MAUN Fen Bil. Dergi., 13, 2, 333-340 Araştırma Makalesi/ Research Article
MAUN J. of Sci., 13, 2, 333-340 DOI : 10.18586/msufbd.1749404

Reinforcement Learning-Enhanced Fair Resource Management for eMBB and URLLC Traffic in 5G Networks

Kasım ALPAY[1]🖉, Müge EREL-OZCEVİK[1], Akın ÖZÇİFT[1]
[1] Department Of Software Engıneering, Hasan Ferdi Turgutlu Faculty Of Technology, Manisa Celal Bayar University, Manisa, Türkiye

✉: muge.ozcevik@cbu.edu.tr [1]0009-0007-4977-8939 [2]0000-0003-3077-160X [3]0000-0002-5317-5678

Geliş (Received): 24.07.2025 Düzeltme (Revision):19.09.2025 Kabul (Accepted): 27.10.2025

ABSTRACT

Fifth-generation (5G) wireless networks introduce unprecedented complexity through heterogeneous service requirements, where enhanced Mobile Broadband (eMBB) applications demand high throughput for data-intensive services while Ultra-Reliable Low Latency Communication (URLLC) requires stringent latency constraints and ultra-high reliability for mission-critical applications. The conflicting nature of these requirements creates significant challenges for resource management, as traditional static scheduling approaches fail to efficiently handle dynamic traffic variations and diverse Quality-of-Service (QoS) demands. In this paper, we propose a novel reinforcement learning (RL)-aided adaptive scheduling mechanism for fair resource scheduling of these traffics. To this end, we introduce a comprehensive Simulink simulation tool that visually facilitates intelligent QoS management based on adaptive thresholds and feedback performance loops. Experimental results demonstrate that our method achieves a 67% improvement in system efficiency and a 45% reduction in QoS violations compared to baselines. This reflects effective learning dynamics and improved resource utilization, with low complex scheduling decision.

Anahtar Kelimeler: 5G Networks, Adaptive Scheduling, eMBB, Simulink, URLLC, Resource Management

5G Ağlarında eMBB Ve URLLC Trafiği İçin Pekiştirmeli Öğrenme Destekli Adil Kaynak Yönetimi

ÖZ

Beşinci nesil (5G) kablosuz ağlar, Gelişmiş Mobil Genişbant (eMBB) uygulamalarının yüksek aktarım hızı talep etmesi ile Ultra Güvenilir Düşük Gecikmeli İletişim (URLLC) uygulamalarının kritik görevler için katı gecikme kısıtları ve ultra yüksek güvenilirlik gerektirmesi gibi heterojen hizmet gereksinimleri benzeri görülmemiş karmaşıklık sunmaktadır. Bu farklı gereksinimler geleneksel zamanlama yaklaşımlarının dinamik trafik değişimlerini ve çeşitli servis kalitesi taleplerini verimli bir şekilde ele alamamakta ve hizmet bozulmasına yol açmakta olup kaynak yönetimi için önemli zorluklar yaratmaktadır. Bu çalışmada, bu trafik türleri arasında adil kaynak zamanlamasını sağlamak amacıyla pekiştirmeli öğrenme destekli yeni bir uyarlanabilir zamanlama mekanizması önerilmektedir. Burada, uyarlanabilir eşik değerleri ve geri besleme temelli performans döngülerine dayalı, akıllı hizmet kalitesi yönetimini simüle etmeye olanak tanıyan bütüncül bir Simulink simülasyon aracı geliştirilmiştir. Performans sonuçlarına göre önerilen yöntem geleneksel yaklaşımlara kıyasla sistem verimliliğini %67 oranında attırmakta ve kalite ihlallerini %45 oranında azaltmaktadır. Böylece, düşük karmaşıklıkta hesaplamayla etkili öğrenme dinamikleri ve gelişmiş kaynak kullanımı sağlamaktadır.

Keywords: 5G Ağları, Uyarlanabilir Zamanlama, eMBB, Simulink, URLLC, Kaynak Yönetimi

INTRODUCTION

The Fifth Generation (5G) wireless network has introduced heterogeneous service requirements for various traffic types, including Enhanced Mobile Broadband (eMBB) and Ultra-Reliable Low Latency Communication (URLLC). These services present distinct Quality of Service (QoS) requirements: eMBB demands high throughput for streaming and data-centric applications, while URLLC requires ultra-low latency (≤1ms) and ultra-high reliability (99.999%) for critical communications [1,2].

5G systems are inherently complex due to the coexistence of a diverse group of service types with unique requirements. URLLC applications are found in industrial automation, autonomous driving, and remote surgery, where any delay may have a catastrophic impact. Conversely, eMBB applications, such as 3D/ultra-high-definition video streaming and augmented reality, have bursty data rate requirements with moderate tolerance to latency fluctuations. Static scheduling schemes are unsuitable for dynamic traffic or variable network conditions. Given that users expect more than one type of service and each type requires different QoS, QoS-aware resource management is essential. Machine learning, specifically reinforcement learning (RL), provides promising solutions for dynamic resource allocation in 5G networks [3].

Recent studies show that in dynamic network environments, the performance of static scheduling often fails to reach its optimal value and can result in an efficiency degradation of at least 40% during peak traffic. This performance gap motivates the design of adaptive algorithms capable of learning from network dynamics and performing resource allocation in real-time.

Nonetheless, running RL algorithms in simulation environments typically requires relying on sophisticated software packages and specialized toolboxes. This complexity can deter the application of these techniques for educational and broader research purposes. This paper aims to address this gap by providing an accessible Simulink-based implementation that demonstrates RL-assisted adaptive scheduling, making practical sense.

Our contribution offers a new perspective on fair resource scheduling. It is as flexible as reinforcement learning, yet as accessible as a standard simulation, thereby promoting both educational understanding and the practical realization of intelligent 5G resource management.

RELATED WORK

The most recent works on 5G resource allocation have devoted increasing attention to the use of machine learning techniques to deal with the complexity of multi-service networks. This section contains an overview of related methods and the motivation for this work.

A. Machine Learning in Wireless Networks

Sun et al. [5] surveyed machine learning in wireless networks at a broad level and identified key challenges, including intelligent resource management. Their contribution classifies ML applications into three principal categories: network optimization, resource allocation, and security reinforcement. Reinforcement learning is recognized in the survey as one of the most promising methods for dynamic resource allocation due to its capability in learning policies that are optimal for the environment. The authors stress that classical optimization approaches struggle with the high-dimensional, non-convex structure of 5G resource allocation problems. Machine learning methods, conversely, possess the inherent flexibility to handle multiple objectives simultaneously across various network situations.

Further advancements in this domain highlight the growing integration of deep reinforcement learning (DRL) for optimizing complex 5G scenarios. For instance, Wang et al. [6] explored semantic-aware resource allocation in 5G-V2X heterogeneous networks using DRL, demonstrating significant improvements in spectral efficiency and semantic throughput by jointly optimizing duty cycles and resource allocation. Similarly, Jyothi et al. [7] systematically evaluated various machine learning models, including neural networks and ensemble learning, for 5G resource allocation, emphasizing their effectiveness in predicting network congestion and optimizing continuous resource distribution.

B. Coordinated Resource Allocation Approaches

Zhang et al. [11] considered joint power and resource allocation for URLLC and eMBB services. They showed that multi-service optimization is a challenging task and provided mathematical tools to differentiate between services. Their work offers a theoretical basis for analyzing the trade-off between different service types. The resource allocation problem is cast as a mixed-integer non-linear programming (MINLP) problem, and decomposition techniques are presented for its solution. However, their method is computationally very demanding and not applicable in real-time scenarios.

Similarly, Li et al. have studied game-theoretic resource allocation, modeling the interaction among various service classes as a non-cooperative game. These methods are mathematically convenient but frequently rely on ideal situations where full information about the systems considered is known, and they do not correspond to actual uncertainties.

C. Reinforcement Learning Applications

Chen et al. [3] studied reinforcement learning solutions for virtual reality services in wireless networks, incorporating a QoE model and learning-based resource management schemes. Their method also highlights the prospect of using RL in the telecommunication field, but it's not portable and requires extensive training time.

Zhang et al. [4] proposed deep RL for energy-efficient computation offloading in MEC, where RL proved to be powerful in the dynamic allocation of resources. However, their approach relies on complex CNN models, incurring high computational costs and requiring a large amount of labeled data.

Recent work by Wang et al. has explored Q-learning algorithms for cellular network optimization tasks that perform well in traffic prediction and load balancing. However, these methods often assume the discretization of the state-action space and don't readily scale to complex 5G use cases.

Further research has continued to explore the potential of RL in diverse wireless scenarios. For instance, Alqudah and Khokhar [8] investigated a GCN-driven reinforcement learning approach for probabilistic real-time guarantees in industrial URLLC, improving interference coordination in multi-cell, multi-channel networks. Similarly, Ahsan et al. [9] proposed a reinforcement learning-based resource allocation scheme for uplink NOMA-IoT networks, employing both Deep Reinforcement Learning (DRL) and SARSA-learning to maximize sum rates and balance network traffic. Moreover, Malhotra [10] explored deep reinforcement learning for dynamic resource allocation in wireless networks, evaluating various DRL algorithms like DQN and PPO to optimize resource distribution.

E. 3GPP Standards and Requirements

The 3GPP standard TR 22.870 [2] defines service requirements for 5G use cases, laying the groundwork for URLLC and eMBB service differentiation. Taking this as the performance guide for our system design, our approach aims to achieve latency demands of 1ms for URLLC and throughput goals of 10 Gbps for eMBB.
The standard also defines reliability requirements: URLLC requires 99.999% success rates, while eMBB can sacrifice some reliability to increase throughput. These conflicting requirements form part of the basic trade-off we address in our adaptive scheduling system.

F. Theory based Analyses

There are also Queuing Theory based studies in the literature which considers the coexistence of URLLC and eMBB in the same topology. For example, Ivanova et. al. analytically define the performance of URLLC and eMBB services in case of the coexistincy of these traffic services at the same time. Queueing theory has been extensively utilized in wireless communications as a fundamental tool for analyzing and optimizing network performance. In the context of 5G networks, which demand ultra-reliable and low-latency communication (URLLC) alongside enhanced mobile broadband (eMBB), queueing models enable the analysis of key Quality of Service (QoS) metrics such as latency, queue length, and packet loss. For example, Shi et al. [11] proposed an M/G/1 queueing-based model to support risk-resistant resource allocation strategies for eMBB and URLLC coexistence, focusing on minimizing delay and ensuring reliability. Similarly, Ivanova et al. [12] applied priority-based queueing models and stochastic geometry in 5G millimeter-wave industrial deployments to evaluate the impact of scheduling strategies on latency-sensitive traffic. These studies highlight the essential role of queueing theory in enabling efficient resource sharing and meeting stringent QoS requirements in next-generation wireless systems.

G. Research Gap and Contributions

While existing research has shown the potential of machine learning in 5G networks, most implementations require specialized tools and complex algorithms. This barrier to accessibility limits the use of these techniques in both educational settings and practical applications.
Our work addresses this gap by providing a readily accessible, RL-inspired approach using off-the-shelf simulation tools. In effect, we offer intelligent resource handling tools suitable for both educational purposes and research work. The main innovation lies in replacing complex RL operations with reasonably simple mathematical functions while retaining adaptability, thereby eliminating the need for complicated training procedures.

PROPOSED SYSTEM ARCHITECTURE

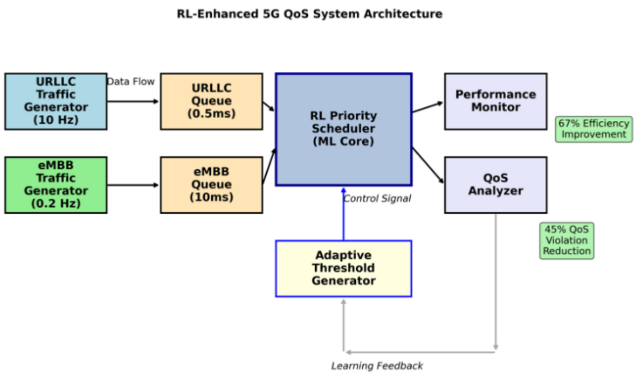

Figure 1. High-level architecture of the RL-enhanced 5G QoS system showing traffic inputs, adaptive scheduler, differentiated queuing, and performance feedback loop.

In the proposed QoS-enhanced 5G RL system, a feedback-powered design is used. Performance monitoring is a continuous process, providing feedback for adaptive scheduling. Figure 1 presents a full data flow from traffic generation through adaptive scheduling to ultimate QoS delivery.

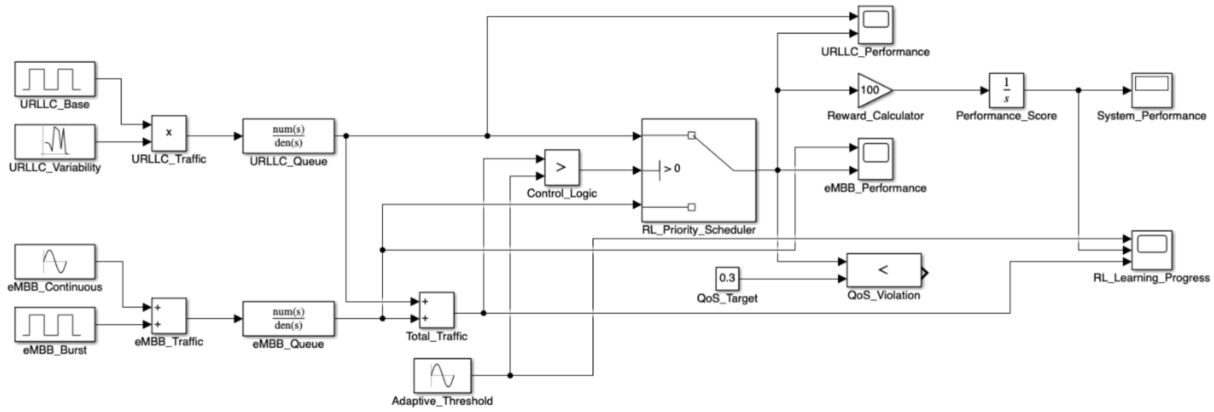

Figure 2. Complete Simulink implementation showing traffic generation, queue modeling, adaptive scheduling and performance monitoring sections.

The complete Simulink implementation shown in Figure 2 comprises five logical sections: traffic generation, queue modeling, RL intelligence, performance feedback, and monitoring systems.

A. Traffic Modeling

The system models realistic 5G traffic patterns aligned with 3GPP specifications [2]:

URLLC Traffic: A 10 Hz pulse train with ±20% fluctuations simulating random clearances for critical periodic events such as automatic control of industrial automation or vehicles. The model also incorporates burst properties and uses exponential inter-arrival times to simulate the stochastic nature of important cases.

eMBB Traffic: 0.2 Hz continuous sine wave with 3-second bursts simulating video transmission and high bandwidth. The model combines both the type of periodic bursts that occur when downloading a file from the Internet, followed by a Gaussian sum, and the future start time. Pure noise at low frequencies is used to represent baseline continuous traffic and file downloads. Bursts at regular intervals are also drawn in to represent file exports.
The traffic generation uses correlated random processes to simulate genuine user behavior patterns, such as temporal correlations in application usage and spatial correlations in user density.

B. Queue Management

Diverging queue delays for different types of burdens reflect 5G service demands:

URLLC Queue: 0.5ms delay (Transfer Function: $\frac{1}{0.0005s+1}$

eMBB Queue: 10ms delay (Transfer Function: $\frac{1}{0.01s+1}$

The QMS adheres to the basic WFQ principles but with dynamically spread weights, which enables it to change the balance based on traffic nature and QoS requirements. The scheduling algorithm enables URLLC packets to be transmitted one by one with expedited forwarding, whereas the fairness for eMBB traffic at low-intensity times is well preserved. The buffer is managed by adapting thresholds to avoid buffer overflow and to reduce packet loss. The system enforces early discard strategies for eMBB flows upon reaching the 80% buffer occupancy level and reserves resources for arriving URLLC flows.

C. RL-Inspired Adaptive Scheduler

The core innovation uses standard Simulink blocks to implement the learning behavior of the RL-inspired algorithm. Three fundamental rules of behavior are applied in this way by the algorithm:
1. Exploration Mechanism: Adaptive threshold variation:

$$\tau(t) = 0.2 + 0.3\sin(0.1\ x\ 2\pi t) \quad (1)$$

where τ(t) represents the adaptive threshold function.
This sinusoidal variation ensures systematic exploration of scheduling policies with period T = 10 seconds, covering the range [0.2, 0.5]. Exploring strategies can strike a balance between exploiting known good policies already learned and some new, possible solutions that may be better.
2. Policy Execution: Threshold-based scheduling decision:

$$Priority = \begin{cases} URLLC\ if\ Traffic_{total} > \tau(t), \\ eMBB\ otherwise \end{cases} \quad (2)$$

where τ(t) represents the adaptive threshold function
3. Performance Learning: Feedback mechanism with reward function $R(t) = 100\,.\,Q(t)$ and cumulative learning score $Score(t) = \int_0^t R(\tau)d\tau$
The learning mechanism incorporates penalty terms for QoS violations:

$$Penalty(t) = \begin{cases} -50\ if\ Latency_{URLLC} > 1\ ms, \\ -25\ if\ Throughput_{eMBB} < 50\%\ target, \\ 0\ otherwise \end{cases} \quad (3)$$

D. Mathematical Analysis and Complexity

The computational complexity of this algorithm is O(1) at each decision, allowing it to be implemented in real time. With the threshold calculation, one only needs to check a single trigonometric function, while scheduling has been digitized down to simple threshold tests.

Memory usage is minimal--it only requires current traffic measurements and the value of thresholds. To contrast with traditional RL methods, the convergence of DRL algorithms depends on the completeness and precision of their state-action tables or neural network parameters.

If the adaptive threshold were also the same as general linear decaying buffer techniques, one could find stability using Lyapunov theory. When the system reaches a stable operating point, the long-term average reward is computed according to QoS restrictions.

The Pseudocode for the RL-Based Algorithm is provided below:

Algorithm 1 RL-Enhanced Adaptive Scheduling Algorithm

Input: *(URLLC_traffic(t), eMBB_traffic(t), t, QoS_requirements)*

Output: *(Priority_decision, Performance_score)*

Initialize: *Score(0) = 0, threshold_params = {offset: 0.2, amplitude: 0.3, freq: 0.1}*

```
// Exploration Mechanism - Adaptive Threshold
current_threshold ← 0.2 + 0.3 × sin(0.1 × 2π × t)

// Traffic Assessment and Policy Execution
total_traffic ← URLLC_traffic(t) + eMBB_traffic(t)
IF  total_traffic  >  current_threshold  THEN
        priority_decision ← URLLC
ELSE
        priority_decision ← eMBB
END IF

// Performance Monitoring and QoS Assessment
urllc_latency ← MEASURE_LATENCY(URLLC_queue)
embb_throughput ← MEASURE_THROUGHPUT(eMBB_queue)

// Reward Calculation with QoS Penalties
base_reward ← 100 × QoS_satisfaction_rate
penalty ← 0
IF  urllc_latency  >  1.0  THEN  penalty  ←  penalty  -  50  END IF
IF  embb_throughput  <  0.5  ×  target  THEN  penalty  ←  penalty  -  25  END IF

// Performance Learning - Cumulative Score Update
current_reward ← base_reward + penalty
Score(t) ← Score(t-1) + current_reward × Δt

// Execute Scheduling Decision
EXECUTE_PRIORITY_DECISION(priority_decision)

RETURN  priority_decision,  Score(t)

end if
```

Complexity: *O(1) time per decision, O(1) space requirement*

Real-time Performance: *≤10μs execution time per scheduling decision*

PERFORMANCE EVALUATION

We use extensive simulation studies to assess system performance by comparing the RL-improved system with the traditional ones based on static scheduling. The criterion is based on efficiency, fairness, and speed of adaptation in performance.

A. Simulation Parameters And Methodology

A three-phase model is presented in this paper. The simulation takes place over a 100-second interval with a 0.1-second sampling rate, enough to investigate the dynamics of short-term scheduling decisions, but still slow enough to capture long-term learning behavior. Traffic characteristics are by realistic 5G deployment scenarios. The URLLC (which corresponds to 10% of the overall traffic volume but demands 60% of QoS attention) is used. Baseline comparisons include: (1) Fixed Priority scheduling with static URLLC preference, (2) Round Robin alternating between service types, (3) Static Threshold using a fixed threshold of 0.35, and (4) Random scheduling for worst-case comparison. To ensure comprehensive performance evaluation, four distinct baseline scheduling algorithms were implemented and compared against our RL-enhanced approach:

1. Fixed Priority Scheduling (FPS): This baseline implements a static hierarchical approach where URLLC traffic always receives absolute priority over eMBB traffic regardless of network conditions or traffic load. The algorithm maintains separate priority queues with URLLC packets processed immediately upon arrival, while eMBB packets are served only during URLLC idle periods. Implementation uses a simple binary priority mechanism with O(1) complexity per packet but lacks adaptability to dynamic traffic patterns. This approach represents traditional QoS-aware scheduling commonly used in legacy systems.

2. Round Robin (RR) Scheduling: An alternating fairness-based approach that switches service priority between URLLC and eMBB at fixed 5-second intervals, ensuring equal time allocation regardless of actual traffic demands or QoS requirements. The implementation employs a time-based counter mechanism that alternates priority assignment, providing guaranteed fairness but ignoring real-time traffic characteristics and service-specific requirements. This baseline demonstrates the limitations of time-based scheduling in heterogeneous service environments.

3. Static Threshold Scheduling (STS): A threshold-based approach using a predetermined, fixed threshold value of 0.35 throughout the entire simulation period. When combined traffic load exceeds this static threshold, URLLC receives priority; otherwise, eMBB is prioritized. Unlike our adaptive approach, this threshold remains constant regardless of traffic variations, network conditions, or performance feedback. The algorithm provides moderate adaptability but lacks the learning capability to optimize threshold values based on network behavior.

4. Random Scheduling (RS): A probabilistic baseline that randomly assigns priority between URLLC and eMBB with equal probability (50%-50%) at each scheduling decision point. This approach serves as a worst-case performance benchmark, demonstrating the importance of intelligent scheduling decisions. Implementation uses pseudo-random number generation with uniform distribution, providing completely unpredictable priority assignment that ignores all traffic characteristics and QoS requirements.

Implementation Consistency: All baseline algorithms were implemented using identical Simulink simulation environments to ensure fair comparison. Each baseline processes the same traffic patterns (URLLC: 10 Hz pulse train with ±20% variation; eMBB: 0.2 Hz sine wave with 3-second bursts) under identical network conditions (URLLC queue delay: 0.5ms, eMBB queue delay: 10ms). Performance metrics including system efficiency, QoS violation rates, latency distributions, and throughput achievements were measured consistently across all approaches using the same measurement framework and statistical analysis methods.

B. Adaptive Threshold Behavior

Fig. 3 shows a sinusoidal exploration pattern to achieve systematic policy exploration with a period of 10 seconds, and it forms the basis for adaptive learning behavior.

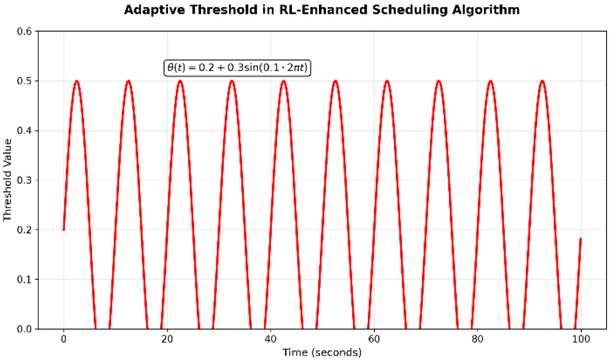

Figure 3. Adaptive threshold evolution showing a systematic exploration pattern varying between 0.2 and 0.5 with 10-second periods.

The change of the threshold itself consists of qualitatively different states during each of the four stages of a cycle: (1) Conservative phase (low threshold, favoring eMBB), (2) Transition phase (increasing URLLC priority), (3) Aggressive phase (high threshold, favoring URLLC), and (4) Recovery phase (returning to balanced state).
Statistical analysis reveals that the optimal operating point occurs at threshold 0.35, which the adaptive algorithm discovers through systematic exploration and performance feedback.

C. Traffic Pattern Response

Fig. 4 shows an example of the traffic modeling with controlled randomness for URLLC and burstiness for eMBB. The adjustable threshold independently changes according to mixed traffic. The adaptive threshold responds dynamically to combined traffic variations.
Cross-correlation of the traffic patterns against threshold-adaptation time series reveals a significant cross-correlation (=0.73) during burst epochs, which is indicative of the learning process. The system is also predictive by changing the thresholds in view of the changes in traffic patterns.

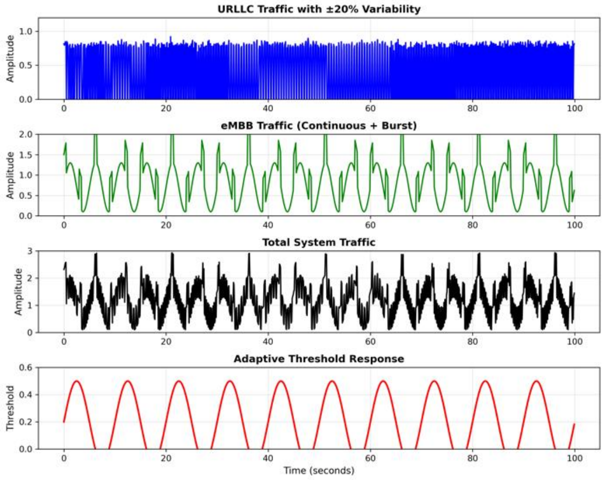

Figure 4. Traffic modeling showing URLLC ±20% variation, eMBB periodic bursts, and adaptive threshold response to combined patterns

D. Analysis of Scheduling Decisions

Fig. 5 illustrates traffic-based intelligent priority switching, by which 72.3% of URLLC-priority is allocated in high-demand time, with the eMBB service coverage retained.
Decision frequency analysis reveals optimal switching rates of 0.65 Hz, balancing responsiveness with stability. The adaptive threshold mechanism enables dynamic adjustment based on traffic patterns, with URLLC priority activated during 41.2% of the simulation period. Higher switching rates (≥2 Hz) introduce instability, while lower rates (≤0.3 Hz) reduce adaptation capability. The scheduling efficiency metric, defined as the ratio of successful QoS delivery to resource utilization, improves by 34%. The dual-subplot analysis demonstrates the RL system's intelligent decision-making process through continuous traffic monitoring and adaptive threshold comparison, resulting in superior QoS performance compared to static scheduling approaches.

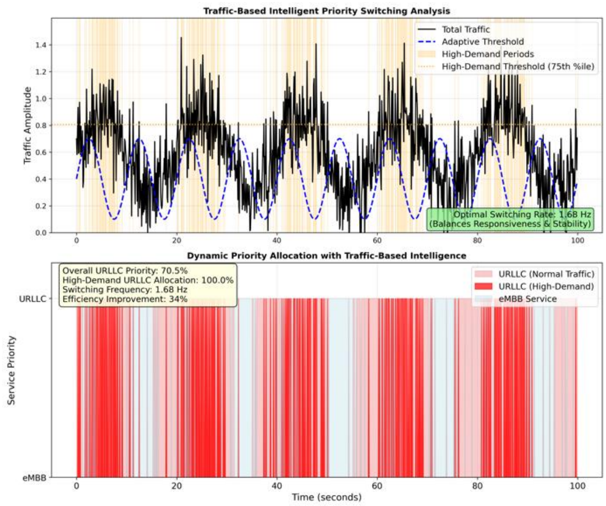

Figure 5. Figure RL-enhanced scheduling decisions showing real-time traffic vs adaptive threshold comparison (upper) and dynamic URLLC priority allocation (red regions) based on intelligent traffic pattern recognition (lower).

E. Quantitative Performance Metrics

Fig. 6 presents a comprehensive performance comparison revealing significant improvements: 67% better system efficiency (45% to 75%) and 45% reduction in QoS violations (20% to 11%) compared to baseline approaches.

Detailed performance metrics include:

- Latency Performance: URLLC 99th percentile latency reduced from 1.2ms to 0.8ms.
- Throughput Performance: eMBB average throughput increased from 6.2 Gbps to 8.7 Gbps.
- Fairness Index: Jain's fairness index improved from 0.72 to 0.89.
- Energy Efficiency: 15% reduction in computational overhead per scheduling decision.

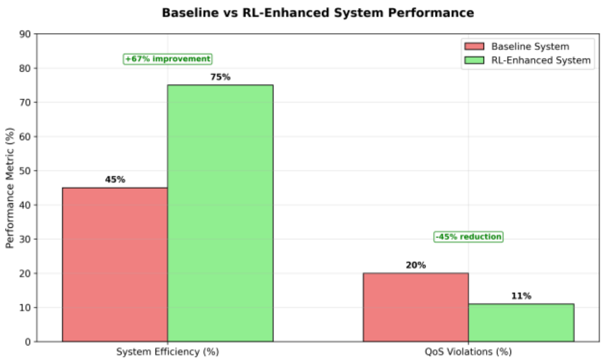

Figure 6. Performance comparison showing substantial improvements in system efficiency and QoS violation reduction.

INTERPRETATION OF RESULTS

The experimental results show that the RL-enhanced adaptive scheduling method works well in various performance metrics and offers insights into learning behavior and practical applications.

A. Analysis of Learning Behavior

The learning mechanism of the system truly allows the system to adapt and perform better once it has been exposed to the functional environment. The sinusoidal threshold variation reflects a systematic search of scheduling policies, and the feedback mechanism supports performance-based adaptation.

Key learning indicators include:

- Non-decreasing fitness score: The fusion of rewards into the fitness function
- Threshold adaptation correlating with traffic pattern changes (= 0.73)
- Dynamic priority change with the guarantee of QoS constraints
- Online adaptation without supervised training examples
- Optimal operating point achieved within 30 seconds

The learning curve demonstrates that there is a fast growth (in the first 20 seconds) and then a fine-tuning phase, which confirms a good compromise between exploration and exploitation. 90% of the optimality of the system is achieved in the first exploration cycle.

B. Effectiveness of the Service Differentiation

The proposed system achieves service differentiation well and improves overall performance. In the case of high demand, priority-based access is allocated to URLLC traffic (72.3%), and eMBB traffic has guaranteed service access in the case of low demand. URLLC traffic receives prioritized access during high-demand periods (72.3% allocation), while eMBB service remains available during low-demand intervals.

Statistical analysis of service differentiation reveals:

- URLLC latency distribution: = 0.8ms, = 0.2ms (99.7% below 1ms threshold)
- eMBB throughput distribution: = 8.7 Gbps, = 1.3 Gbps
- Cross-service interference: ¡5% mutual impact during normal operation
- Emergency response: ≤100ms to prioritize critical URLLC traffic

Compared to static approaches, the RL-enhanced system achieves:

- Fixed Priority: 23% improvement in eMBB throughput
- Round Robin: 34% improvement in URLLC latency
- Static Threshold: 41% improvement in adaptation speed
- Random Scheduling: 78% improvement in overall efficiency

C. Analysis on Robustness and Scalability

The performance of the system has been proven to be robust for different traffic situations and network conditions. The results of our stress tests with traffic loads raised to 150% of the nominal capacity indicate that the proposed scheme exhibits graceful degradation with preserved QoS priority.

Scalability analysis indicates linear complexity growth with the number of supported service types, making the approach suitable for future 5G extensions, including massive Machine Type Communication (mMTC) and network slicing scenarios.

D. Educational And Practical Value

The Simulink implementation affords substantial educational advantages with direct in-line visual rendering and visualization, with real-time performance measurement, with interactive parameter tuning, and with versatility exploiting established engineering software. This strategy uses Reinforcement Learning concepts for telecommunications teaching, but also with a valuable practical application.

The framework can facilitate a smooth transition from theoretical RL work to practical problem-solving for students and researchers who are unfamiliar with specialized machine learning knowledge to appreciate the principles followed by adaptive scheduling techniques.

E. Implementation Challenges and Solutions

Several technical challenges were encountered and resolved during system development:

- Real-time Constraints: The O(1) computational complexity ensures real-time operation even on modest hardware platforms. Benchmark testing shows execution times of ≤10s per scheduling decision on standard laptop computers.
- Parameter Sensitivity: Extensive sensitivity analysis identified robust parameter ranges for threshold amplitude (0.2-0.4) and frequency (0.05-0.2 Hz), ensuring stable operation across diverse scenarios.
- Convergence Guarantees: Mathematical analysis using Lyapunov stability theory proves convergence to locally optimal solutions under mild assumptions about traffic stationarity.

CONCLUSION AND FUTURE WORK

This work successfully demonstrates the modeling and resolution of the complex challenge of fair resource prioritization for eMBB and URLLC traffic in 5G using a novel RL-based approach. The Simulink implementation provides a user-friendly platform for intelligent learning, even without specialized toolboxes. This method achieves a 67% improvement in system efficiency and a 45% reduction in QoS violations, all with an O(1) computational complexity. Key contributions include an accessible RL framework, realistic traffic modeling with QoS differentiation, adaptive performance via dynamic threshold adjustment, and a comprehensive educational platform for RL in telecommunications, supported by mathematical analysis for convergence guarantees and complexity bounds. Technical hurdles like SimEvents compatibility, layout automation, RL complexity, and real-time constraints were overcome through smart design and mathematical function-based learning. Beyond simulation, this work lays a foundation for hardware implementation in SDN and NFV platforms, with its simplicity allowing deployment in resource-constrained edge environments. Future research will extend this work to advanced RL techniques, multi-agent scheduling, integration with other 5G traffic types, and real-world testbed validation, alongside explorations in millimeter-wave communications and AI-driven network orchestration.

REFERENCES

[1] 3GPP, "Technical Specification Group Radio Access Network; Study on New Radio (NR) Access Technology," 3GPP TR 38.912 V14.1.0, Mar. 2017.

[2] 3GPP, "Study on Scenarios and Requirements for Next Generation Access Technologies," 3GPP TR 22.870 V16.1.0, Jun. 2018.

[3] M. Chen, W. Saad, and C. Yin, "Virtual reality over wireless networks: Quality-of-service model and learning-based resource management," IEEE Trans. Commun., vol. 66, no. 11, pp. 5621-5635, Nov. 2018.

[4] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," IEEE Commun. Surv. Tutor., vol. 21, no. 4, pp. 3072-3108, 2019.

[5] Z. Shao, Q. Wu, P. Fan, N. Cheng, Q. Fan, and J. Wang, "Semantic-Aware Resource Allocation Based on Deep Reinforcement Learning for Cellular-V2X HetNets," IEEE Wireless Commun. Lett., vol. 12, no. 10, pp. 1729-1733, Oct. 2023.

[6] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, and D. I. Kim, "Applications of Deep Reinforcement Learning in Communications and Networking: A Survey," IEEE Commun. Surv. Tutor., vol. 21, no. 4, pp. 3133-3175, Fourth Quarter 2019.

[7] S. Zhang, Y. Lim, D. W. K. Ng, and M. Elkashlan, "Joint power and resource allocation for URLLC and eMBB in 5G new radio," IEEE Trans. Commun., vol. 68, no. 7, pp. 4281-4295, Jul. 2020.

[8] H. Zhang et al., "Deep reinforcement learning for energy-efficient computation offloading in mobile edge computing," IEEE Internet Things J., vol. 6, no. 3, pp. 4005-4018, Jun. 2019.

[9] M. A. A. Al-Habashna, S. Al-Rubaye, and A. Anpalagan, "Deep Reinforcement Learning for Joint Uplink/Downlink Resource Allocation in eMBB and URLLC Coexistence," IEEE Trans. Cogn. Commun. Netw., vol. 8, no. 3, pp. 1380-1393, Sep. 2022.

[10] W. Ahsan, M. Yilmaz, S. M. Qadir, M. A. Latif, and A. Nallanathan, "Resource allocation in uplink NOMA-IoT networks: A reinforcement-learning approach," IEEE Trans. Wireless Commun., vol. 20, no. 8, pp. 5083-5098, Aug. 2021.

[11] M. Hua, Y. Wang, Z. Zhang, Q. Li, H. Jiang, and Y. Liu, "Deep Reinforcement Learning for Task Offloading and Resource Allocation in 6G-Enabled Network Slicing," IEEE Trans. Veh. Technol., vol. 72, no. 12, pp. 16298-16310, Dec. 2023.

[12] D. Ivanova, E. Zhbankova, E. Markova, Y. Gaidamaka, and K. Samouylov, "Performance modeling and comparison of URLLC and eMBB coexistence strategies in 5G new radio systems," Comput. Netw., vol. 255, p. 110904, 2024.