are necessary due to complexities such as interfacing the new system architecture with diverse hardware platforms, emerging security issues, and legal requirements surrounding transactions involving personal and sensitive data. Future work holds significant potential, particularly in extending the research to more elaborate distributed learning paradigms, cross-domain integration, and the selection of comprehensive frameworks that include sustainable metric calculations. As networks expand to address high-density user environments, integrating cloud and edge intelligence while maintaining dynamic, learning-driven resource management becomes essential. Continued cooperation among the research community, industry stakeholders, and standards-making organizations will be crucial in achieving higher levels of conformity across nations, thus fulfilling the promise of next-generation slicing, provided it is not impeded by high complexity and security considerations.

The article illustrates how reinforcement learning, combined with multi-objective optimization, can meet the complex and dynamic communication needs of today and tomorrow. The practical applications span various areas, including self-driving car systems, factories, hospitals, and emergency services. Consequently, sustained efforts toward progressively enhancing and synergizing these approaches for the development of robust commercial and mission-critical networks will pave the way for smarter digital platforms, fostering innovation and differentiated services.

References

Afolabi, I., Bagaa, M., Boumezer, W., and Taleb, T. (2021). Toward a Real Deployment of Network Services Orchestration and Configuration Convergence Framework for 5G Network Slices. *IEEE Network,* 35 (1), 242-250. https://doi.org/:10.1109/MNET.011.2000146

Afrianto, Y., Sukoco, H., and Wahjuni, S. (2018). Weighted Round Robin Load Balancer to Enhance Web Server Cluster in OpenFlow Networks. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*.

Afrianto, Y., and Wahjuni, S. (2018). Weighted Round Robin Load Balancer to Enhance Web Server Cluster in OpenFlow Networks. *Telkomnika*

(Telecommunication Computing Electronics and Control), 16, 1402-1408. https://doi.org/:10.12928/TELKOMNIKA.v16i3.5601

Ageyev, D., Yarkin, D., and Nameer, Q. (2014). Traffic aggregation and EPS network planning problem. *First International Scientific-Practical Conference Problems of Infocommunications Science and Technology,* 14-17 Oct. https://doi.org/:10.1109/INFOCOMMST.2014.6992316.

Azimi, Y., Yousefi, S., Kalbkhani, H., and Kunz, T. (2022). Applications of Machine Learning in Resource Management for RAN-Slicing in 5G and Beyond Networks: A Survey. *IEEE Access,* 10, 106581-106612. https://doi.org/:10.1109/ACCESS.2022.3210254

Babbar, H., Rani, S., AlZubi, A. A., Singh, A., Nasser, N., and Ali, A. (2022). Role of Network Slicing in Software Defined Networking for 5G: Use Cases and Future Directions. *IEEE Wireless Communications,* 29 (1), 112-118. https://doi.org/:10.1109/MWC.001.2100318

Chen, M., Yang, J., Hao, Y., Mao, S., and Hwang, K. (2017). A 5G Cognitive System for Healthcare. *Big Data and Cognitive Computing*, 1 (1). https://doi.org/:10.3390/bdcc1010002.

Escolar, A. M., Alcaraz-Calero, J. M., Salva-Garcia, P., Bernabe, J. B., and Wang, Q. (2021). Adaptive Network Slicing in Multi-Tenant 5G IoT Networks. *IEEE Access,* 9, 14048-14069. https://doi.org/:10.1109/ACCESS.2021.3051940

Felici-Castell, S., García-Pineda, M., Segura-Garcia, J., Fayos-Jordan, R., and Lopez-Ballester, J. (2021). Adaptive live video streaming on low-cost wireless multihop networks for road traffic surveillance in smart cities. *Future Generation Computer Systems,* 115, 741-755. https://doi.org/:https://doi.org/10.1016/j.future.2020.10.010

Foukas, X., Patounas, G., Elmokashfi, A., and Marina, M. K. (2017). Network Slicing in 5G: Survey and Challenges. *IEEE Communications Magazine,* 55 (5), 94-100. https://doi.org/:10.1109/MCOM.2017.1600951

Hussain, F., Hussain, R., Anpalagan, A., and Benslimane, A. (2020). A New Block-Based Reinforcement Learning Approach for Distributed Resource Allocation in Clustered IoT Networks. *IEEE Transactions on Vehicular Technology,* 69 (3), 2891-2904. https://doi.org/:10.1109/TVT.2020.2965796

Khan, A. A., Abolhasan, M., Ni, W., Lipman, J., and Jamalipour, A. (2021). An End-to-End (E2E) Network Slicing Framework for 5G Vehicular Ad-Hoc Networks. *IEEE Transactions on Vehicular Technology,* 70 (7), 7103-7112. https://doi.org/:10.1109/TVT.2021.3084735

Khanh, Q. V., Hoai, N. V., Manh, L. D., Le, A. N., and Jeon, G. (2022). Wireless Communication Technologies for IoT in 5G: Vision, Applications, and Challenges. *Wireless Communications and Mobile Computing,* 2022, 3229294. https://doi.org/:10.1155/2022/3229294

Ksentini, A., and Frangoudis, P. A. (2020). Toward Slicing-Enabled Multi-Access Edge Computing in 5G. *IEEE Network,* 34 (2), 99-105. https://doi.org/:10.1109/MNET.001.1900261

Lieto, A., Malanchini, I., Mandelli, S., Moro, E., and Capone, A. (2022). Strategic Network Slicing Management in Radio Access Networks. *IEEE Transactions on Mobile Computing,* 21 (4), 1434-1448. https://doi.org/:10.1109/TMC.2020.3025027

Luo, Y., Jiang, M., Zhang, D., and Effenberger, F. (2023). Field Trial of Network Slicing in 5G and PON-Enabled Industrial Networks. *IEEE Wireless Communications,* 30 (1), 78-85. https://doi.org/:10.1109/MWC.002.2200215

Mai, T., Yao, H., Zhang, N., He, W., Guo, D., and Guizani, M. (2022). Transfer Reinforcement Learning Aided Distributed Network Slicing Optimization in Industrial IoT. *IEEE Transactions on Industrial Informatics,* 18 (6), 4308-4316. https://doi.org/:10.1109/TII.2021.3132136

Nota, A., Saidi, S., Overbeck, D., Kurtz, F., and Wietfeld, C. (2022). Providing Response Times Guarantees for Mixed-Criticality Network Slicing in 5G. *Design, Automation & Test in Europe Conference & Exhibition (DATE),* 14-23 March. https://doi.org/:10.23919/DATE54114.2022.9774503.

Polyakov, N., Yarkina, N., and Samouylov, K. (2021). A simulator for analyzing a network slicing policy with SLA-based performance isolation of slices. *Discrete and Continuous Models and Applied Computational Science,* 29, 36-52. https://doi.org/:10.22363/2658-4670-2021-29-1-36-52

Polyakov N.A., Y. N. V., Samouylov K.E. (2021). A simulator for analyzing a network slicing policy with SLA-based performance isolation of slices. *Discrete and Continuous Models and Applied Computational Science* 29 (1), 36-52. https://doi.org/10.22363/2658-4670-2021-29-1-36-52

Qasim, N., Jawad, A., Jawad, H., Khlaponin, Y., and Nikitchyn, O. (2022). Devising a traffic control method for unmanned aerial vehicles with the use of gNB-IOT in 5G. *Eastern-European Journal of Enterprise Technologies,* 3, 53-59. https://doi.org/:10.15587/1729-4061.2022.260084

Salih, M. M., Khaleel, B. M., Qasim, N. H., Ahmed, W. S., Kondakova, S., and Abdullah, M. Y. (2024). Capacity, Spectral and Energy Efficiency of OMA and NOMA Systems. *35th Conference of Open Innovations Association (FRUCT).* https://doi.org/:10.23919/FRUCT61870.2024.10516394.

Sieliukov A.V., Q. N. H., Khlaponin Y.I. (2022). Conceptual model of the mobile communication network. *The Workshop on Emerging Technology Trends on the Smart Industry and the Internet of Things «TTSIIT»*, 20-22.

Singh, S., Babu, C. R., Ramana, K., Ra, I.-H., and Yoon, B. (2022). BENS−B5G: Blockchain-Enabled Network Slicing in 5G and Beyond-5G (B5G) Networks. *Sensors*, 22 (16). https://doi.org/:10.3390/s22166068.

Wang, H., Wu, Y., Min, G., and Miao, W. (2022). A Graph Neural Network-Based Digital Twin for Network Slicing Management. *IEEE Transactions on Industrial Informatics,* 18 (2), 1367-1376. https://doi.org/:10.1109/TII.2020.3047843

Wu, Y., Dai, H. N., Wang, H., Xiong, Z., and Guo, S. (2022). A Survey of Intelligent Network Slicing Management for Industrial IoT: Integrated Approaches for Smart Transportation, Smart Energy, and Smart Factory. *IEEE Communications Surveys & Tutorials,* 24 (2), 1175-1211. https://doi.org/:10.1109/COMST.2022.3158270

Yadav, A. K., Wijethilaka, S., Braeken, A., Misra, M., and Liyanage, M. (2023). An Enhanced Cross-Network-Slice Authentication Protocol for 5G. *IEEE Transactions on Sustainable Computing,* 8 (4), 555-573. https://doi.org/:10.1109/TSUSC.2023.3283615

Yan, D., Ng, B. K., Ke, W., and Lam, C. T. (2023). Deep Reinforcement Learning Based Resource Allocation for Network Slicing With Massive MIMO. *IEEE Access,* 11, 75899-75911. https://doi.org/:10.1109/ACCESS.2023.3296851