ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

AI-ENHANCED
DYNAMIC PREEMPTIVE RESOURCE ALLOCATION
IN NEXT GENERATION CELLULAR NETWORKS

Ph.D. THESIS

Ege ENGIN

Electronics and Communications Engineering Department

Electronics Engineering Programme

JUNE 2025

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

AI-ENHANCED
DYNAMIC PREEMPTIVE RESOURCE ALLOCATION
IN NEXT GENERATION CELLULAR NETWORKS

Ph.D. THESIS

Ege ENGIN
(504182214)

Electronics and Communications Engineering Department

Electronics Engineering Programme

Thesis Advisor: Prof. Dr. Hakan Ali ÇIRPAN

JUNE 2025

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YENİ NESİL HÜCRESEL AĞLARDA YAPAY ZEKA DESTEKLİ DİNAMİK ÖNCELİKLİ KAYNAK TAHSİSİ

DOKTORA TEZİ

Ege ENGIN
(504182214)

Elektronik ve Haberleşme Mühendisliği

Elektronik Mühendisliği Programı

Tez Danışmanı: Prof. Dr. Hakan Ali ÇIRPAN

HAZİRAN 2025

Ege ENGIN, a Ph.D. student of ITU Graduate School student ID 504182214 successfully defended the thesis entitled “AI-ENHANCED DYNAMIC PREEMPTIVE RESOURCE ALLOCATION IN NEXT GENERATION CELLULAR NETWORKS”, which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Dr. Hakan Ali ÇIRPAN** Istanbul Technical University	....
Jury Members :	**Prof. Dr. Ender Mete EKŞİOĞLU** Istanbul Technical University	....
	Assis. Prof. Dr. Semiha Tedik BAŞARAN Istanbul Technical University	....
	Dr. İbrahim HÖKELEK TÜBİTAK BİLGEM	....
	Dr. Ömer Faruk GEMİCİ Ericsson	....

Date of Submission : 2 May 2025
Date of Defense : 19 June 2025

FOREWORD

This thesis is the result of a long and demanding journey, and I would like to express my heartfelt gratitude to those who supported, guided, and encouraged me throughout the process.

First and foremost, I extend my deepest appreciation to my thesis advisor, Prof. Dr. Hakan Ali ÇIRPAN, whose wisdom, vast knowledge, and practical approach have been invaluable. His patience, understanding of my full-time work commitments, and unwavering trust in my abilities gave me the confidence to persist and succeed in this endeavor. His guidance has truly been a cornerstone of my academic achievements. I am also grateful to Dr. İbrahim HÖKELEK, with whom I had the privilege of publishing three research papers. His ability to listen carefully, provide sharp insights, and offer critical suggestions in a timely manner played a key role in enhancing the quality of my work. His support and encouragement were invaluable throughout my thesis journey.

To my wife, Melis, my greatest companion and constant source of strength: thank you for your love, patience, and endless support. You have been my anchor in every challenge and my greatest joy in every achievement. To my son, Deniz, who brings immense happiness and love to my life: I am deeply thankful for your patience and understanding during the times I had to focus on my research. You make every moment of life brighter and more meaningful. I also owe sincere thanks to my parents, Ahmet and Serpil, for their unconditional love, support, and encouragement throughout my life. Your belief in me has always been a source of motivation and strength, and I am deeply grateful for everything you have done.

I would like to express my gratitude to ASELSAN, where I had the opportunity to attend insightful lectures during my work hours, gaining valuable knowledge that contributed significantly to my research. I am also thankful to my supervisors and colleagues at Aselsan, who supported me academically and professionally, and to my friends there, who helped ease my stress during challenging times. Finally, I extend my thanks to my current colleagues at E-Space and my former colleagues at Ericsson for their constructive reviews of my publications, thought-provoking technical discussions, and shared learning experiences, all of which enriched my perspective and contributed to my growth.

To everyone who played a part in this journey, offering guidance, support, or encouragement: thank you. This thesis would not have been possible without you.

June 2025 Ege ENGIN

TABLE OF CONTENTS

ABBREVIATIONS

3GPP	: 3rd Generation Partnership Project
5G	: Fifth-Generation
5G NR	: Fifth-Generation New-Radio
5GB	: 5G And Beyond
6G	: Sixth-Generation
ACK	: Acknowledgment
AI	: Artificial Intelligence
CQI	: Channel Quality Indicator
CSI	: Channel State Information
CSI-RS	: Channel State Information Reference Signal
DCI	: Downlink Control Information
DQN	: Deep Q Network
eMBB	: Enhanced Mobile Broadband
eNodeB	: Evolved Node B
gNodeB	: Next Generation Node B
RAN	: Radio Access Network
RL	: Reinforcement Learning
SINR	: Signal-to-Interference-and-Noise Ratio
URLLC	: Ultra-Reliable Low-Latency Communication
V2V	: Vehicle-to-Vehicle
V2X	: Vehicle-to-Everything
QoS	: Quality-of-Service
FIFO	: First-In-First-Out
MCS	: Modulation and Coding Scheme
mMTC	: massive machine-type communication

SYMBOLS

$\mathscr{R}_i^{\text{eMBB}}$	: Throughput for eMBB user i.
$P_{\text{success},i}(a)$	: Probability of successful decoding for user i under preemption action a.
$N_{\text{RB},i}$	: Number of resource blocks allocated to user i.
N_{sc}	: Number of subcarriers per resource block.
$N_{\text{sym},i}$	: Number of OFDM symbols allocated to user i.
$Q_{m,i}$	: Modulation order for user i.
C_i	: Code rate for user i.
T_{slot}	: Slot duration.
OH_i	: Overhead factor for user i.
$\tau_{\text{max}}^{\text{URLLC}}$	: Maximum permissible end-to-end delay for URLLC traffic.
$P_{\text{success}}^{\text{URLLC}}$	: Minimum required probability of successful packet delivery for URLLC traffic.
$\tau^{\text{URLLC}}(a)$	: End-to-end delay for URLLC traffic under preemption action a.
$\mathbb{P}^{\text{URLLC}}(\text{success} \mid a)$	: Probability of successful URLLC transmission scheduled based on preemption action a.
p	: Preemption pattern in the time domain (one of 10 standardized symbol patterns).
d	: Resource block density for preemption in the frequency domain $\{2,5,10\}$.
$a = ((p,b))$	: Preemption action.
$\mathscr{A}$	: Set of all feasible preemption actions.
$\mathscr{P}^{\text{URLLC}}$	: Set of all feasible preemption patterns.
N_{RB}	: Total number of resource blocks in a system bandwidth.
$(s_1, s_2, \ldots, s_n)$	: Indices of OFDM symbols within a slot forming a preemption pattern.
$n \in \{2,4,7\}$	: Possible number of OFDM symbols for a mini-slot.
$\mathscr{D}^{\text{URLLC}}$	: Set of possible resource block sets for preemption.
$(d_1, d_2, \ldots, d_k)$	: Indices of resource blocks selected for preemption.

LIST OF TABLES

LIST OF FIGURES

Page

AI-ENHANCED
DYNAMIC PREEMPTIVE RESOURCE ALLOCATION
IN NEXT GENERATION CELLULAR NETWORKS

SUMMARY

The rapid expansion of diverse slices and applications in 5G and beyond (5GB) communication systems has necessitated the fulfillment of stringent key performance indicators (KPIs), including higher data rates, reduced latency, improved spectral efficiency, enhanced energy efficiency, and expanded network capacity. Each service category is defined to satisfy one or more different KPIs. These specific needs are addressed with network slices, which divide network resources into dedicated slices. Likewise, Radio Access Network (RAN) slicing plays a crucial role in managing resources at the radio access level, enabling the coexistence and efficient sharing of resources across multiple slices. The individual needs and demands of the two primary slices, namely, Enhanced Mobile Broadband (eMBB) and Ultra-Reliable Low-Latency Communication (URLLC), create a complex optimization problem where traditional scheduling approaches often fail to achieve satisfactory performance for both traffic types simultaneously.

In response to this challenge, this thesis adopts a systematic, multi-phase approach to enhance scheduling performance incrementally. Initially, a heuristic scheduling approach is proposed to establish a baseline and facilitate easy integration with existing systems. This approach utilizes the allowed latency of URLLC slices to optimize concurrent scheduling by mitigating the adverse impact of preemption on the perceived quality of the eMBB slices. The Patient and Intelligent Preemption (PIPE) algorithm demonstrates how strategic delaying of URLLC scheduling decisions can minimize unnecessary preemption of eMBB resources while still meeting strict URLLC deadlines. The algorithm introduces a novel preemption allowance rate that defines how many times an eMBB codeblock can be preempted before requiring retransmission, effectively creating a buffer against throughput degradation. Evaluation through link-level simulations shows PIPE's ability to maintain eMBB throughput within 5% degradation while fully satisfying URLLC latency constraints, outperforming traditional scheduling methods like aggressive and threshold-proportional approaches.

Building upon the heuristic foundation, the research progresses to develop a 3GPP-compliant preemption mechanism that incorporates all aspects of the radio access network. This enhanced approach introduces selective preemption targeting low-priority eMBB resources, such as those using lower modulation orders or having higher error-correction redundancy. The mechanism includes comprehensive modeling of preemption indication signaling, which enables efficient buffer management at user equipment by providing timely notifications about punctured resources. The system-level implementation spans the entire protocol stack from physical layer

processing to MAC scheduling and HARQ operations, ensuring realistic performance evaluation under diverse network conditions. Simulation results demonstrate the mechanism's ability to maintain URLLC latency below 1ms while minimizing eMBB throughput degradation, even under high URLLC traffic loads. The solution maintains low computational complexity and requires minimal control channel overhead, making it practical for real-world deployments in dense network scenarios.

The final phase of the research introduces an artificial intelligence-driven approach using Deep Reinforcement Learning (DRL) to further optimize the preemption decisions. The Context-Aware Multi-DQN Scheduling (CA-M-DQN) framework represents a significant advancement over static scheduling algorithms by employing multiple Deep Q-Network models, each specialized for different Channel Quality Indicator (CQI) levels and URLLC packet sizes. This modular design allows the system to adapt dynamically to varying network conditions and traffic patterns. The framework incorporates a carefully designed hybrid action space with 28 discrete preemption strategies that comply with 3GPP mini-slot configurations, along with a multi-objective reward function that balances URLLC latency compliance against eMBB throughput preservation. System-level validation shows that CA-M-DQN achieves superior eMBB throughput compared to heuristic baselines while maintaining URLLC reliability with block error rates below 1%, meeting the most stringent 3GPP requirements. The success of this approach lies in its ability to learn optimal policies through interactions with the network environment, continuously improving its decision-making capabilities without requiring explicit mathematical models of the underlying system dynamics.

The practical implications of this research extend across various deployment scenarios in 5GB networks. The proposed solutions demonstrate scalability for high-density urban environments and industrial IoT applications where diverse traffic types must coexist. Energy efficiency improvements are achieved through reduced retransmissions and optimized resource allocation, contributing to more sustainable network operations. The strict adherence to 3GPP standards ensures compatibility with existing infrastructure, lowering barriers to adoption for network operators. Furthermore, the frameworks' modular designs provide flexibility to accommodate future network enhancements and new service categories beyond eMBB and URLLC, such as massive Machine-Type Communications (mMTC) and extreme URLLC (xURLLC).

Looking ahead, several promising directions emerge for extending this research. Multi-Agent Reinforcement Learning (MARL) could address the challenges of inter-cell interference in multi-cell deployments by enabling coordinated resource allocation across multiple base stations. Hardware validation through software-defined radio (SDR) testbeds would provide valuable insights into real-world performance characteristics and implementation challenges. The integration of predictive capabilities using advanced machine learning techniques could further enhance the system's ability to anticipate traffic patterns and optimize resource allocation proactively. Additionally, the framework could be extended to support emerging use cases in 6G networks, such as holographic communications and digital twins, where even more demanding latency and reliability requirements are expected.

This thesis makes significant contributions to the field of wireless communications by developing practical solutions to the complex problem of resource allocation in heterogeneous 5GB networks. The systematic progression from heuristic methods to AI-driven approaches demonstrates how incremental enhancements can lead to substantial performance improvements. By combining theoretical rigor with practical implementation considerations, the research bridges the gap between academic concepts and real-world deployment requirements. The proposed frameworks establish a solid foundation for future innovations in intelligent radio resource management, paving the way for self-optimizing networks that can adapt to the ever-evolving demands of next-generation wireless applications and services. The integration of multi-layer preemption mechanisms with advanced reinforcement learning techniques represents a paradigm shift in how network slicing can be managed, offering unprecedented flexibility and efficiency in resource utilization while maintaining strict quality-of-service guarantees for diverse traffic types.

**YENİ NESİL HÜCRESEL AĞLARDA
YAPAY ZEKA DESTEKLİ
DİNAMİK ÖNCELİKLİ KAYNAK TAHSİSİ**

ÖZET

5G ve ötesi (5GB) haberleşme sistemlerinde çeşitlenen dilimlerin ve uygulamaların hızla yaygınlaşması, daha yüksek veri hızları, düşük gecikme, geliştirilmiş spektral verimlilik, artırılmış enerji verimliliği ve genişletilmiş ağ kapasitesi gibi zorlu temel performans göstergelerinin (KPI'lar) sağlanmasını zorunlu kılmıştır. Her hizmet kategorisi, bir veya daha fazla KPI'ı karşılayacak şekilde tanımlanmıştır. Bu özel gereksinimler, ağ kaynaklarının özel dilimlere bölünmesiyle sağlanan ağ dilimleriyle karşılanmaktadır. Benzer şekilde, Radyo Erişim Ağı (RAN) dilimleme, radyo erişim seviyesindeki kaynakların yönetilmesinde kilit rol oynar; birden fazla dilimin aynı anda var olmasına ve kaynakların verimli şekilde paylaşılmasına olanak tanır. Geliştirilmiş Mobil Genişbant (eMBB) ve Ultra-Güvenilir Düşük Gecikmeli Haberleşme (URLLC) olmak üzere iki ana dilimin farklı ihtiyaçları, geleneksel zamanlayıcıların her iki trafik türü için aynı anda tatmin edici performans sağlayamadığı karmaşık bir optimizasyon problemi ortaya çıkarır.

Bu zorluğa yanıt olarak, bu tez zamanlayıcı performansını kademeli olarak iyileştirmeyi amaçlayan sistematik ve çok aşamalı bir yaklaşım benimsemektedir. İlk olarak, mevcut sistemlerle kolay entegrasyonu sağlamak ve bir karşılaştırma temeli oluşturmak amacıyla sezgisel (heuristic) bir zamanlama yöntemi önerilmektedir. Bu yöntem, URLLC dilimlerinin izin verilen gecikmesini kullanarak eşzamanlı zamanlamayı optimize eder ve eMBB dilimlerinin algılanan kalitesi üzerindeki önceliklendirme (preemption) etkisini azaltır. "Sabırlı ve Akıllı Önceliklendirme" (PIPE) algoritması, URLLC zamanlama kararlarının stratejik olarak ertelenmesinin, eMBB kaynaklarının gereksiz önceliklendirmesini en aza indirirken URLLC gecikme gerekliliklerini de karşılayabileceğini göstermektedir. Algoritma, bir eMBB kod bloğunun yeniden iletim gerektirmeden kaç kez önceliklendirilebileceğini tanımlayan yenilikçi bir "önceliklendirme toleransı oranı" sunar ve bu sayede veri aktarım oranındaki düşüşlere karşı bir tampon oluşturur. Bağlantı seviyesi simülasyonları, PIPE algoritmasının eMBB aktarım hızında yalnızca %5'ten az bir kayıpla URLLC gecikme kısıtlarını tam olarak karşılayabildiğini ve agresif ya da eşik oranlı geleneksel zamanlama yöntemlerinden daha iyi performans gösterdiğini ortaya koymaktadır.

Sezgisel temelin üzerine inşa edilen araştırma, radyo erişim ağının tüm bileşenlerini kapsayan 3GPP uyumlu bir önceliklendirme mekanizması geliştirmeye doğru ilerlemektedir. Bu gelişmiş yaklaşım, düşük modülasyon derecelerine sahip veya yüksek hata düzeltme yedekliliği içeren düşük öncelikli eMBB kaynaklarına odaklanan seçici önceliklendirmeyi tanıtır. Mekanizma, kullanıcı ekipmanında (UE) zamanında bilgi sağlanarak tampon yönetimini kolaylaştıran önceliklendirme gösterimi sinyalleşmesinin kapsamlı modellenmesini içermektedir. Sistem düzeyinde

uygulama, fiziksel katman işleminden MAC zamanlamasına ve HARQ işlemlerine kadar tüm protokol yığını boyunca gerçekleştirilir ve çeşitli ağ koşulları altında gerçekçi bir performans değerlendirmesi yapılmasını sağlar. Simülasyon sonuçları, yüksek URLLC trafik yüklerinde bile URLLC gecikmesinin 1 ms altında tutulduğunu ve eMBB aktarım hızındaki kaybın en aza indirildiğini göstermektedir. Bu çözüm, düşük hesaplama karmaşıklığına ve minimum kontrol kanalı yüküne sahiptir; bu da onu yoğun ağ senaryolarında gerçek dünya uygulamaları için pratik kılar.

Araştırmanın son aşaması, önceliklendirme kararlarını daha da optimize etmek amacıyla Yapay Zeka tabanlı bir yaklaşımı, Derin Pekiştirmeli Öğrenme (Deep Reinforcement Learning – DRL) kullanarak sunar. Bağlam Farkındalıklı Çoklu DQN Zamanlayıcı (CA-M-DQN) çerçevesi, her biri farklı Kanal Kalite Göstergesi (CQI) seviyeleri ve URLLC paket boyutları için uzmanlaşmış birden fazla Derin Q-Ağ modelini kullanarak, statik zamanlayıcılara kıyasla önemli bir ilerleme sağlar. Bu modüler tasarım, sistemin değişken ağ koşulları ve trafik desenlerine dinamik olarak uyum sağlamasına olanak tanır. Çerçeve, 3GPP mini-yuvalı yapılandırmalarına uygun 28 ayrık önceliklendirme stratejisini içeren dikkatle tasarlanmış hibrit bir eylem uzayı ve URLLC gecikme uygunluğu ile eMBB veri hızı koruması arasında denge kuran çok amaçlı bir ödül fonksiyonu içermektedir. Sistem düzeyinde yapılan doğrulamalar, CA-M-DQN'in sezgisel yaklaşımlara kıyasla daha yüksek eMBB veri hızı sağladığını ve aynı zamanda %1'in altında blok hata oranları ile URLLC güvenilirliğini koruduğunu göstermektedir; bu da 3GPP'nin en katı gereksinimlerini karşılamaktadır. Bu yaklaşımın başarısı, sistemin matematiksel modellerine ihtiyaç duymadan ağ ortamıyla etkileşimler yoluyla en iyi politika kararlarını öğrenme yeteneğinde yatmaktadır.

Bu araştırmanın pratik etkileri, 5GB ağlarında çeşitli kurulum senaryoları boyunca uzanmaktadır. Önerilen çözümler, yüksek yoğunluklu kentsel ortamlar ve farklı trafik türlerinin bir arada bulunduğu endüstriyel IoT uygulamaları için ölçeklenebilirlik sunmaktadır. Daha az yeniden iletim ve optimize edilmiş kaynak tahsisi sayesinde enerji verimliliğinde iyileşmeler sağlanmakta ve bu da daha sürdürülebilir ağ operasyonlarına katkıda bulunmaktadır. 3GPP standartlarına sıkı sıkıya bağlı kalınması, mevcut altyapıyla uyumluluğu garanti ederek ağ operatörleri için benimseme engellerini azaltmaktadır. Ayrıca, çerçevelerin modüler tasarımları, eMBB ve URLLC ötesindeki yeni hizmet kategorileri – örneğin büyük ölçekli Makine Türü Haberleşme (mMTC) ve aşırı güvenilir düşük gecikmeli iletişim (xURLLC) – için esneklik sağlamaktadır.

Geleceğe yönelik olarak, bu araştırmanın genişletilebileceği birkaç umut verici yön ortaya çıkmaktadır. Çok Etmenli Pekiştirmeli Öğrenme (Multi-Agent Reinforcement Learning – MARL), çok hücreli yerleşimlerde hücreler arası girişim sorunlarını çözmek için kullanılabilir ve birden fazla baz istasyonu arasında koordine kaynak tahsisine olanak tanıyabilir. Yazılım tanımlı radyo (SDR) test yatakları aracılığıyla yapılacak donanım doğrulamaları, gerçek dünya performansı ve uygulama zorlukları hakkında değerli bilgiler sunacaktır. Gelişmiş makine öğrenimi teknikleriyle öngörü yeteneklerinin entegrasyonu, sistemin trafik desenlerini önceden tahmin ederek kaynak tahsisini proaktif olarak optimize etmesini daha da geliştirebilir. Ek olarak, çerçeve, holografik iletişim ve dijital ikizler gibi daha zorlu gecikme ve güvenilirlik gereksinimlerinin beklendiği 6G ağlarındaki yeni kullanım senaryolarını destekleyecek şekilde genişletilebilir.

Bu tez, heterojen 5GB ağlarında kaynak tahsisi gibi karmaşık bir probleme pratik çözümler geliştirerek kablosuz iletişim alanına önemli katkılarda bulunmaktadır. Sezgisel yöntemlerden yapay zeka destekli yaklaşımlara doğru sistematik ilerleme, kademeli iyileştirmelerin nasıl önemli performans artışlarına yol açabileceğini göstermektedir. Teorik titizlik ile pratik uygulama ihtiyaçlarının birleşimi, akademik kavramlar ile gerçek dünya kurulum gereksinimleri arasındaki boşluğu kapatmaktadır. Önerilen çerçeveler, akıllı radyo kaynak yönetiminde gelecekteki yenilikler için sağlam bir temel oluşturmakta ve kendini optimize edebilen, gelecek nesil kablosuz uygulama ve hizmetlerin sürekli değişen taleplerine uyum sağlayabilen ağların önünü açmaktadır. Çok katmanlı önceliklendirme mekanizmalarının gelişmiş pekiştirmeli öğrenme teknikleriyle entegrasyonu, ağ dilimlemenin yönetilme biçiminde bir paradigma değişimini temsil etmekte ve kaynak kullanımında benzeri görülmemiş bir esneklik ve verimlilik sunarken farklı trafik türleri için katı hizmet kalitesi garantilerinin korunmasını sağlamaktadır.

1. INTRODUCTION

The Third Generation Partnership Project (3GPP) standardizes the Fifth Generation New Radio (5G NR), a significant advancement in mobile technology. The new system is designed to meet the growing demand for higher data rates, reduced latency, and more efficient utilization of wireless spectrum. 5G NR introduces network slicing, enabling multiple virtual networks to be established on a single physical network infrastructure. They facilitated different applications with different requirements to be supported simultaneously through separate slices. The primary service categories of 5G can be listed as follows, which can be executed in separate slices: Enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communications (URLLC), and Massive Machine-Type Communications (mMTC). eMBB offers high-speed internet for use in applications such as video streaming and downloading large files. URLLC facilitates time-sensitive tasks, such as remote surgery or automated factories, which typically require precise and low-latency transmissions. mMTC connects large numbers of intelligent devices for services such as smart cities and nature monitoring. However, running these different services simultaneously presents technical issues. While eMBB requires large bandwidth for high-speed downloads, URLLC necessitates guaranteed low latency and resiliency. These two conflicting needs complicate the allocation of network resources because it is possible to optimize one slice at the expense of decreasing the perceived quality of other slices. Researchers are developing intelligent scheduling tools to effectively trade these needs, because traditional scheduling tools are highly challenged when employed in 5G NR networks. The flexibility of these algorithms to various layers of the 5G protocol stack and different network conditions can be considered a second challenge. Apart from this, dynamic channel conditions across varying traffic loads vying for shared radio resources are an extra consideration. It poses an additional requirement for adaptively scheduling solutions that are contingent on real-time dynamic resource assignment.

When eMBB and URLLC slices share the same resource, their scheduling becomes complex using legacy scheduling methods because they have conflicting throughput and latency requirements. 3GPP introduced the preemption mechanism as a solution to this issue in the 5G NR specifications [4]. Higher-priority slices, such as URLLC, are allowed to preempt allocations of ongoing lower-priority slices, such as eMBB, in a standardized preemption mechanism. It supports the ultra-low-latency requirement of URLLC slices without prior reservation of such resources. It provides dynamic and on-demand resource allocation. It is beneficial where URLLC traffic occurs in bursts and should be served with low latency. In this thesis, two heuristic algorithms are proposed that utilize the preemption mechanism to optimize the throughput performance of eMBB slices while maintaining the latency constraint of URLLC slices for 5G and beyond (5GB) networks. The first algorithm, presented in Chapter 3, is simulated using a link-level simulator to examine its performance under simulated physical-layer conditions. The extended version of this algorithm considers Channel Quality Indicator (CQI) feedback information when making preemption decisions, as presented in Chapter 4. This improved version is, moreover, simulated using a 3GPP-based system-level simulator to examine its end-to-end performance and applicability under actual network conditions. To apply these end-to-end investigations, the preemption mechanism, encompassing all physical layer dynamics, i.e., transmitter-side preemption, receiver-side soft buffer flushing, and associated signaling procedures, such as Downlink Control Information (DCI) and the Preemption Indication (PI), has been integrated into the MATLAB-based 3GPP system-level simulator.

The latest advancement in Artificial Intelligence (AI), specifically Deep Reinforcement Learning (DRL), effectively addresses complex optimization problems, including scheduling and resource allocation. DRL enables schedulers to learn optimal actions through interactions with the environment without requiring an explicit model of network dynamics. The learning capability makes DRL eligible to address the dynamic, multi-objective nature of 5GB resource allocation. In this thesis, we propose an AI-based framework for dynamically making preemption decisions in 5GB networks. This framework addresses the coexistence of eMBB and

URLLC slices, encompassing traffic awareness, buffer state, and channel conditions in making DRL-based preemption decisions. Simulation results show that the proposed scheduling framework effectively maintains the stringent latency demands of URLLC slices while incurring only an acceptable decrease in eMBB throughput, demonstrating its viability for practical implementation. This framework combines AI-driven preemptive scheduling with system-level modeling of 5G NR Radio Access Network (RAN) to create a robust and scalable approach to intelligent radio resource management. The proposed two heuristic and one AI-based methods can serve as alternative approaches to resource allocation for coexisting slices with different requirements in the next generation of self-optimizing wireless networks, where adaptability will be essential.

Collectively, the central research question is tried to be answered throughout this thesis:

To what extent can an AI-driven scheduling framework, leveraging both heuristic and DRL-based strategies, optimize radio resource allocation in 5G NR systems to concurrently maximize throughput, minimize latency, and ensure high reliability while also providing scalability for future network generations?

1.1 Thesis Objectives and Contributions

This research aims to develop sophisticated DRL strategies for optimal resource allocation in 5G NR networks. The primary objectives and contributions are summarized below.

- Implementation of an optimized preemptive scheduling strategy that effectively balances URLLC latency constraints with eMBB throughput demands, using intelligent preemption based on traffic urgency and real-time resource utilization.

- Introduction of an innovative AI-driven scheduling framework leveraging DRL for adaptive network resource management, capable of dynamically balancing conflicting QoS requirements across heterogeneous traffic types.

- Development of an enhanced DQN model that integrates context-awareness and real-time feedback mechanisms to improve performance in dynamic

5G environments, featuring novel state representation techniques, attention mechanisms, and multi-objective optimization capabilities.

- Design of a novel multi-dimensional reward function tailored for multi-objective optimization in radio resource allocation, which adaptively adjusts performance metric weights based on current network conditions and service priorities.

- Extensive experimental validation demonstrating the quantitative superiority of the proposed AI-based approach over conventional scheduling algorithms across a range of performance metrics and network scenarios.

- Provision of practical implementation guidelines for integrating AI-based scheduling solutions into existing 5G infrastructure, addressing challenges related to computational complexity, real-time processing, and deployment feasibility.

1.2 Organization of the Thesis

The thesis is organized as follows:

- Chapter 2 provides a review of resource allocation in 5G NR. It covers the details of 5G NR scheduling, preemption mechanism offered by 3GPP, and scheduling mechanisms. It also outlines the fundamentals of AI methods and their application to wireless communications.

- Chapter 3 presents a critical analysis of the current literature on AI-driven network optimization, with a focus on reinforcement learning approaches to resource management in 5G networks. This chapter highlights the strengths and limitations of existing methodologies and identifies key research gaps.

- Chapter 4 details the proposed DRL-based scheduling framework, including the system model, methodological approach, and design considerations. It covers the mathematical formulation of the reinforcement learning problem, the architecture of the DQN model, and the development of the corresponding state representations and multi-dimensional reward functions.

- Chapter 5 analyzes the simulation results and performance evaluations, comparing the proposed AI-based scheduling algorithm with conventional techniques. Detailed discussions on throughput distributions, latency profiles, reliability coefficients, and resource utilization are provided.

- Chapter 6 introduces an advanced Context-Aware Multi-DQN-Driven Preemptive Scheduling paradigm for network slicing, presenting system-level evaluations and discussing real-world deployment implications. This chapter examines the framework's potential to support efficient network slicing and service differentiation in 5G and beyond.

- Chapter 7 synthesizes the key findings of the research, discusses practical implications for network operators and equipment vendors, and outlines potential directions for future research in AI-enhanced wireless communications.

This thesis makes a significant contribution to AI-powered wireless networks by demonstrating how deep reinforcement learning can transform resource allocation strategies in 5G and beyond. Integrating advanced AI methodologies with state-of-the-art scheduling techniques lays the groundwork for the next generation of intelligent, self-optimizing cellular networks capable of meeting the evolving demands of future wireless applications and services.

2. BACKGROUND

The rapid expansion of diverse slices and applications in 5G and beyond (5GB) communication systems has necessitated the fulfillment of stringent key performance indicators (KPIs), including higher data rates, reduced latency, improved spectral efficiency, enhanced energy efficiency, and expanded network capacity [5]. Each service category is defined to satisfy one or more different KPIs. These specific needs are addressed with network slices, which divide network resources into dedicated slices. Likewise, Radio Access Network (RAN) slicing plays a crucial role in managing resources at the radio access level, enabling the coexistence and efficient sharing of resources across multiple slices. The individual needs and demands of the two primary slices, namely, Enhanced Mobile Broadband (eMBB) and Ultra-Reliable Low-Latency Communication (URLLC), are illustrated in Fig. 2.1. eMBB, being one among the three essential 5G service categories, is designed to deliver enhanced mobile broadband slices with the capability of handling applications such as video calls, ultra-high-definition (UHD) video streaming, mobile cloud computing, and fixed wireless access [6,7]. This high usage leads to an exponential increase in data traffic, and it needs to be effectively addressed by eMBB slices to provide substantially higher data rates [8]. On the other hand, URLLC slices are tailored for mission-critical applications, such as remote surgery, public safety infrastructure, industrial automation, and automotive communications, for which ultra-high reliability and low latency are of critical concern [9].

URLLC users are commonly characterized as generating bursty traffic with short packet lengths, calling for rapid scheduling decisions. They often need to be taken within a single Transmission Time Interval (TTI) of 1 ms or even shorter intervals, e.g., two to four symbols per mini-slot [1]. Rapid scheduling is required to meet URLLC's stringent latency requirements, which involve end-to-end delays of less than 1 ms. In

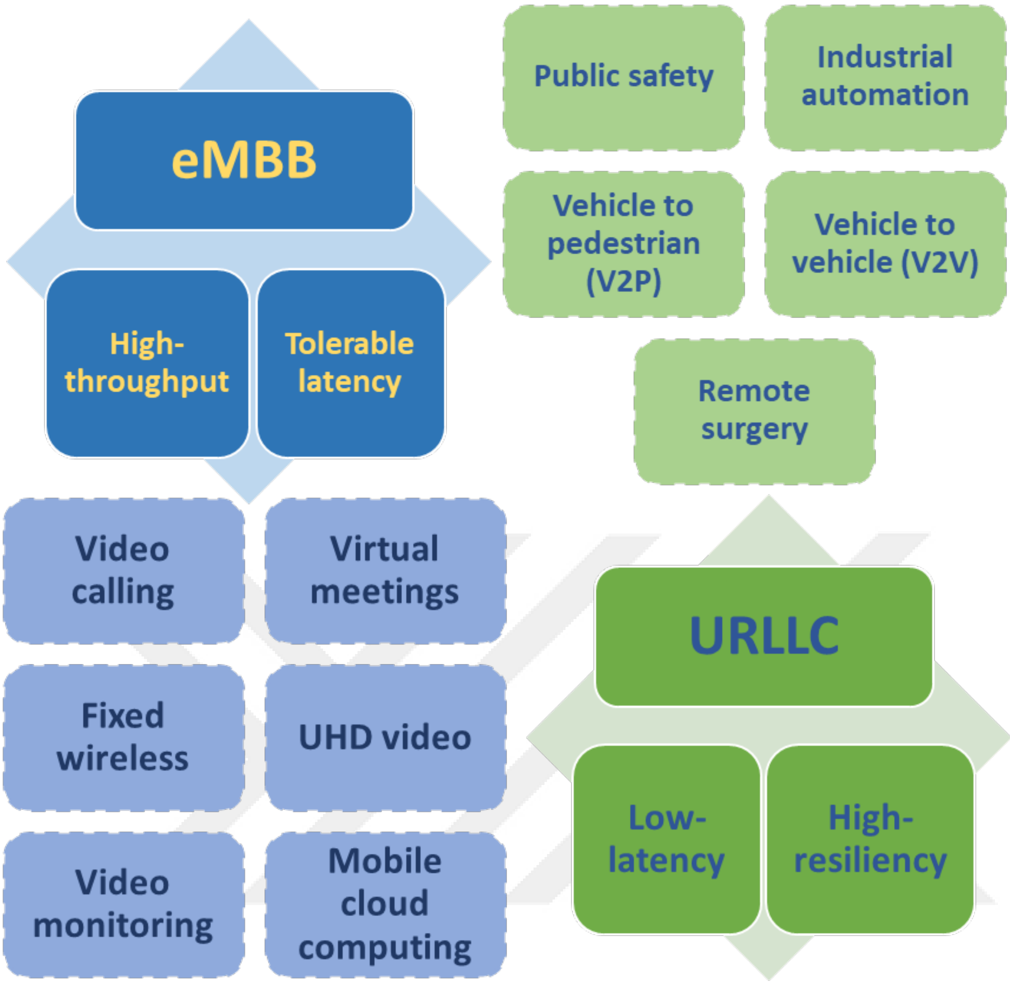

Figure 2.1: Primary slices in 5GB networks: eMBB and URLLC

addition, URLLC communications must provide very high reliability and low latency, as shown in Table 2.1.

Table 2.1: Latency and reliability requirements for URLLC slices [1]

Use Case	Max Latency	Reliability
Factory Automation	≤ 1 ms	99.999%
Autonomous Vehicles	≤ 10 ms	99.999%
Remote Surgery	≤ 1 ms	99.9999%
Smart Grids	≤ 5 ms	99.999%
Industrial Robotics (Motion Control)	≤ 1 ms	99.9999%
Smart Factories (Process Automation)	≤ 10 ms	99.999%

The coexistence of URLLC and eMBB slices in downlink traffic introduces significant challenges in radio resource allocation. Prioritizing URLLC users to meet stringent requirements often leads to performance degradation for eMBB slices. This chapter first presents the resource grid, downlink data processing, and the scheduling mechanisms for coexisting slices in 5G NR, addressing this issue. We then introduce the fundamentals of artificial intelligence (AI) necessary to understand the AI-driven preemption methods presented in subsequent chapters.

2.1 Resource Allocation in 5G NR

The resource allocation in 5G NR aims to optimize Quality of Service (QoS) by efficiently distributing radio resources among users. With limited time and frequency resources, the scheduler implementation within the 5G New Radio (NR) base station (gNB) has crucial importance. As depicted in Fig. 2.2, the 5G radio resource grid is structured into Resource Blocks (RBs), representing the smallest scheduling unit. Each RB consists of 12 consecutive subcarriers. Furthermore, the grid comprises Resource Elements (REs), where each RE corresponds to one sub-carrier and one symbol. REs are the fundamental units for transmitting modulated bits, with the number of bits per RE determined by the modulation and coding scheme (MCS).

The minimal scheduling unit in 5G NR is a slot, consisting of 14 symbols in the time domain (or 12 symbols when an extended cyclic prefix is used). The duration of a slot is determined by the numerology and sub-carrier spacing (SCS). For instance, with numerology 0, the sub-carrier spacing is 15 kHz, resulting in a slot duration of 1 ms. In contrast, with numerology 1, the sub-carrier spacing increases to 30 kHz, reducing the slot duration to 0.5 ms. This pattern continues for higher numerologies, where the slot duration decreases proportionally with increasing SCS [4,7,10].

2.2 Downlink Data Processing in 5G NR

A Transport Block (TB) is a data packet used for communication between the physical layer and the Media Access Control (MAC) layer in 5G NR wireless networks. The downlink data processing for TBs involves several steps before transmission over the air interface, as illustrated in Fig. 2.3. Initially, up to 2 TBs undergo physical layer

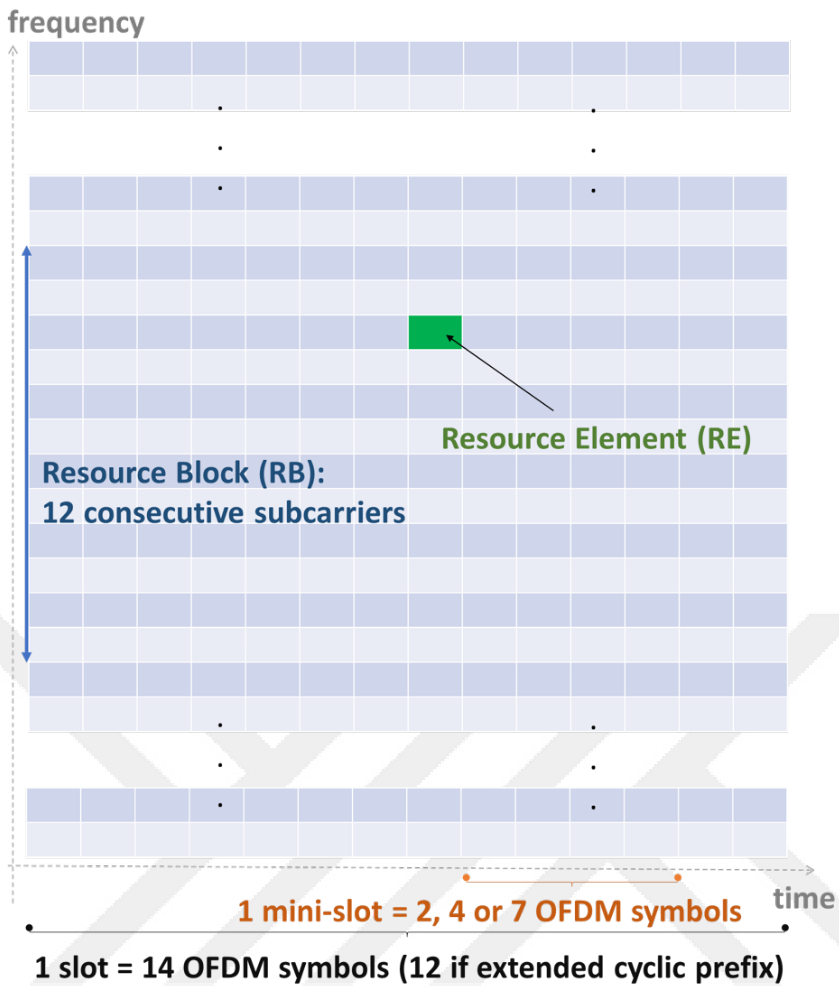

Figure 2.2: Illustration of the 5G NR resource grid

processing at the transmitter before being mapped onto the Physical Downlink Shared Channel (PDSCH). In the first step, a Cyclic Redundancy Check (CRC) is appended to the TB for error detection. It is followed by selecting the appropriate Low-Density Parity-Check (LDPC) base graph based on the transport block size to allow proper forward error correction reliability. Then, each TB is segmented into code blocks. For each code block, a CRC is attached to enable the detection of failing code blocks after transmission. LDPC coding and rate matching are applied to each code block, which are then concatenated and scrambled using a unique identifier corresponding to the UE, UE group, and usage. The scrambled data is modulated according to channel conditions and throughput requirements. The modulation type and the coding conditions are controlled by the Modulation and Coding Scheme (MCS). Finally, the modulated data is transmitted to the receiving UE. These procedures are detailed in 3GPP TS 38.212 [11] for efficient and reliable data transmission.

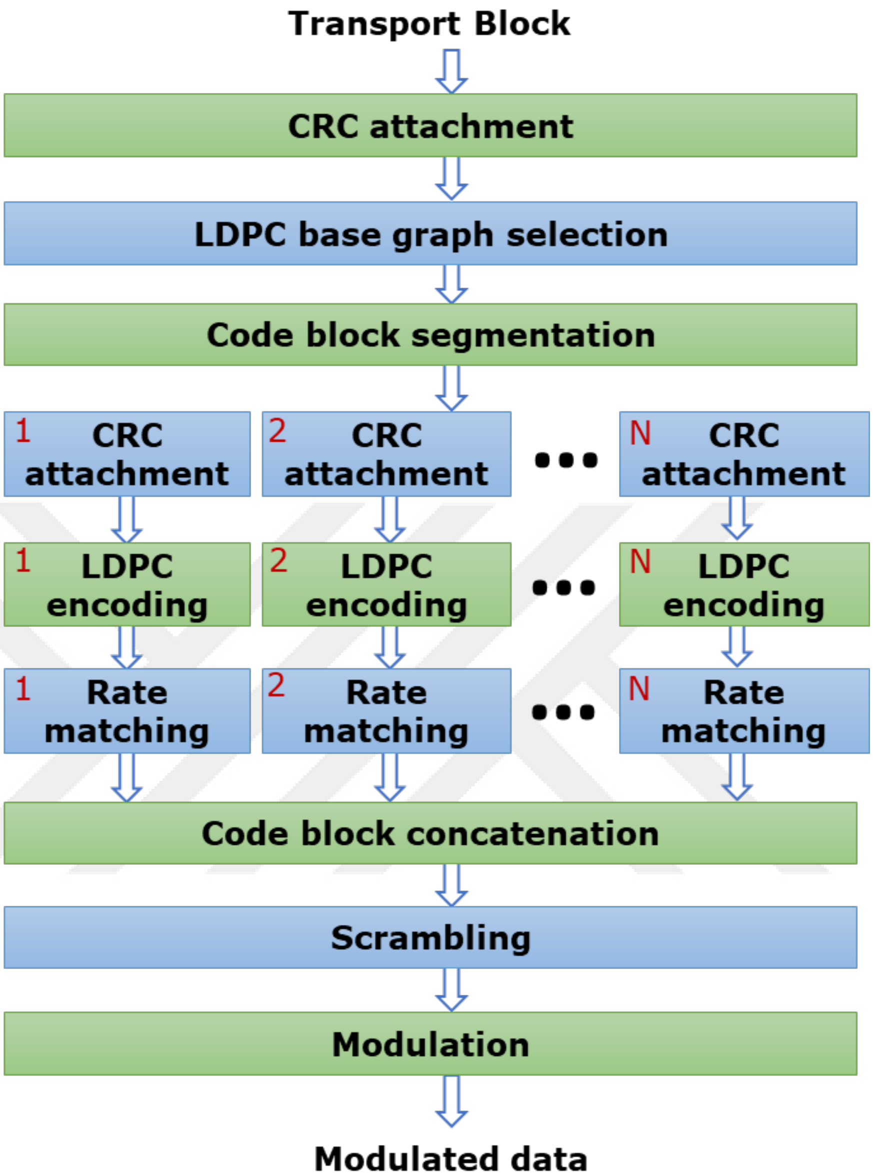

Figure 2.3: Downlink data processing chain in 5G NR

The resources in the orthogonal frequency-division multiplexing (OFDM) resource grid are allocated in both the frequency and time domains. Time-domain allocation determines the transmission duration for each user, while frequency-domain allocation depends on factors such as carrier frequency, available bandwidth, and sub-carrier spacing. The scheduler, located within the 5G NR base station (gNB), is responsible for allocating these limited radio resources to user equipment (UEs) in an efficient manner. The most granular unit of 5G NR is a Resource Block (RB), which is composed of 12 consecutive subcarriers. Unit of a subcarrier and a symbol unit are referred to as a Resource Element (RE), which serves as the building block component for modulated bits in 5G NR. Each RE can have several bits depending on the selected Modulation and Coding Scheme (MCS) [12]. The smallest 5G NR scheduling unit is a slot consisting of 14 symbols in the time domain (or 12 symbols when the extended cyclic prefix is used). The duration of a slot is determined by the numerology and sub-carrier spacing (SCS). For example, with numerology 0, the SCS is 15 kHz, resulting in a slot duration of 1 ms. Similarly, with numerology 1, the SCS increases to 30 kHz, reducing the slot duration to 0.5 ms, and so on [10].

A gNB can send the location and size information of downlink data to a UE through two scheduling alternatives: semi-persistent scheduling and dynamic scheduling. In dynamic scheduling, the gNB uses Downlink Control Information (DCI) on the Physical Downlink Control Channel (PDCCH) to transmit this information. The UE must periodically check the PDCCH for data assignments according to its Discontinuous Reception (DRX) setting. On the other hand, in semi-persistent scheduling, the UE is pre-configured with periodicity information about the configured grants or assignments using the Radio Resource Control (RRC), and these grants are provided periodically for PDSCH transmission upon activation via Downlink Control Information (DCI). Both types of scheduling possibilities can be valid for both types of PDSCH mapping, i.e., Type A and Type B. Type A is targeted towards traditional slot definition since the transmission of the PDSCH will start from the 2^{nd} or 3^{rd} symbol. Type B, on the other hand, will be used for mini-slot scheduling, which supports up to 2, 4, and 7 symbols of PDSCH and can begin from any symbol. The ultra-low latency requirements of URLLC services can only be met using Type B scheduling.

2.3 Preemption in 5G NR

To address the diverse requirements of 5G slices, 3GPP introduced a preemption mechanism to facilitate the coexistence of low-latency URLLC slices and high-throughput eMBB slices [4]. This mechanism enables URLLC transmissions to interrupt ongoing eMBB transmissions by dynamically reclaiming eMBB resources when immediate access is necessary to meet URLLC's stringent latency constraints. While preemption prioritizes URLLC's urgent needs, it must be carefully managed to mitigate adverse effects on overall network performance, particularly in terms of eMBB throughput.

In practice, the 3GPP preemption mechanism operates through Downlink Control Information (DCI) signaling, which conveys preemption details to UEs, indicating which eMBB resource blocks have been reallocated to URLLC transmissions. This real-time adjustment enables eMBB users to modify their decoding processes, thereby limiting performance degradation [13]. The optimal implementation of preemption requires the careful selection of lower-priority or low-quality eMBB transmissions to ensure minimal impact on eMBB quality while satisfying URLLC's latency demands, which is crucial for applications such as autonomous vehicles and industrial automation.

In the gNB's scheduling operations, various traffic types and user priorities must be considered. As illustrated in Fig. 2.4, the scheduler can dynamically preempt ongoing PDSCH data by reallocating eMBB resources to URLLC transmissions. This preemption mechanism supports latency-sensitive applications by facilitating the Preemption/Puncturing Indication (PI), which is carried in the PDCCH DCI. PI informs UEs of punctured eMBB resources in the upcoming slot [4]. Upon receiving this indication, the UE discards the affected data in its buffers and prepares for retransmissions. If a UE fails to decode the punctured data or does not receive a PI, it may request HARQ retransmissions, which can result in potential negative acknowledgments (NACKs). This mechanism supports low-latency URLLC transmissions, but excessive HARQ retransmissions can degrade eMBB performance. The definition of efficiency in those resource allocation strategies can be defined by their ability to balance latency, reliability, and throughput.

Fig. 2.5 demonstrates a code block group (CBG)-based preemption scenario. A UE receives a downlink PDSCH transmission composed of multiple code blocks (CBs) grouped into CBG0 and CBG1. When URLLC data, requiring immediate low-latency delivery, arrives at the gNB under high network load, the gNB preempts part of the ongoing eMBB transmission by allocating CB4 and CB5 to the URLLC data, thereby puncturing eMBB resources. Consequently, the UE may fail to decode the affected

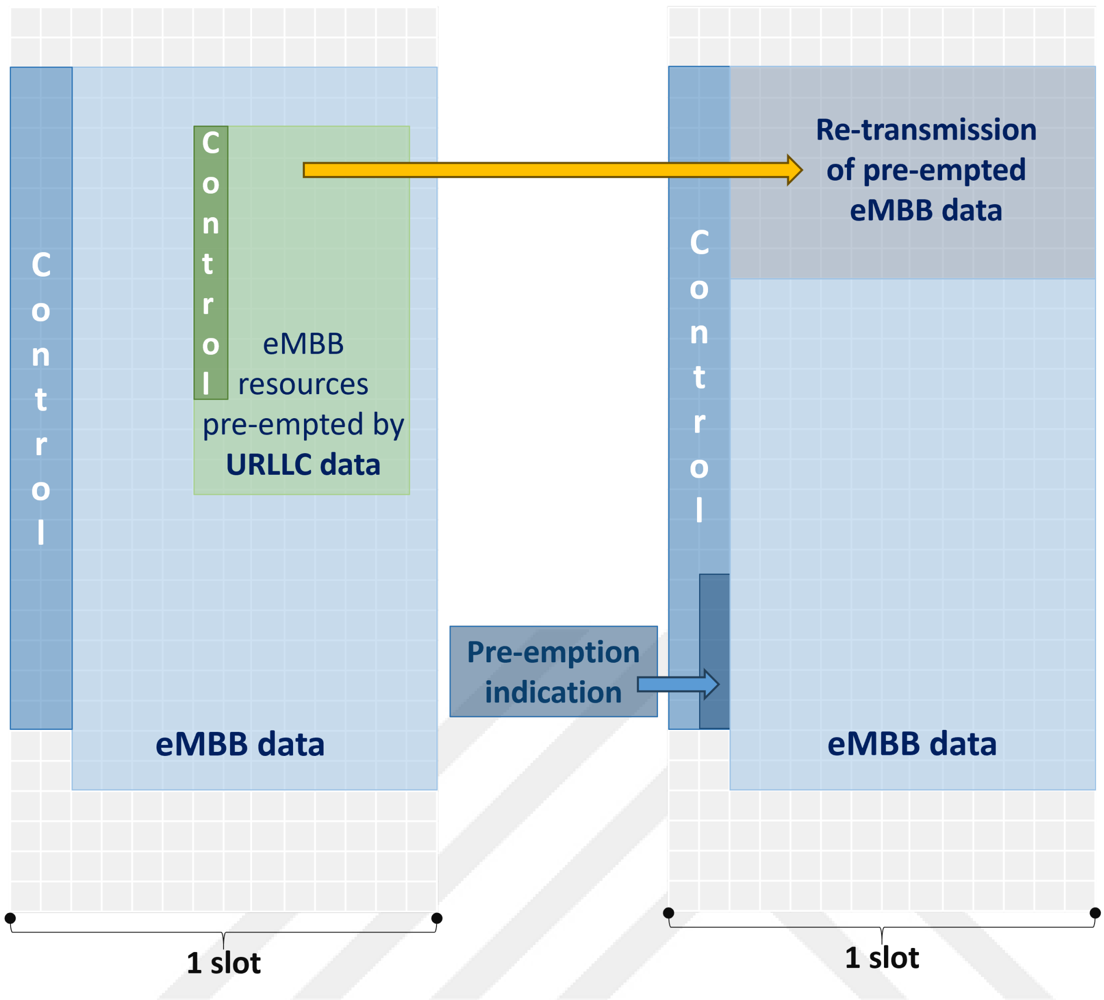

Figure 2.4: Preemption of eMBB slice to support URLLC slice

CBs in CBG1. The gNB can retransmit only the impacted CBs in subsequent slots or rely on the PI to notify the UE of punctured resources, thereby mitigating throughput loss. This selective retransmission reduces unnecessary overhead and supports efficient HARQ operations, minimizing eMBB throughput degradation while preserving timely URLLC packet delivery.

Table 2.2 summarizes the 3GPP-defined preemption mechanism, including decision criteria and implementation considerations for managing resource allocation between eMBB and URLLC slices in constrained networks.

Conventional preemption approaches can establish a baseline for scheduling eMBB and URLLC slices concurrently. However, their effectiveness may be degraded by the dynamic behavior of traffic, as well as fluctuations in resource availability in 5G networks. AI presents a viable approach through adaptability, enabling

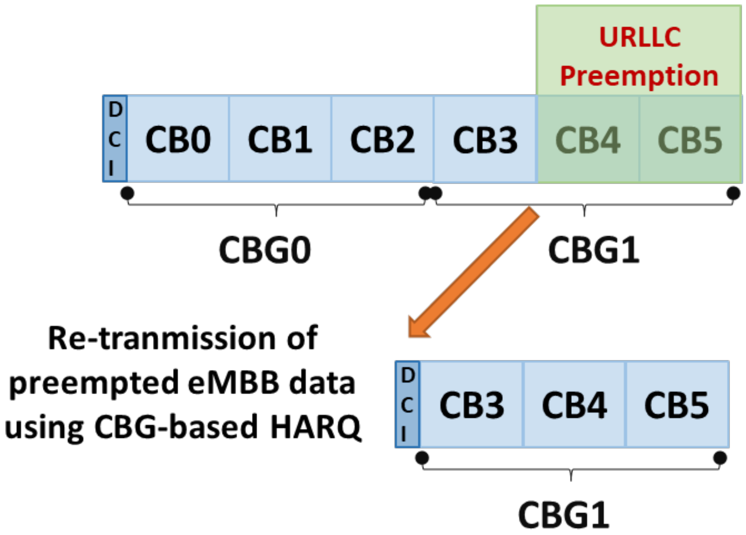

Figure 2.5: CBG-based preemption

Table 2.2: Summary of preemption mechanism in 5G NR.

Step	Description
1	If the URLLC slice's latency requirements can be met with an N-slot advance decision, the data must be transmitted at the earliest opportunity.
2	If the latency constraints cannot be fulfilled, two options are available:
2.1	Identify an unoccupied resource for transmission.
2.2	Preempt already allocated resources.
3	When eMBB data is preempted by URLLC transmissions:
3.1	The eMBB receiver may correct errors using built-in error correction mechanisms.
3.2	If error correction fails, retransmission of the affected data is initiated.

decision-making over data at will. Notably, reinforcement learning (RL) and deep Q-networks (DQN) can be leveraged to enhance the efficiency of preemption mechanisms.

The following section provides key AI terminology and mathematical formulas to form the foundation of the RL- and DQN-based methods presented in the subsequent chapters. The AI solutions presented here aim to maximize resource utilization, improve preemption efficiency, and address the critical trade-off between eMBB throughput and URLLC latency.

2.4 Artificial Intelligence

Artificial Intelligence (AI) has emerged as a revolutionary paradigm to address complex resource allocation challenges in next-generation wireless networks. The inherent complexity, heterogeneity, and dynamism of 5G and future networks, characterized by varying service requirements, time-evolving channel conditions, and multi-dimensional objectives, render traditional model-based approaches less appropriate [14,15]. AI-driven techniques, particularly those based on reinforcement learning, offer promising alternatives by enabling systems to learn optimal resource management policies directly from experience, adapt to time-varying environments, and balance conflicting performance metrics without requiring explicit mathematical models of the underlying system dynamics [16]. This section presents the theoretical foundations and advanced implementations of reinforcement learning paradigms that form the backbone of our proposed resource allocation framework for eMBB-URLLC multiplexing.

Machine learning approaches can be broadly categorized into three paradigms based on their learning methodology [17]:

- Supervised Learning: Models learn a mapping function from labeled input-output pairs, requiring extensive training data with ground truth annotations. While effective for classification and regression tasks with stable patterns, supervised approaches are limited in dynamic environments where optimal decisions depend on sequential interactions and delayed feedback [18].

- Unsupervised Learning: These techniques discover inherent structures and patterns in unlabeled data through clustering, dimensionality reduction, and representation learning. Although valuable for preprocessing and feature extraction, unsupervised methods alone cannot directly address the sequential decision-making problems typical in resource allocation.

- Reinforcement Learning: RL enables agents to learn optimal behaviors through trial-and-error interactions with an environment [16], making it particularly suited for sequential decision-making under uncertainty—the predominant characteristic of wireless resource management [19].

For the eMBB-URLLC multiplexing problem, reinforcement learning offers distinct advantages: it can operate without explicit knowledge of the underlying system model, adapt to non-stationary network conditions, balance multiple competing objectives through properly designed reward functions, and learn complex state-action mappings that capture intricate interdependencies between network parameters [20].

2.4.1 Reinforcement learning

Reinforcement learning (RL) formalizes sequential decision-making through Markov Decision Processes (MDPs), which are defined by the tuple $\langle \mathscr{S}, \mathscr{A}, \mathscr{P}, \mathscr{R}, \gamma \rangle$ [21]. Here:

- $\mathscr{S}$: The state space represents the complete network conditions, including channel quality indicators, buffer states, traffic demands, and Quality of Service (QoS) metrics.
- $\mathscr{A}$: The action space encompasses all possible resource allocation decisions, such as Resource Block (RB) assignments, power allocations, and scheduling priorities.
- $\mathscr{P} : \mathscr{S} \times \mathscr{A} \times \mathscr{S} \to [0,1]$: The state transition probability function models the dynamics of the network environment.
- $\mathscr{R} : \mathscr{S} \times \mathscr{A} \to \mathbb{R}$: The reward function encodes the optimization objectives, such as maximizing throughput, minimizing latency, and ensuring reliability guarantees.
- $\gamma \in [0,1]$: The discount factor balances immediate rewards against long-term performance.

The agent's objective in RL is to find an optimal policy $\pi^* : \mathscr{S} \to \mathscr{A}$ that maximizes the expected cumulative discounted reward, expressed as $\mathbb{E}_\pi \left[\sum_{t=0}^{\infty} \gamma^t r_t\right]$ [16]. The value function $V^\pi(s)$, which represents the expected return when starting in state s and following policy π thereafter, is defined as:

$$V^{\pi}(s) = \mathbb{E}_{\pi}\left[\sum_{k=0}^{\infty}\gamma^{k} r_{t+k} \mid s_t = s\right] \tag{2.1}$$

Similarly, the action-value function $Q^{\pi}(s,a)$ quantifies the expected return after taking action a in state s and following policy π thereafter [22]:

$$Q^{\pi}(s,a) = \mathbb{E}_{\pi}\left[\sum_{k=0}^{\infty}\gamma^{k} r_{t+k} \mid s_t = s, a_t = a\right] \tag{2.2}$$

The optimal policy can be derived from the optimal action-value function $Q^{*}(s,a)$, which satisfies the Bellman optimality equation [23]:

$$Q^{*}(s,a) = \mathbb{E}\left[r_t + \gamma \max_{a'} Q^{*}(s_{t+1},a') \mid s_t = s, a_t = a\right] \tag{2.3}$$

The concept of a reinforcement learning architecture is illustrated in Fig. 2.6.

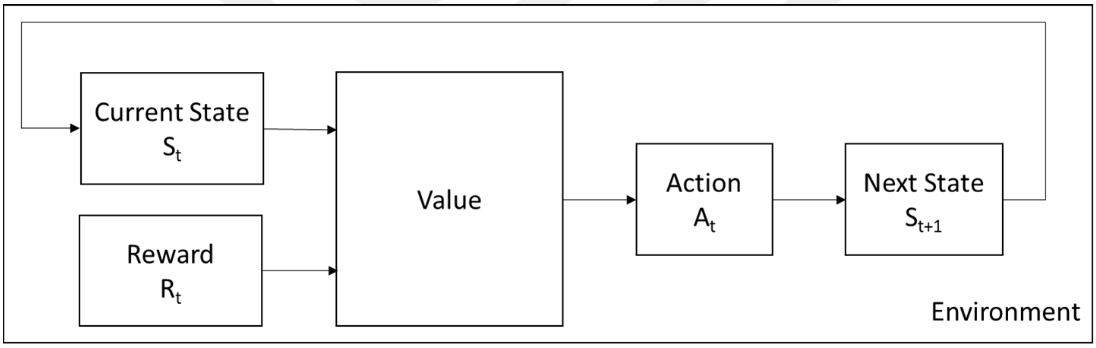

Figure 2.6: Reinforcement Learning Architecture

The Q-learning algorithm, a model-free off-policy reinforcement learning method introduced by Watkins [24], iteratively approximates the optimal action-value function $Q^{*}(s,a)$ through temporal difference updates:

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\left[r_t + \gamma \max_{a'} Q(s_{t+1},a') - Q(s_t,a_t)\right] \tag{2.4}$$

Here, $\alpha \in (0,1]$ denotes the learning rate, which controls the step size of the updates, and γ represents the discount factor, which determines the importance of future rewards.

For wireless resource allocation, the Q-learning process involves the following steps [25]:

1. Observation: Monitoring the network state, including channel conditions, buffer occupancy, and traffic demands.
2. Action Selection: Making resource allocation decisions based on an exploration-exploitation strategy, such as the ε-greedy approach [16].
3. Execution: Implementing the allocation decisions and observing the resulting performance outcomes.
4. Learning: Updating the Q-function based on the observed rewards and transitioning to new network states.

The concept of a Q-learning architecture is illustrated in Fig. 2.7.

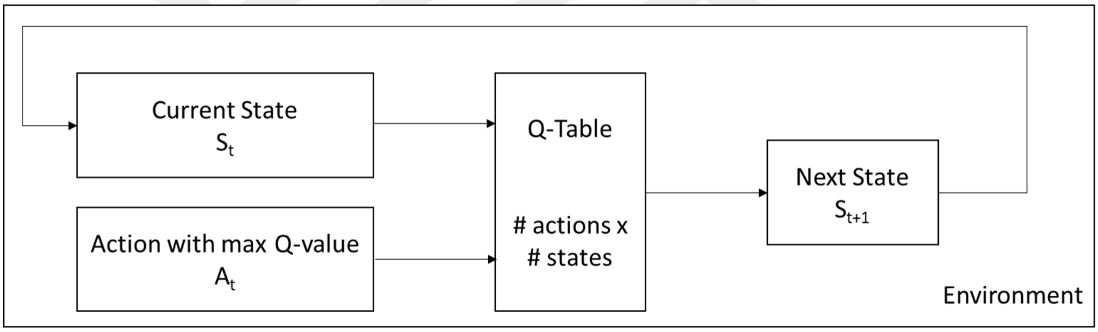

Figure 2.7: Q-Learning Architecture

2.4.2 Deep reinforcement learning

When dealing with high-dimensional state spaces, typical in wireless networks, traditional tabular Q-learning becomes computationally infeasible. Deep Q-Networks (DQN) [26]–[29] address these scalability limitations by leveraging parametric approximation through deep neural networks:

$$Q(s,a;\theta) \approx Q^*(s,a)1001[26] \tag{2.5}$$

Here, θ represents the parameters of the neural network. This function approximation enables generalization across similar states and actions, significantly improving learning efficiency in large state spaces [26].

The DQN architecture incorporates several pivotal innovations to overcome the inherent instability of function approximation in reinforcement learning [26]:

1. Experience Replay: A buffer $\mathscr{D}$ stores transitions (s_t, a_t, r_t, s_{t+1}) to:
    - Break temporal correlations in the training data,
    - Improve sample efficiency by enabling multiple updates from each experience and
    - Reduce variance in updates through mini-batch sampling.
2. Target Network: A delayed copy θ^- of the main network parameters is periodically updated (every C steps) according to $\theta^- \leftarrow \theta$. This stabilizes Q-targets and prevents harmful oscillations during training.

The neural network parameters θ are optimized by minimizing the loss function proposed by Mnih et al. [26]:

$$\mathscr{L}(\theta) = \mathbb{E}_{(s,a,r,s')\sim\mathscr{D}}\left[\left(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta)\right)^2\right] \tag{2.6}$$

The architecture of a Deep Q-Network (DQN) is illustrated in Fig. 2.8. This network outputs Q-values for each possible resource allocation action. The agent selects actions using an exploration-exploitation strategy, such as the ε-greedy approach, implements resource allocation decisions, and observes the resulting rewards. The experience replay buffer stores transitions (s_t, a_t, r_t, s_{t+1}), from which mini-batches are sampled for training. This process breaks temporal correlations and improves sample efficiency [26,30]. The target network, which is updated periodically from the online network, provides stable Q-targets during training, preventing harmful oscillations in the learning process [26]. This architecture enables efficient learning of complex resource allocation policies, even in high-dimensional state spaces with non-stationary dynamics typical of 5G wireless environments [28].

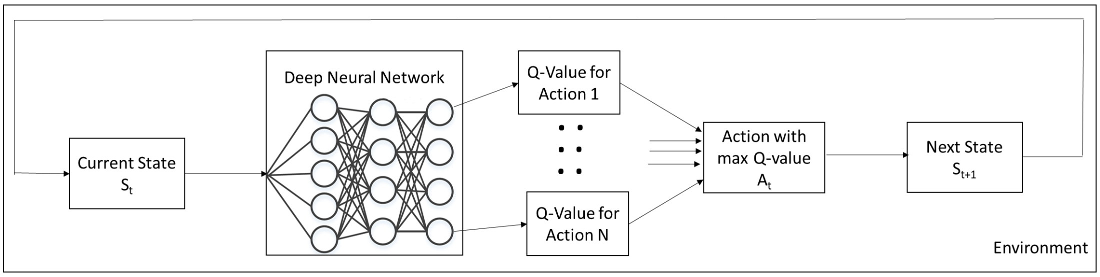

Figure 2.8: Architecture of Deep Q-Network (DQN)

2.4.3 Advanced DRL techniques

Modern deep reinforcement learning (DRL) frameworks incorporate several enhancements to the base DQN algorithm to improve stability, efficiency, and performance:

- Double DQN [27,31]: Decouples action selection and evaluation to mitigate overestimation bias in Q-values:

$$y_t = r_t + \gamma Q(s_{t+1}, \arg\max_{a'} Q(s_{t+1}, a'; \theta); \theta^-) \tag{2.7}$$

- Prioritized Experience Replay [32]: Samples transitions with probability proportional to the magnitude of the temporal difference (TD) error:

$$P(i) = \frac{p_i^\alpha}{\sum_k p_k^\alpha} \tag{2.8}$$

  where $p_i = |\delta_i| + \varepsilon$ represents the priority of transition i, and δ_i is the TD-error.

- Dueling Network Architecture [33]: Separately estimates state values $V(s;\theta)$ and action advantages $A(s,a;\theta)$:

$$Q(s,a;\theta) = V(s;\theta_V) + \left(A(s,a;\theta_A) - \frac{1}{|\mathscr{A}|} \sum_{a'} A(s,a';\theta_A) \right) \tag{2.9}$$

  This decomposition improves learning efficiency by enabling the network to learn state values independently of specific actions.

- Distributional RL [34]: Models the entire distribution of returns rather than just their expectation, capturing uncertainty and risk in the learning process.

2.4.4 Policy optimization methods

Beyond value-based methods like DQN, policy optimization approaches directly learn the policy function $\pi_\theta(a|s)$, which maps states to action probabilities:

- Policy Gradient Methods [35]: Update policy parameters in the direction of $\nabla_\theta \mathbb{E}_{\pi_\theta}[G_t]$ using the policy gradient theorem:

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta}[\nabla_\theta \log \pi_\theta(a|s) Q^{\pi_\theta}(s,a)] \tag{2.10}$$

- Actor-Critic Algorithms [36]: Combine value function approximation (critic) with direct policy optimization (actor) to reduce variance while maintaining manageable bias.

- Proximal Policy Optimization (PPO) [37]: Restricts policy updates to ensure stable learning:

$$L^{CLIP}(\theta) = \mathbb{E}_t \left[\min(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}_t)\right] \tag{2.11}$$

  where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$ is the probability ratio between the new and old policies.

3. RELATED WORK

Resource allocation constitutes a fundamental challenge that permeates numerous technological domains. As telecommunications systems undergo continuous evolution, driven by advancements in technology and escalating user demands for innovative services, the complexity of achieving optimal scheduling mechanisms has intensified significantly. This issue is particularly salient in the context of 5GB systems, where the expansion of available bandwidth and the pursuit of higher data rates necessitate the deployment of larger grid structures and increased layer counts. Consequently, these requirements exacerbate computational workloads and amplify the inherent complexity of the underlying optimization problems. In response to these concerns, the research world has increasingly turned to artificial intelligence-based methodologies, specifically deep learning techniques. The objective here has been to address the complex and multifaceted demands of modern resource allocation models.

Table 3.1 categorizes recent research studies on 5GB resource allocation into five approaches: proportional fair scheduling, supervised learning, deep learning, reinforcement learning, and deep reinforcement learning. Categorization not only highlights the diversity of methodologies employed but also provides an organized structure within which to consider the development and progress of resource allocation techniques in modern telecommunications networks. Through the organization of the literature in this manner, the table facilitates a comparative review of the strengths, limitations, and relevance of each approach, thereby enhancing the development of a more comprehensive understanding of the field.

Reinforcement learning (RL) has emerged as a prominent methodology in solving resource allocation issues in 5G systems [63]–[69]. For the LTE mission-critical services case, as demonstrated by [59], the eNodeB (LTE base station) is modeled as the agent, and the action space is defined by resource block assignments. The state in this model is constructed in accordance with the delay minimization goal, where state

Table 3.1: Summary of research on resource allocation problem.

Scheduling Method	Ref. No	Application	Pub. Date
Proportional Fair	[38]	5G eMBB	2019
	[39]	5G NR-U	2022
Supervised Learning	[40,41]	5G Cloud RAN	2016, 2018
	[42]	6G Resource Allocation	2023
	[43]	5G Network Slicing	2023
Deep Learning	[44]	5G UDN	2018
	[45]	5G Cloud RAN	2018
	[46]	5G URLLC	2020
	[47]	5G Radio Resource Management	2022
	[48]	6G Spectrum Allocation	2023
RL	[49]	5G D2D Communication	2016
	[50]	5G Cloud RAN	2017
	[51,52]	5G RAN	2017, 2018
	[53]	5G, Smart Grid	2019
	[54]	5G Network Slicing	2019
	[**?**,55]	5G RAN Network Slicing	2020
	[56]	5G Fog-RAN	2020
	[57,58]	5G V2V Communication	2020
	[59]	LTE Mission Critical	2020
	[60]	5G mmWave	2020
	[61]	5G V2X Communication	2022
	[62]	6G Integrated Sensing	2023
DRL	[63]	V2V Communication	2018
	[64,65]	UDN	2018, 2021
	[66]	5G Network Slicing	2018
	[67]	5G Mobile	2019
	[68]	6G Multi-cell Networks	2019
	[69]	5G URLLC	2020
	[70]	RAN	2020
	[71]	5G Network Slicing	2021
	[72]	5G URLLC/eMBB	2022
	[73]	6G Integrated Sensing and Communication	2022
	[74]	6G Network Resource Allocation	2023
	[75]	6G Federated Resource Allocation	2024

0 indicates that the average delay is less than a target threshold, and state 1 otherwise. The reward function is a sigmoid of the accumulated delay, giving a differentiable and continuous measure of system performance.

The agent is represented by the shift to 5G NR gNodeB (5G NR base station) in most studies [2,54,56]–[58]. This change illustrates the architectural shift between LTE and 5G networks, demonstrating the adaptability of reinforcement learning techniques to evolving network paradigms. Utilizing the gNodeB as the agent, these contributions demonstrate how RL structures can be formulated to counter the increased complexity and dynamicity of 5G systems, further solidifying RL as an effective facilitator of efficient resource allocation in future networks.

Vehicle-to-vehicle (V2V) communication is a significant area of application for 5GB resource allocation. Contributions by [57] and [58] focus on this area but employ different state representation approaches. While [57] defines the state in terms of relative positions of other vehicles with respect to a given vehicle, [58] defines it in terms of percentage of uplink/downlink data rate over channel capacity at a given instant. Developments in vehicle-to-everything (V2X) communications, as shown by Wang *et al.* [61], have extended these approaches to encompass joint resource allocation between communication and sensing needs. Their system optimizes resource blocks and power allocation, meeting strict latency requirements for safety-critical applications, and represents a breakthrough in preparedness for autonomous driving applications of 5GB systems.

Ultra-dense networks (UDNs) represent another area of application for 5GB system resource allocation, and reinforcement learning has been employed in various works to address the corresponding problems. In [2], there is an event trigger-based Q-learning algorithm wherein agents are the 5G network small cells. The state space and action space refer to the user's available channel conditions and power levels, respectively. The reward function is optimized to reduce multiple performance measures, including QoS constraints, energy efficiency, and signal-to-interference-and-noise ratio (SINR) at the user. This multi-objective approach maximizes network performance by considering both user satisfaction and resource efficiency. Apart from that, [65] treats

the state space as traffic, base station transmission power for small cell, and the state of channels at a moment in time. Indicators on small cell base station links to users, assignment of power to connected users, and energy efficiency values per subcarrier after transformation are employed as the action space. Due to its energy efficiency focus, the reward function tends to prioritize this parameter. Conversely,
cite8403505 sets another objective, namely maximizing computational capacity, compared to energy efficiency. All such works demonstrate the flexibility of reinforcement learning design in meeting varied optimization objectives in UDN settings and thus unveiling its potential for addressing the intricate needs of 5G networks.

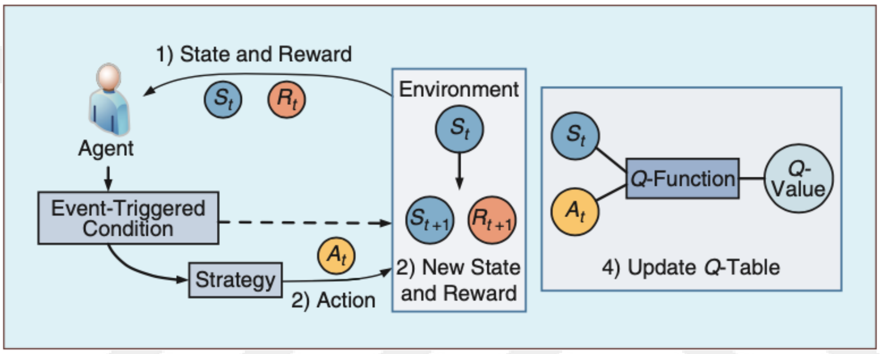

Figure 3.1: The event-triggered Q-learning algorithm process [2]

Radio Access Networks (RAN) have garnered significant attention in the research literature [**?**,50]–[52,54]–[56], with specialized variants such as Cloud-RAN and Fog-RAN being extensively investigated. For instance, [50] explores Cloud-RAN. [56] focuses on Fog-RAN and leverages Q-learning to optimize the summation of average QoS utilization across Fog-RAN nodes. The state space in this framework is defined by several key parameters, including the total allocated resource parts, average QoS utility, average CPU utilization, and the allocated CPU resources for the F-RAN node. The action space is discretized, comprising actions that represent changes in resource allocation ranging from -90% to $+90\%$, with increments of 10%. This structured approach enables precise control over resource allocation, ensuring efficient utilization and enhanced QoS in Fog-RAN environments.

Network slicing is another fundamental aspect of 5G RAN dynamic resource allocation, studied by [54] and [55]. Here, the state is represented as available resources and arriving resource requests, and the action is represented as allocating those resources. The reward function is the network utility that is calculated in terms of the weighted sum of fulfilled requests and weights assigned in proportion to their respective priorities. For comparison, [54] employs a combination of parameters to describe the state, including the slice flow ratio and the total resource requests. Recent network slicing success has been demonstrated by Lei *et al.* [71], who introduced a novel deep reinforcement learning approach that dynamically allocates resources across network slices based on real-time traffic requirements and slice demands. Their solution employs a two-agent design where one agent optimizes resource utilization and the other optimizes fairness between various classes of services. Chen *et al.* [43] also proposed a supervised learning approach that predicts traffic patterns among various network slices and enables proactive resource allocation with a significant reduction in SLA violations when compared to reactive schemes.

For multi-cell 6G networks, [68] proposes a framework that optimizes overall network throughput. This is achieved by selecting actions that assign the least number of power values to all subbands for each cell. Relying on this support, Yang *et al.* [73] have introduced a multi-agent deep reinforcement learning framework for 6G network integrated sensing and communications. Their framework optimizes resource allocation between communication and sensing operations with up to 40% improvement in spectral efficiency compared to conventional methods. Similarly, Wu *et al.* [62] presented a hierarchical reinforcement learning methodology that decomposes the complex 6G resource allocation problem into solvable sub-problems and enables simpler training and adaptation to dynamic network conditions.

Among all deep reinforcement learning (DRL)-based 5G resource allocation methods, [70] is of most relevance to our work. This paper studies a single base station, N users downlink TDMA system where average transmission power is to be minimized under average data rate constraints for all the users. Similar to our approach, it employs a Deep Q-Network (DQN) for reinforcement learning. More recent research from Sharma *et al.* [74] has introduced graph neural networks to DRL designs so that there can be improved modeling of 6G interference patterns and network topology in the resource allocation system. Their design performs better where traditional DRL approaches cannot efficiently manage the state space of large dimensions.

The challenge of the co-scheduling of eMBB and URLLC traffic has been addressed in various research studies [14,15,27]–[29,76,77]. All of these papers address the problem from a different perspective, observing the complexity and multi-dimensionality of this scheduling task. Abdelsadek *et al.* [14] were interested in minimizing eMBB data loss at the cost of ensuring minimum QoS requirements. Their solution also addresses URLLC transmission reliability by addressing transmission errors with finite block-length coding through deep supervised learning techniques. Existing research by Kim *et al.* [72] has advanced the solution further with a transformer-based architecture for detecting temporal relations in traffic patterns, leading to a 25% reduction in URLLC latency without compromising eMBB throughput.

On the other hand, Huang *et al.* [27] proposed a model-free deep reinforcement learning-based solution to counteract the negative impact of preemptive preemption for eMBB users. Huang *et al.* [28] investigated another study of eMBB throughput maximization by choosing MCS to offer link reliability to eMBB users based on the concepts of deep learning. [29] addresses the co-existence of URLLC and eMBB traffic over 5G networks, which is a baseline to our first heuristic algorithm. It designs separate scheduler algorithms for URLLC and eMBB scenarios. The eMBB scheduler uses a rate-dependent utility function, while the URLLC scheduler uses a Deep Reinforcement Learning (DRL) algorithm to optimize QoS. To enable concurrent scheduling, it uses a preemption policy that preempts eMBB packets for URLLC traffic. The eMBB codeword uses an inner erasure code with the capability of correcting up to a certain number of erased mini-slots. By preemption of the eMBB codeword, the base station is able to allocate particular resources to URLLC traffic. In addition, the algorithm uses a first-in-first-out (FIFO) queue for URLLC packets to enable efficient handling of high-priority traffic.

The transition from 5G to 6G introduces new opportunities and challenges in terms of resource allocation. Federated learning methodologies are one of the primary developments in 6G resource allocation research. Liu *et al.* [75] introduced a federated deep reinforcement learning framework that supports collaborative training among a group of edge devices with preservation of data privacy. Their approach exhibits fine adaptability to diverse network topologies, with nearly optimal allocation of resources with much reduced signaling overhead. Similarly, Liu *et al.* [42] developed a supervised learning approach that couples centralized training and distributed inference and supports real-time resource allocation decisions at the edge of the network. Zhang et al. [48] developed an interpretable deep learning-based spectrum allocation model for 6G networks to address conventional deep learning models. The method produces explainable allocation decisions with improved performance, an important consideration in regulatory compliance and operator confidence in automated resource allocation networks.

Recent research has focused more on preemption as a preemption technique for optimal utilization of resources in 5G networks. Work by researchers such as [78]–[80] has investigated methods of mitigating the effect of preemption on eMBB throughput. Building on these foundations, Chai *et al.* [47] introduced a predictive deep learning approach that anticipates traffic patterns to optimize preemption decisions, achieving a significant reduction in eMBB throughput degradation while maintaining URLLC reliability requirements.

To ensure practical applicability in real-world deployments, such preemption strategies and scheduling mechanisms should align with 3GPP specifications [4,81,82]. These standards are critical for meeting the stringent reliability and latency requirements of URLLC while maintaining the performance of eMBB services. While numerous studies have explored reinforcement learning-based scheduling approaches, relatively few adhere to 3GPP standards regarding preemption patterns or resource grid configurations. In contrast, our framework strictly complies with 3GPP TS 38.214 [7] for Channel Quality Indicator (CQI) to SINR mapping and TS 38.321 [5] for preemption signaling, ensuring full compatibility with 5G NR specifications. For instance, our standardized symbol patterns (Figure 6.1) and mini-slot scheduling mechanisms are directly derived from 3GPP guidelines, facilitating seamless integration with commercial gNodeB implementations. Furthermore, our AI-based approach, described in Chapter 6, builds upon and enhances the framework established in Chapter 5 by integrating a context-aware multi-Deep Q-Network (DQN) architecture into a 5G system-level simulator. This simulator incorporates advanced preemption techniques and leverages the benefits of training multiple DQN models based on CQI and packet sizes. This approach enables dynamic adaptation of preemption decisions based on real-time channel conditions, significantly improving the system's responsiveness to varying network demands.

4. LINK-LEVEL EVALUATION OF HEURISTIC PREEMPTIVE SCHEDULING ALGORITHM

With network services evolving at an ever-faster pace, 5G and beyond (5GB) systems face a crucial challenge in serving multiple slices concurrently. With increasingly varied quality-of-service (QoS) requirements per slice, scheduling decisions become more complex. The coexistence of two of the essential slices, eMBB and URLLC, is examined throughout this thesis. While eMBB services require high throughput, URLLC requires ultra-low latency and high reliability. These distinct needs of different slices result in a challenging compromise. The majority of current methods favor one type of traffic over the other, resulting in inefficient spectrum utilization. This chapter presents the baseline research for the procedures outlined in this thesis. In this chapter, a preemptive heuristic scheduling methodology is examined through a link-level simulator to investigate the impact of preemption on the coexistence problem. Moreover, this chapter also provides insights into the performance of preemptive scheduling, laying the groundwork for the QoS-aware and learning-based approaches that are later employed in subsequent chapters.

4.1 System Model: Resource Allocation Using Preemption

For downlink resource allocation in an OFDM system consisting of user equipments (UEs), we introduce two schedulers. eMBB and URLLC schedulers allocate resources at a slot and mini-slot level, respectively. The allocation of mini-slot level allows for a finer granularity in serving URLLC slices with lower latency. As common schedulers need to allocate resources some slots prior to transmission, preemption of eMBB resources may be inevitable to satisfy ultra-low latency requirements of URLLC slices. We propose a preemption-based heuristic algorithm called *Patient and Intelligent Preemption* (PIPE) [83]. In the PIPE algorithm, eMBB and URLLC schedulers operate concurrently. The intelligence in the algorithm name stems from its ability to make preemption decisions, aiming to minimize eMBB throughput loss. The patience in PIPE comes from strategically delaying URLLC scheduling to minimize eMBB throughput loss while satisfying URLLC latency constraints. With the link-level simulations, it is shown that PIPE achieves superior performance compared to conventional methods, maintaining robust operation across varying URLLC traffic loads. The algorithm effectively balances throughput and latency requirements. PIPE can serve as an alternative to preemptive-based resource allocation strategies for 5GB networks.

The eMBB scheduler allocates downlink data in code blocks with a water-filling approach. The modulation orders are determined by Channel Quality Indicators (CQIs). To eliminate the need for dynamic link adaptation for this initial study, we assume error-free CQI reporting and quasi-static channel conditions within a frame. This work focuses primarily on the coexistence of eMBB and URLLC slices rather than maximizing the eMBB scheduler's performance.

We extend the LTE Rel-12 CQI-MCS mapping from [3] by integrating NR specifications from 3GPP 38.214 [7] as in the modified Table 4.1. This table also shows the observation ratios based on CQI reporting frequencies from [3]'s multi-site LTE measurements. Moreover, we introduce a novel preemption allowance rate. This rate quantifies the permissible eMBB codewords to be preempted by URLLC codewords per code block so that eMBB codewords are received error-free with the help of forward error correction algorithms.

Table 4.1: CQI vs MCS mapping and observation ratios.

CQI	MCS	Modulation	Preemption All.	Observation %
4-7	4-8	QPSK	5 RB (10 bits)	8.89
8-9	11-15	16 QAM	2 RB (8 bits)	18.51
10-12	18-22	64 QAM	1 RB (6 bits)	53.69
13-15	24-28	64 QAM	0 RB (0 bits)	18.92

URLLC slices must be served with strict delay constraints. In our framework, URLLC slices are assumed to be served with traffic that consists of single resource block (RB) packets. This downlink traffic is generated through a memoryless Bernoulli process with arrival probability p_u as in Eq. 4.1. The URLLC scheduler performs per-symbol monitoring of URLLC queues. Then, it initiates preemption of eMBB resources when necessary, while simultaneously determining the optimal frequency allocation.

The Patient and Intelligent preemption (PIPE) algorithm first examines whether any eMBB resource blocks have exceeded their preemption allowance threshold. If the threshold is exceeded, that means these RBs are prioritized for preemption. When this condition isn't met, the algorithm's patience is demonstrated through its ability to strategically delay scheduling when URLLC packets still have an available latency budget. If a URLLC packet must be sent immediately, it preempts the eMBB data that would have the least impact on overall performance, specifically using a lower modulation and coding scheme.

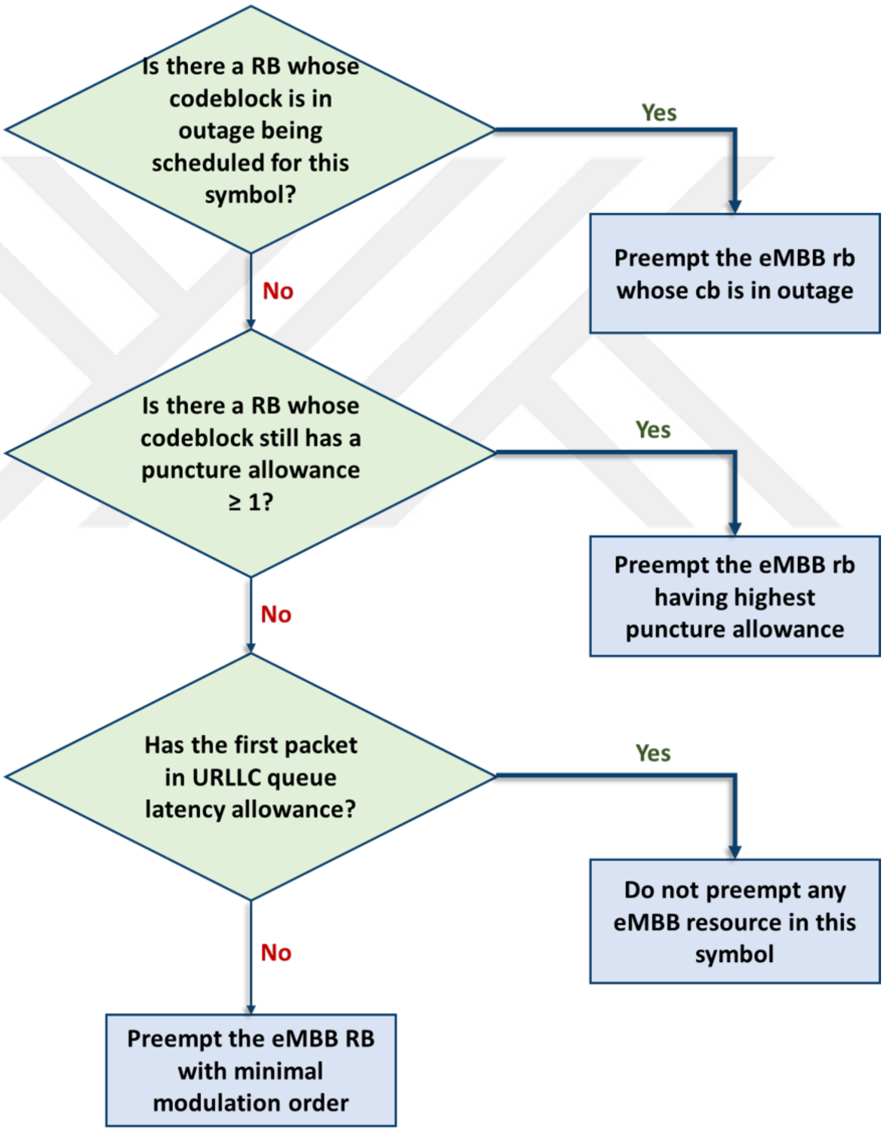

Figure 4.1: Flowchart of PIPE

Fig. 4.2 illustrates the eMBB scheduler's resource allocation for four users across two slots. Each user's resource share is represented by a distinct color: yellow, orange, blue, and gray, respectively. The allocation percentages remain consistent at 8.89%, 18.51%, 53.69%, and 18.92%, per slot, respectively, though their spatial distribution varies randomly between scheduling instances. Green-colored segments indicate URLLC preemption of eMBB resources due to higher priority traffic. The absence of preemption in some regions suggests either no pending URLLC traffic or the PIPE algorithm's patience decision to postpone preemption for more optimal future resource utilization.

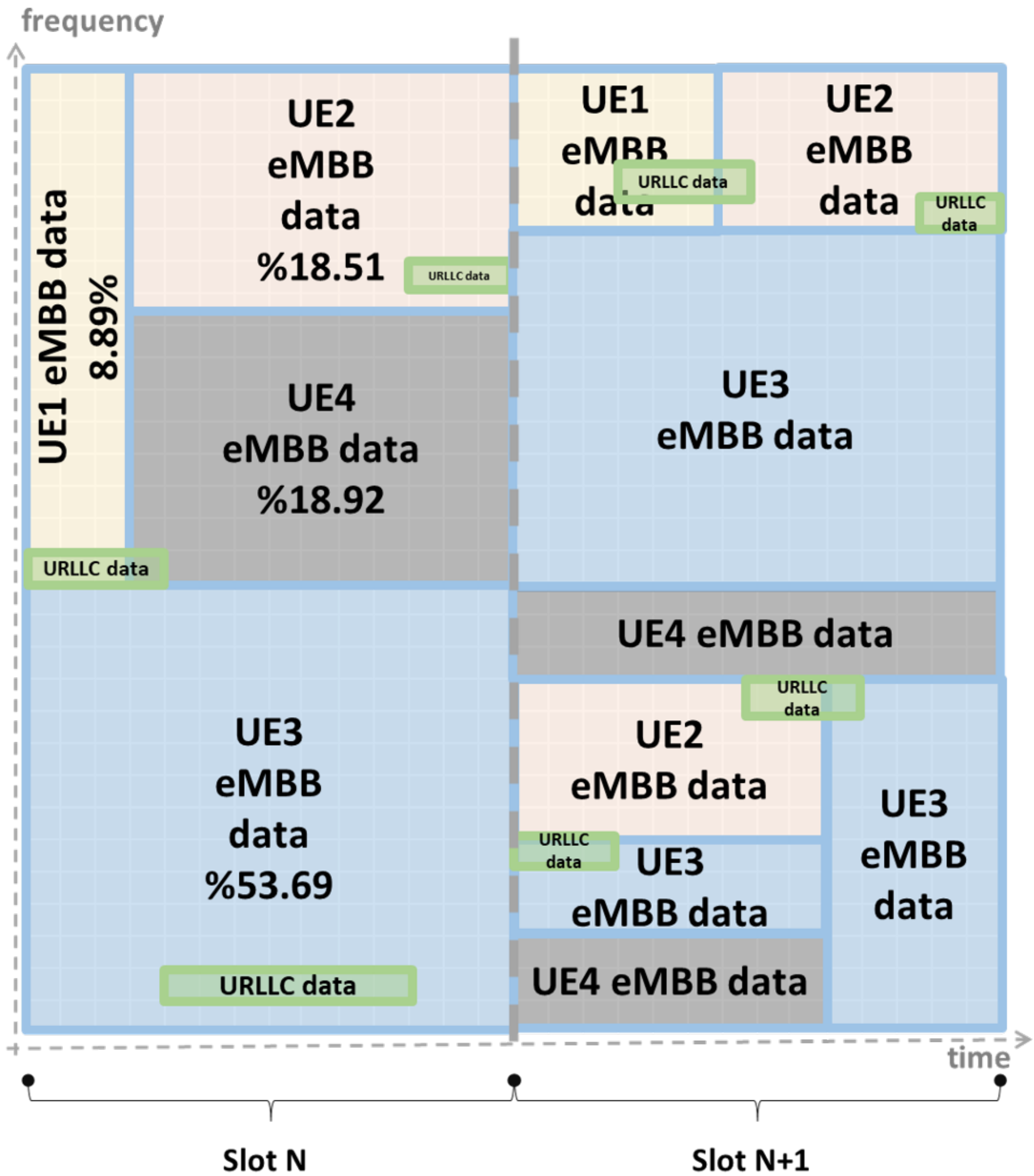

Figure 4.2: Resource allocation illustration of PIPE

4.2 Simulation Results

The simulation results are handled using a physical-level 5G link-level simulator. The experimental configuration utilizes 12 frequency resources per frame, comprising four eMBB users (modulation orders: 2, 4, 6, and 6) and one URLLC user. The users' observation ratios are 8.89, 18.51, 53.69, and 18.92, with corresponding preemption allowances of 5, 2, 1, and 0 resource blocks. We impose a maximum delay of 7 symbols for URLLC packets. The URLLC traffic load, p_u, is defined in Eq. 4.1, which follows a uniform distribution across simulations.

$$p_u \in \{0.1, 0.2, 0.3, 0.4, 0.5\} \tag{4.1}$$

For performance comparison, we evaluate the PIPE algorithm approach against four URLLC scheduling methods from [29]: Deep Reinforcement Learning (DRL) based, aggressive, threshold proportional (TP), and TP-smart. The aggressive method immediately allocates URLLC blocks to any available frequency. TP selects frequencies with codewords having the highest preemption cutoffs. TP-smart first checks for eMBB codewords in outage, utilizing those resources when available; otherwise, it behaves similarly to TP. This smart variant demonstrates optimal performance for instantaneous transfers while minimizing eMBB disruption.

Figure 4.3 demonstrates that PIPE achieves superior total throughput compared to other algorithms, while all methods meet the URLLC latency constraints. TP-smart and DRL show the next best performance. All URLLC scheduling algorithms maintain good performance when the arrival rate $p_u \leq 0.3$, showing no throughput reduction for any user. However, increasing the URLLC traffic load causes degradation in eMBB throughput across all algorithms. The proposed PIPE algorithm achieves the highest eMBB throughput, particularly for UEs utilizing higher modulation orders, resulting in minimal throughput reduction.

Simulation results demonstrate that the aggressive scheduling algorithm performs poorly in comparison to both TP-smart and DRL, which achieve significantly better results in all scenarios. However, TP-smart and DRL have certain limitations, particularly in their use of modulation order and their tendency to preempt all UEs

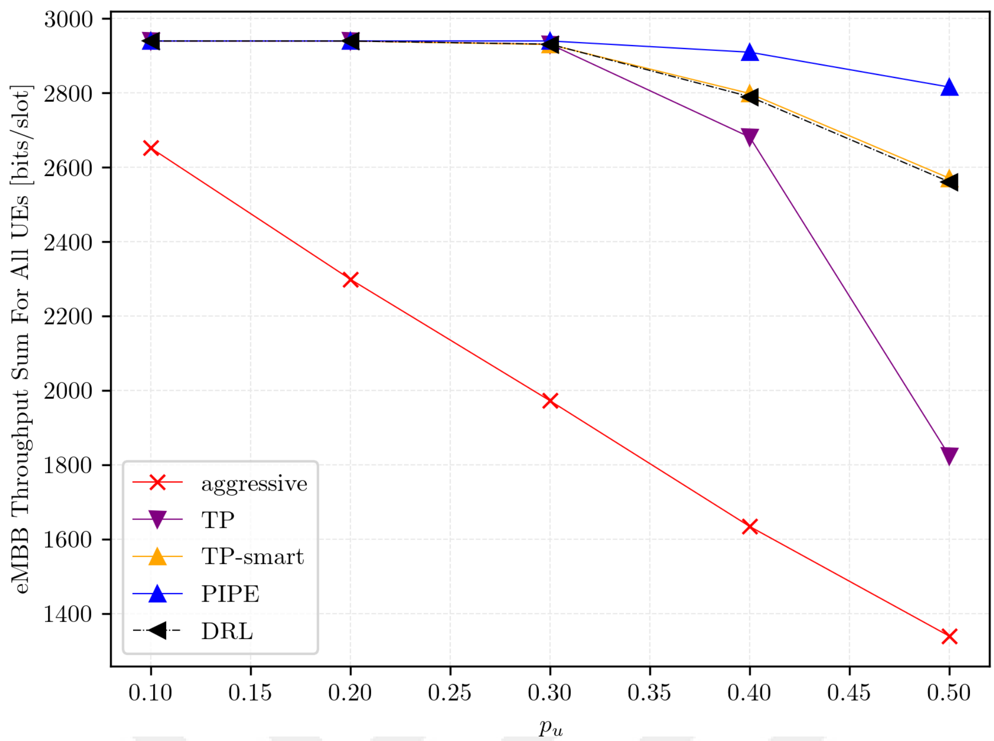

Figure 4.3: Throughput of all scheduled users

with the same priority, which reduces their effectiveness relative to our proposed algorithm. It is confirmed that all algorithms meet the latency requirements for URLLC packets, i.e., no remaining unsent URLLC traffic is allowed within the specified latency. Additionally, Fig. 4.4 highlights how PIPE transmits URLLC packets more patiently compared to other algorithms, which attempt to send them immediately, even when delaying transmission would still comply with URLLC latency constraints.

4.3 Discussion

The PIPE algorithm uses code block group transmission with 5G NR-compliant preemption. The goal of the algorithm is to balance the throughput of eMBB slices while meeting the latency requirements of URLLC slices. Previous methods lack a presented preemption allowance rate. This rate shows a threshold of preemption occurrences for each eMBB codeword to be received without requiring retransmission. Hence, exceeding it results in retransmission, which ultimately reduces eMBB throughput. The success of this algorithm is verified using a link-level simulator, where the frequency of transmissions per user is determined through real-time observation ratios, and throughput calculations utilize CQI-based modulation orders. Simulations

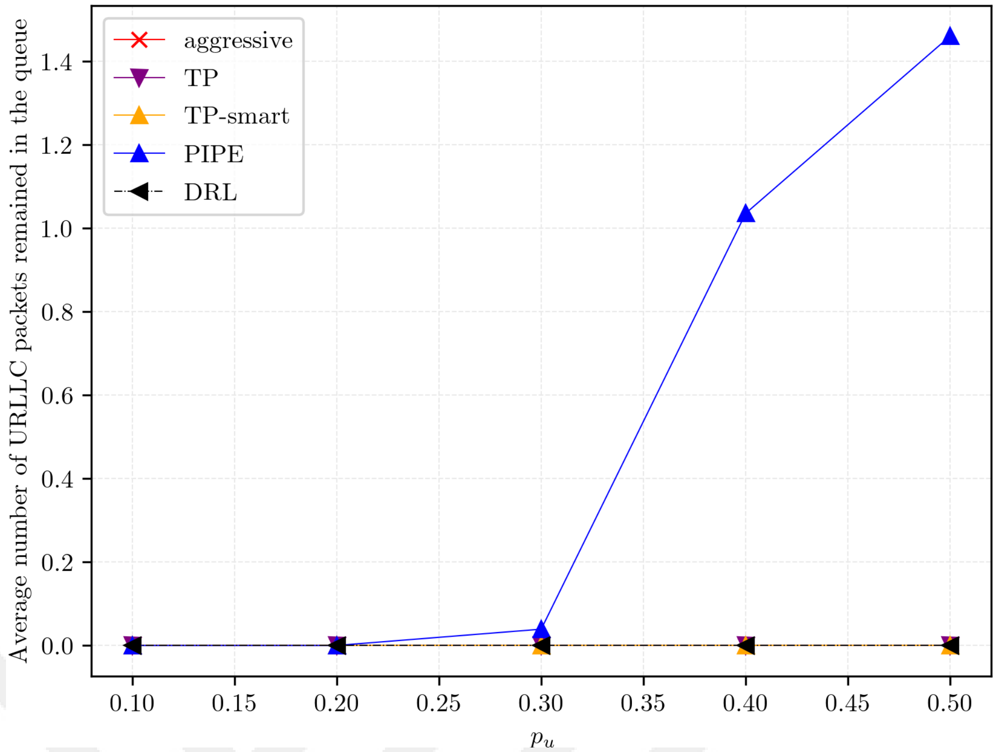

Figure 4.4: Average residual of algorithms

verify the minimal eMBB loss while meeting URLLC deadlines. However, this verification is limited since the simulator is not 3GPP compliant and lacks system dynamics. Moreover, this algorithm lacks the use of any channel quality information and buffer knowledge about URLLC transmissions, which can help further optimize the balance of eMBB throughput and URLLC latency. The realization of these drawbacks in this initial research directs us to develop an improved version of the algorithm, introduced in Chapter 4, which is tested with a 3GPP-compliant system simulator. This version includes a preemption mechanism that details all aspects and implements all radio access layers.

5. 3GPP-CONFORMANT RAN-LEVEL EVALUATION OF HEURISTIC PREEMPTIVE SCHEDULING ALGORITHM

The initial version lacked a fully defined system model, deviated from 3GPP standards, and did not include preemption indication information to notify UEs of preempted resources. In this expanded study, we present a well-defined system model that integrates the complete physical layer processing chain, including the Low-Density Parity-Check (LDPC) coder, enabling simulations that closely emulate real-world conditions. This enhancement not only improves the fidelity of our simulation results but also significantly enhances their practical applicability, transforming this work into a comprehensive system-level evaluation that surpasses the initial findings of the conference paper. The main contributions of this study are summarized as follows.

- A 3GPP-based preemption Mechanism for Coexistence: A novel 3GPP-compliant preemption mechanism that efficiently manages the coexistence of eMBB and URLLC slices is proposed. The method dynamically reallocates resources to ensure both slices coexist without compromising their respective QoS requirements, particularly in resource-constrained 5GB networks. Our preemption approach is applicable across various scheduling methods, including round-robin and priority-based scheduling. Furthermore, the proposed method can be seamlessly integrated into other techniques, such as proportional fair algorithms, hybrid scheduling, and semi-persistent scheduling (SPS).

- A Selective Preemption Strategy for Minimizing eMBB Throughput Degradation: A selective preemption strategy that targets low-priority radio resources for preemption is also proposed to minimize the detrimental impact on the eMBB throughput. This strategy ensures that the URLLC slice meets strict low-latency traffic demands while keeping the throughput degradation of the eMBB slice to a minimal level. Additionally, we propose a preemption indication message to be utilized to flush the buffers of punctured eMBB users. This feature considerably improves decoding performance by enhancing eMBB recovery and reducing the need for retransmissions.

- Low Computational Complexity and Low Control Overhead: Advanced scheduling techniques (e.g., SPS and grant-free methods) typically require either high computational complexity or high control channel overhead. The proposed computationally lightweight preemption mechanism achieves comparable or superior performance without the need for complex scheduling operations, making it highly scalable and efficient for practical deployment. A minimal signaling overhead for control channel information exchange plays a critical role in improving network efficiency with strict latency and throughput guarantees.

- Comprehensive System-Level Validation With Superior Performance: We validated the proposed method through extensive system-level simulations covering Layers 1, 2, and 3 of 5G networks. The simulation results imply that the preemption method significantly outperforms traditional scheduling and preemption methods (e.g., round-robin and priority-aware scheduling) in terms of effectively managing the trade-off between the latency of URLLC and the throughput of eMBB. Furthermore, the flexibility of the URLLC scheduling strategy ensures its applicability to a wide range of scheduling algorithms, including hybrid and grant-free approaches. This is an important feature for standardization purposes, where the method ensures future-ready scalability and adaptability for dynamic traffic management in 5GB networks.

5.1 System Model

For downlink multiplexing of eMBB and URLLC, three main solutions have emerged: first, transmitting indicator signals to identify preempted time or frequency resources accurately [84], [85], which adds overhead; second, employing blind classification without indicators to detect URLLC signals [86], improving bandwidth efficiency at the expense of more computational power; and third, combining a lightweight indicator with blind detection to locate URLLC signals [84].

The preemption mechanism in the 5G standard satisfies the low-latency requirement of URLLC data by allowing the URLLC data queue to be processed before the finalization of ongoing resource allocation [87]. However, multiple HARQ re-transmissions can affect the quality of experience of the eMBB users and result in throughput degradation. Hence, an efficient resource allocation algorithm is needed to optimize both slices' latency-reliability and throughput. These challenges highlight the need for innovative solutions to the resource management problem. Several strategies can be employed to identify the optimal eMBB resources for preemption.

Fig. 5.1 illustrates the interaction between different protocol layers of the UE and gNB in a 5G NR network, emphasizing the dynamic scheduling and preemption process. The UE monitors and reports its channel conditions and buffer status to the gNB, which are crucial inputs for the MAC scheduler at the gNB. The MAC layer incorporates HARQ for error correction, while the RLC (Radio Link Control) layer utilizes ARQ (Automatic Repeat reQuest) for reliable data transmission. Based on the information provided by the UE, the scheduler at the gNB allocates appropriate radio resources and may issue preemption indications, which dictate the priority of slice and resource reallocation as needed. This bidirectional exchange of control information over the air interface ensures efficient and adaptive resource management, optimizing the network's performance to meet varying QoS requirements in real-time.

5.2 Proposed Preemption Method

In this heuristic preemptive approach, eMBB and URLLC schedulers allocate downlink data resources in a 5GB OFDM system with multiple UEs having different slices in a single cell. Schedulers work concurrently, independently, using a slot and mini-slot basis allocation. Mini-slot resource allocation shown in Fig. 5.2. The eMBB scheduler allocates eMBB downlink data on possible resources per slot for high throughput. URLLC scheduler uses a preemption mechanism for 5G downlink data allocation to meet strict low-latency and reliability constraints. URLLC traffic usually has a latency shorter than the slot duration. eMBB scheduler allocates resources first, then URLLC scheduler finds appropriate resource blocks to puncture and replace with

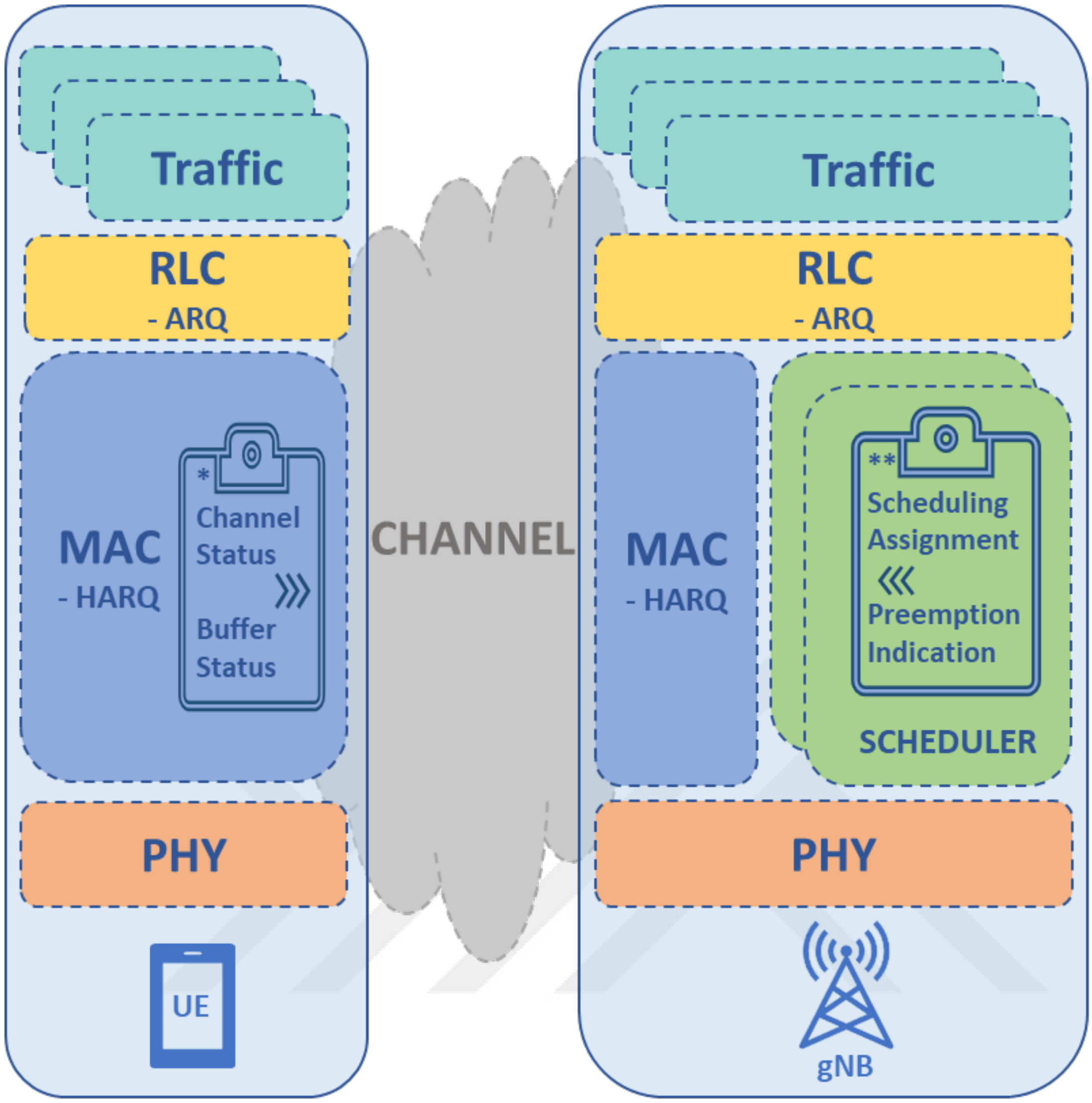

Figure 5.1: 3GPP-Conformant RAN-Level Evaluation

URLLC data. Study aims to minimize eMBB throughput reduction while meeting URLLC latency requirements.

I denotes the number of eMBB users, and N represents the number of URLLC users. Available bandwidth is $\mathscr{B}$ divided into RBs. The downlink traffic of eMBB slices is modeled under the full-buffer assumption, i.e, there is always data to be sent to eMBB users. CQI derived from real-time measurements, which is adapted from LTE networks and modified for 5G [3] as in Fig. 5.4. The average CQI of each user is determined based on the user's distance from the gNB. MCS is dynamically adjusted as follows:

$$\mathrm{MCS}_i = f_{\mathrm{eMBB}}(\mathrm{CQI}_i), \quad \forall i \in I, \tag{5.1}$$

Table 5.1: CQI tables of URLLC and eMBB slices.

	URLLC		eMBB	
CQI Index	Modulation	Code Rate x 1024	Modulation	Code Rate x 1024
0	out of range		out of range	
1	QPSK	30	QPSK	78
2	QPSK	50	QPSK	120
3	QPSK	78	QPSK	193
4	QPSK	120	QPSK	308
5	QPSK	193	QPSK	449
6	QPSK	308	QPSK	602
7	QPSK	449	16QAM	378
8	QPSK	602	16QAM	490
9	16QAM	378	16QAM	616
10	16QAM	490	64QAM	466
11	16QAM	616	64QAM	567
12	64QAM	466	64QAM	666
13	64QAM	567	64QAM	772
14	64QAM	666	64QAM	873
15	64QAM	772	64QAM	948

where $f_{\text{eMBB}}(\cdot)$ is CQI-to-MCS mapping function for eMBB traffic.

The downlink traffic of URLLC slices is generated according to real-time requirements as in [88]. MCS of URLLC users n is assigned based on CQI values in Table 5.1:

$$\text{MCS}_n = f_{\text{URLLC}}(\text{CQI}_n), \quad \forall n \in N, \tag{5.2}$$

where $f_{\text{URLLC}}(\cdot)$ is CQI-to-MCS mapping function for the URLLC traffic.

eMBB users are scheduled in round-robin (RR) manner in each time slot for fair resource allocation:

$$\text{RBG}_k(t) = \text{RR}(\mathscr{B}, k), \quad \forall k \in K, \tag{5.3}$$

where $\text{RR}(\cdot)$ is round-robin scheduler, $\text{RBG}_k(t)$ represents resource block groups assigned to user k in time slot t.

The URLLC scheduling process begins by identifying empty RBs remaining after eMBB scheduling. If any, these RBs are allocated to URLLC users. When insufficient free resources exist, the system initiates preemption of eMBB traffic, prioritizing RBGs

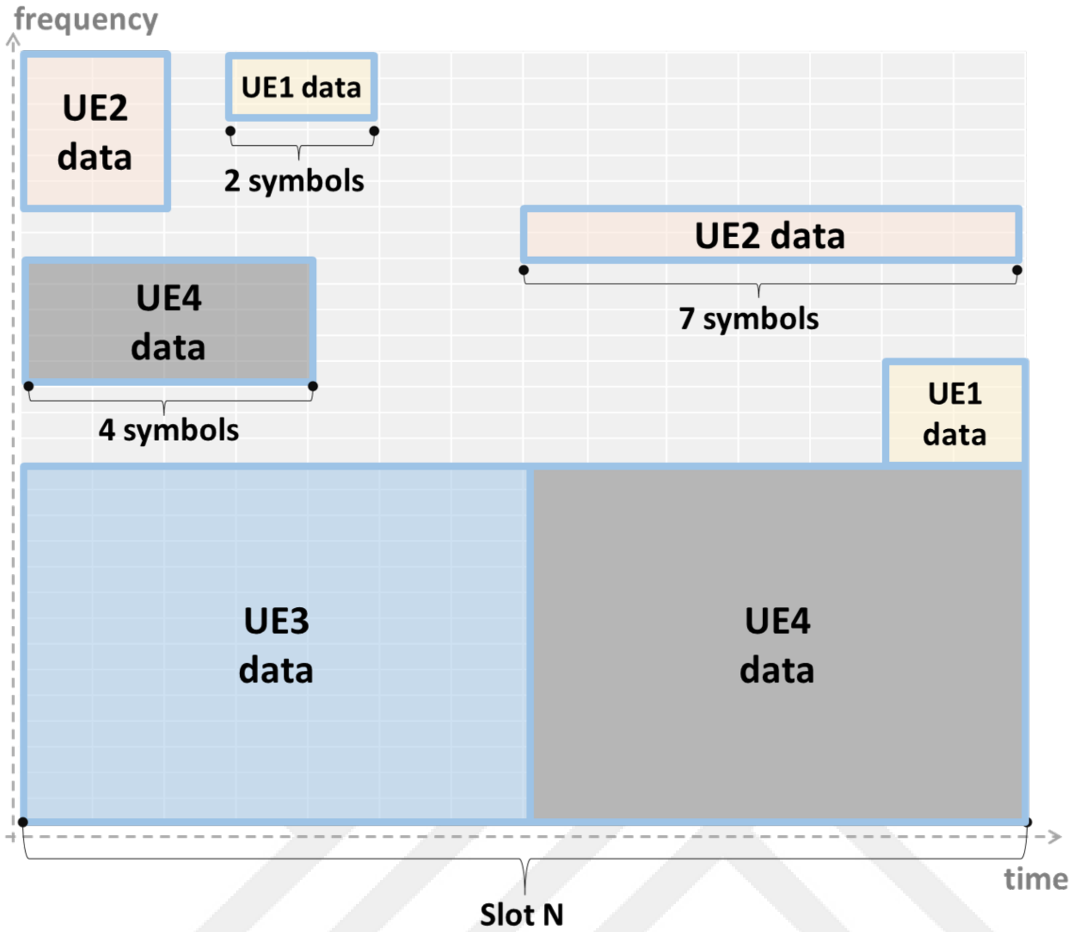

Figure 5.2: Mini-slot resource allocation example in 5G

where URLLC users report the highest Channel Quality Indicator (CQI) values. For any given URLLC user n, the set of RBGs with maximum CQI is determined through CSI reports as:

$$\mathscr{B}_n^{\text{high CQI}} = \left\{ rbg \in \mathscr{B} \mid \mathrm{CQI}_n(rbg) = \max_{rbg' \in \mathscr{B}} \mathrm{CQI}_n(rbg') \right\}. \tag{5.4}$$

The calculation of required resource blocks for URLLC traffic considers both the buffer status and transport block size (TBS), which depends on the CQI of preempted resource blocks. The number of necessary resource blocks for user n is derived from:

$$N_{\mathrm{RB},n} = \left\lceil \frac{\mathrm{BufferStatus}_n}{\mathrm{TBS}_{\text{per RB},n}} \right\rceil, \tag{5.5}$$

where the TBS per resource block follows the 3GPP 38.214 specification [7]:

$$\mathrm{TBS}_{\text{per RB},n} = 8 \times \left\lceil \frac{N'_{\text{info}} + 24}{8} \right\rceil - 24. \tag{5.6}$$

$$N'_{\text{info}} = \begin{cases} \max\left(24, 2^k \cdot \left\lceil \frac{N_{\text{info}}}{2^k} \right\rceil\right) & \text{for } N_{\text{info}} \leq 3824, \\ \max\left(3840, 2^k \cdot \text{round}\left(\frac{N_{\text{info}}-24}{2^k}\right)\right) & \text{otherwise.} \end{cases} \tag{5.7}$$

$$N_{\text{info}} = N_{\text{RE}} \times R \times Q_m \times v, \tag{5.8}$$

$$k = \begin{cases} \max\left(3, \lfloor \log_2(N_{\text{info}}) \rfloor - 6\right) & \text{for } N_{\text{info}} \leq 3824, \\ \lfloor \log_2(N_{\text{info}} - 24) \rfloor - 5 & \text{otherwise.} \end{cases} \tag{5.9}$$

The number of resource elements per RB is fixed at 12, while the coding rate R and modulation order Q_m (such as 2 for QPSK or 4 for 16-QAM) are determined through CQI tables. The number of transmission layers v and the TBS before rate matching N_{info} and after rate matching N'_{info} complete the set of parameters governing this process.

URLLC slices are characterized by higher priority over eMBB transmissions due to their stringent QoS requirements, reflected in lower 5QI values. When the system bandwidth is already allocated for eMBB slices, the system identifies eMBB users with minimal 5QI values:

$$U_{\text{lQ}} = \{k \in \{1, 2, \ldots, N\} \mid 5\text{QI}_k = \min_{j \in \{1,2,\ldots,N\}} 5\text{QI}_j\} \tag{5.10}$$

From this group, the algorithm selects resources where the affected eMBB user demonstrates the lowest Modulation and Coding Scheme (MCS) value:

$$rbg^* = \arg \min_{rbg \in \mathscr{B}_n^{\text{high CQI}}} \text{MCS}_k(rbg), \quad \text{for } k \in U_{\text{lQ}} \tag{5.11}$$

This approach targets eMBB transmissions with lower spectral efficiency but greater error resilience. LDPC coding's effectiveness at reduced code rates is used for this purpose. The system communicates these allocations through Downlink Control Information (DCI) messages. Special preemption indicators inform UEs about resource modifications, enabling adaptive reception strategies that may involve HARQ buffer management to minimize retransmission overhead.

The resource blocks preempted by the URLLC user n are selected among the resource blocks with the highest CQI values of the URLLC user n. From this set, the resource blocks having the lowest MCS values of eMBB users are preferred.

$$\mathscr{P}_n = \left\{ rbg^* \mid rbg^* \in \mathscr{B}_n^{\text{high CQI}}, \mathrm{MCS}_k(rbg^*) = \min(\mathrm{MCS}_k) \right\}. \tag{5.12}$$

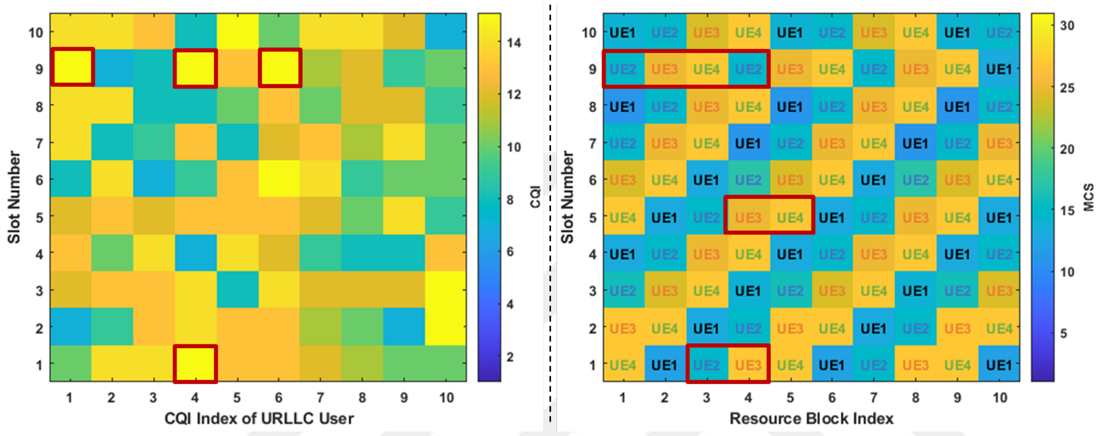

Figure 5.3: Scheduling model illustrating resource preemption and assignment

Fig. 5.3 illustrates an example snapshot (10 RBs, 10 slots) of resource preemption and block assignment in eMBB and URLLC slice coexistence. The left subplot shows CQI index for URLLC user across different resource blocks and slots. Higher CQI values (lighter colors) indicate blocks likely to be selected for URLLC transmission. URLLC packets can be scheduled in Slots 1 and 9 with the highest CQIs to ensure consistent resiliency highlighted by red rectangles on the left plot. The right subplot shows MCS assignments for eMBB users across resource blocks and slots, each block labeled for the respective eMBB user (UE1, UE2, UE3, UE4) and color-coded by MCS value. Red-outlined rectangles indicate preempted blocks. The resource blocks with the highest CQI reported from the URLLC user and the lowest MCS assigned for the scheduled eMBB users are prioritized for preemption. This selection process aims to minimize the reduction in overall eMBB throughput. The preempted packets can often be recovered with effective forward error correction mechanisms like LDPC for

5G PDSCH. Otherwise, the retransmissions are requested through HARQ feedback. Maximum achievable throughput (bps) in 5G NR, as defined in [89], calculated as:

$$\mathrm{Thr} = N_{\mathrm{l}} \cdot N_{\mathrm{PRB}} \cdot N_{\mathrm{sc}} \cdot N_{\mathrm{sym}} \cdot Q_m \cdot R \cdot \frac{1}{T_{\mathrm{slot}}} \cdot (1 - OH) \tag{5.13}$$

In our simulation scenarios, the number of spatial layers N_{l} is set to 1. Each transmission utilizes up to 52 Physical Resource Blocks (PRBs), denoted as N_{PRB}, with each PRB containing $N_{\mathrm{sc}} = 12$ subcarriers. The OFDM structure employs $N_{\mathrm{sym}} = 14$ symbols per slot when using normal cyclic prefix, which matches our simulation configuration. The modulation scheme, represented by Q_m, varies according to transmission requirements (with QPSK using 2 and 16-QAM using 4). The coding rate R, ranging between 0 and 1, is selected from appropriate MCS tables (Table 5.1). Time parameters include the slot duration T_{slot} fixed at 1ms for our numerology. Additionally, the calculation accounts for system overhead through the OH parameter, which typically assumes values between 0.14 and 0.18 depending on operational conditions.

Total throughput of a eMBB user after resource preemption can be expressed as:

$$\begin{aligned} \mathrm{Throughput}_{\mathrm{UE}} = P_{\mathrm{success}} \cdot N_{\mathrm{PRB,UE}} \cdot N_{\mathrm{sc}} \cdot N_{\mathrm{sym,UE}} \\ \cdot Q_m \cdot R_{\mathrm{UE}} \cdot \frac{1}{T_{\mathrm{slot}}} \cdot (1 - OH_{\mathrm{UE}}) \end{aligned} \tag{5.14}$$

where P_{success}, successful decoding probability after preemption, is function of coding rate R_{UE}, preempted resource block count $N_{\mathrm{preempted,UE}}$, and preemption position $P_{\mathrm{preemption,UE}}$ [90]. $N_{\mathrm{PRB}_{\mathrm{eMBB}}}$ denotes resource blocks allocated to eMBB users before preemption, $N_{\mathrm{preempted}}$ represents preempted resource blocks allocated to URLLC users [91].

$P_{\mathrm{success}}(R, N_{\mathrm{preempted}}, P_{\mathrm{preemption}})$ models the probability that an eMBB packet can be decoded. The lower coding rate R provides more redundancy and error correction [90], hence increases the successful decoding probability. The impact of preemption to throughput depends on the preempted resource block count $N_{\mathrm{preempted}}$, and whether preemption affects critical or non-critical packet portions $P_{\mathrm{preemption}}$ [91].

5.3 Simulation Results

The simulation environment models a 5G NR cell using MATLAB's 5G Toolbox supporting symbol-based scheduling in TDD mode. It includes PHY, MAC, RLC, and PDCP layer implementations. We enhanced the toolbox for our research objectives. The default toolbox lacks eMBB or URLLC slice definitions. We introduced this differentiation by assigning slot-based scheduling for eMBB traffic and mini-slot scheduling for URLLC traffic, following 5G NR specifications. The default configuration uses a single CQI table for eMBB traffic. We added a URLLC-specific CQI table for appropriate modulation and coding schemes. We implemented a preemption indication mechanism informing eMBB users when resources are preempted by URLLC transmissions, flushing soft buffers to reflect preemption impact per 3GPP standards. Different scheduling strategies and custom scheduler metrics were incorporated for realistic multi-service scenario modeling. Our simulation environment extends capabilities described in [92], enabling comprehensive scheduling strategy evaluation for eMBB and URLLC coexistence. We adopted the DDDSU slot pattern consisting of three downlink slots, one special slot (combination of downlink, guard period, and uplink), and one uplink slot [93]. This pattern balances downlink-dominant traffic and uplink feedback, suitable for eMBB/URLLC slices, commonly used worldwide, especially in Europe [94]. URLLC slice requirements based on [88], packet sizes 50-80 bytes within 100 MHz bandwidth (273 RBs). Since simulation uses 52 RBs bandwidth, URLLC packets range from 6 to 24 bytes. The code base supports multiple UEs with different traffic profiles, allowing diverse simulation setups. Channel parameters configured according 5G NR clustered delay line model, specifically CDL-C with 30 nanoseconds delay spread. Meeting stringent URLLC reliability requirements, the specialized MCS table ([7] Table 5.2.2.1-4) targets a block error rate of 0.00001, ensuring highly reliable downlink transmission through advanced link adaptation. Simulation includes a single gNB with 10 MHz bandwidth (52 Resource Blocks, 15 kHz subcarrier spacing) serving four eMBB users, one URLLC user. All users utilize the full channel bandwidth in a single Bandwidth Part. Simulation runs 3 seconds, covering 300 frames. We imported the CQI versus MCS mapping table from [3], showing LTE Rel-12 modulation under real-time traffic measurements. Observation ratios represent CQI reception percentages from overall UEs in a multi-site LTE network, illustrated in Fig. 5.4.

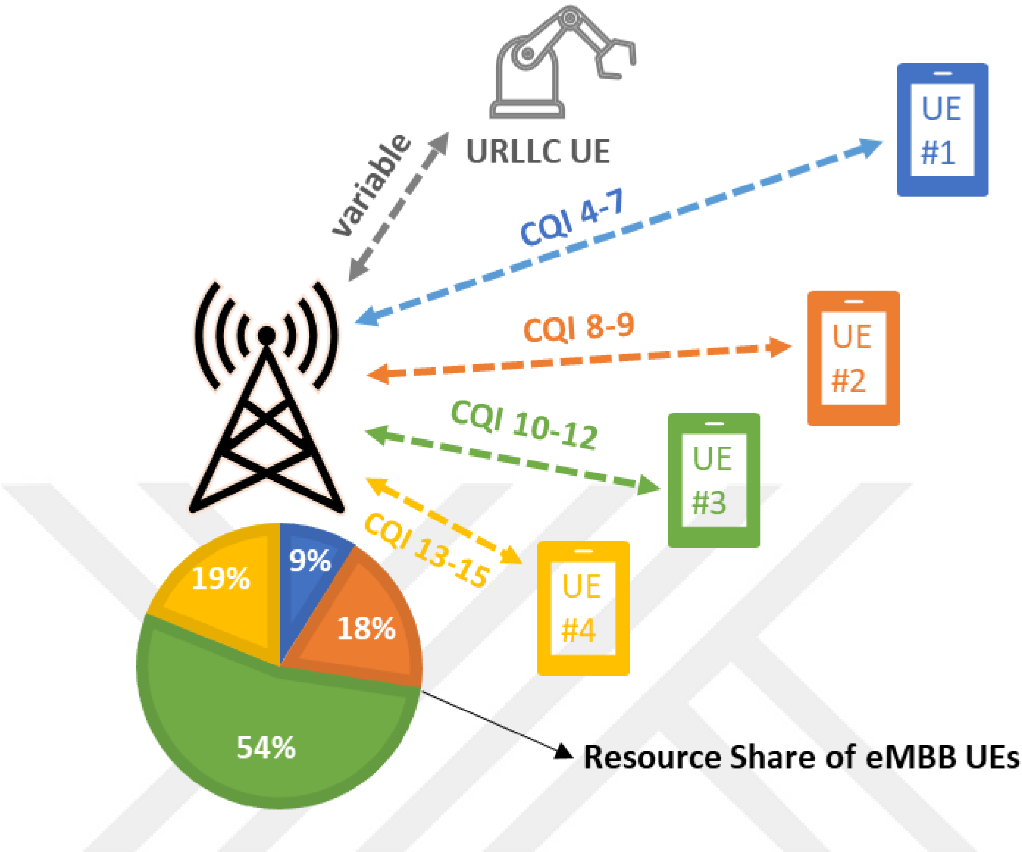

Figure 5.4: Resource share in simulations

Figure 5.5 presents CQI and MCS index histograms from simulations. It confirms the accuracy of CQI reporting for each user. Assigned MCS values confirm the similarity with the real-time scenarios.

Metrics include throughput, measuring average data rate for eMBB traffic; latency representing end-to-end delay experienced by URLLC traffic; and reliability quantified by packet error rate for URLLC traffic. Results provide insights into the system capability meeting stringent 5G slice requirements. We compare a preemptive scheduler with round-robin and priority-aware schedulers. The round-robin scheduler allocates resources cyclically among users without considering traffic priorities. The priority-aware scheduler prioritizes URLLC slices over eMBB slices and allocates resources as in the round-robin scheduler. Since eMBB/URLLC traffic is scheduled

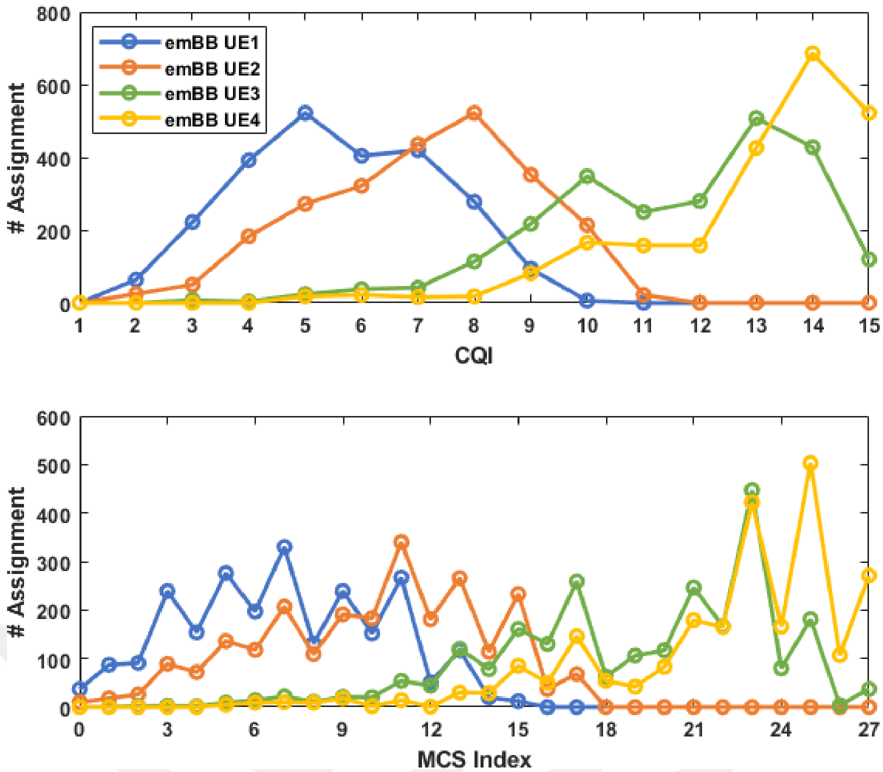

Figure 5.5: CQI reporting of eMBB users and MCS assignments

on a slot/mini-slot basis, respectively, scheduling URLLC first may leave unused resources.

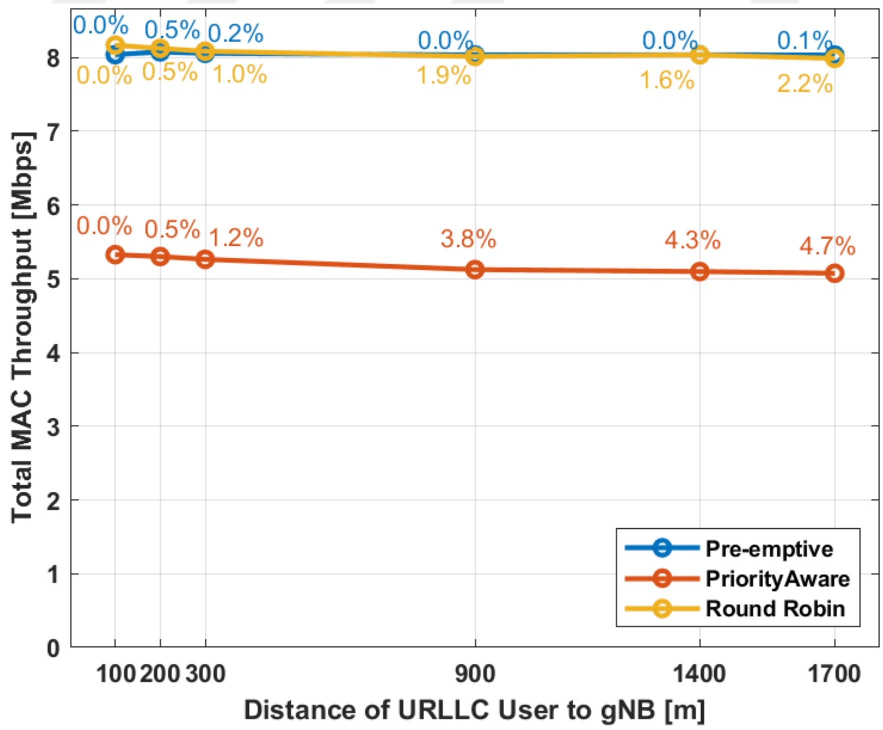

Figure 5.6: MAC throughput vs distance of the URLLC user to the gNB

Fig. 5.6 demonstrates total MAC throughput versus URLLC user distance from gNB. Percentages indicate the relative throughput decrease in each scheduler compared to the throughput when the URLLC user is 100 meters away from the gNB. These percentages show how increasing URLLC distance affects network performance under each scheduling algorithm. Results reveal preemptive scheduler maintains consistent throughput across all distances, due to effective preempted eMBB resource correction.

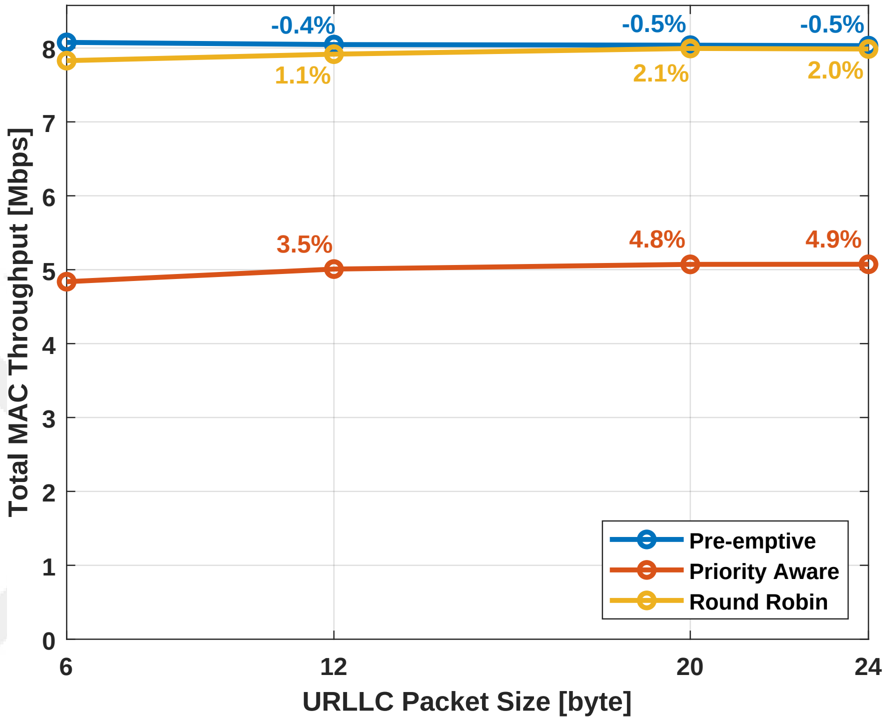

Figure 5.7: MAC throughput vs URLLC packet size

Fig. 5.7 shows total MAC throughput for the compared scheduling algorithms with different URLLC traffic packet sizes. Percentage throughput increase/decrease written at packet size points relative to packet size 16 for each algorithm (+/-). The total MAC throughput degrades slightly with increasing URLLC packet size, which verifies the proposed scheduler's robustness.

URLLC latency is significantly reduced to meet stringent low-latency requirements due to immediate resource reallocation enabled by preemption. Simulation assigns each URLLC transmission latency constraint randomly selected from a predefined set (0.15ms, 0.3ms, 0.5ms, 1ms with probabilities 0.1, 0.1, 0.5, 0.3). This ensures heterogeneous latency constraints reflecting real-world traffic conditions. Fig. 5.8 shows the latency distribution CDF with different scheduler types. As in Fig. 5.9, the preemptive scheduler ensures successful transmission for all URLLC transmissions due to the proper preemption decisions with assigned MCSs based on reported CQIs and the conservative MCS table specified for URLLC. Round-robin and priority-aware schedulers cause a subset of URLLC packets to extend the latency beyond the allowed threshold.

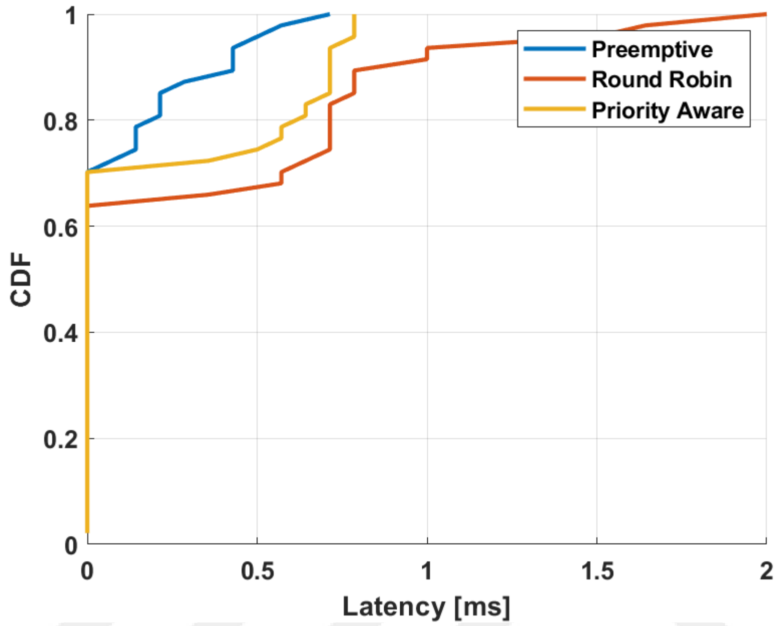

Figure 5.8: CDF of URLLC latency

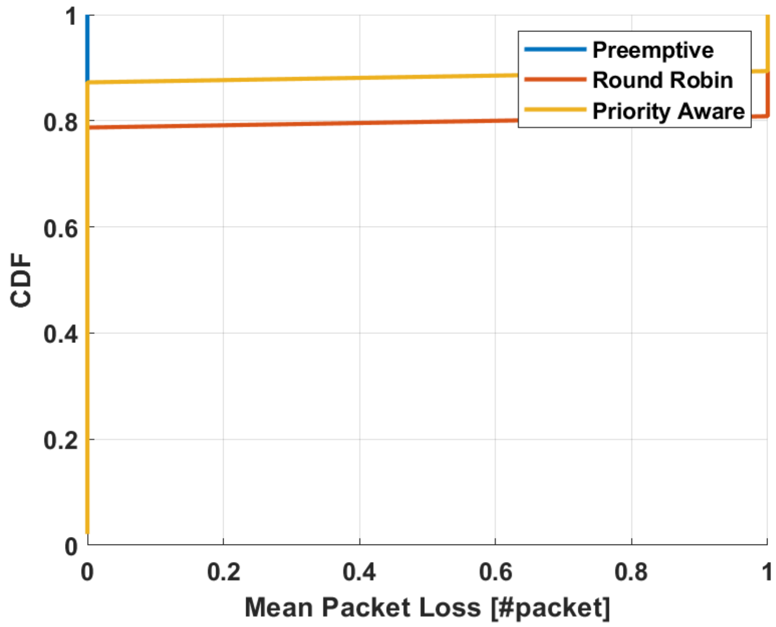

Figure 5.9: Mean packet loss of the URLLC user

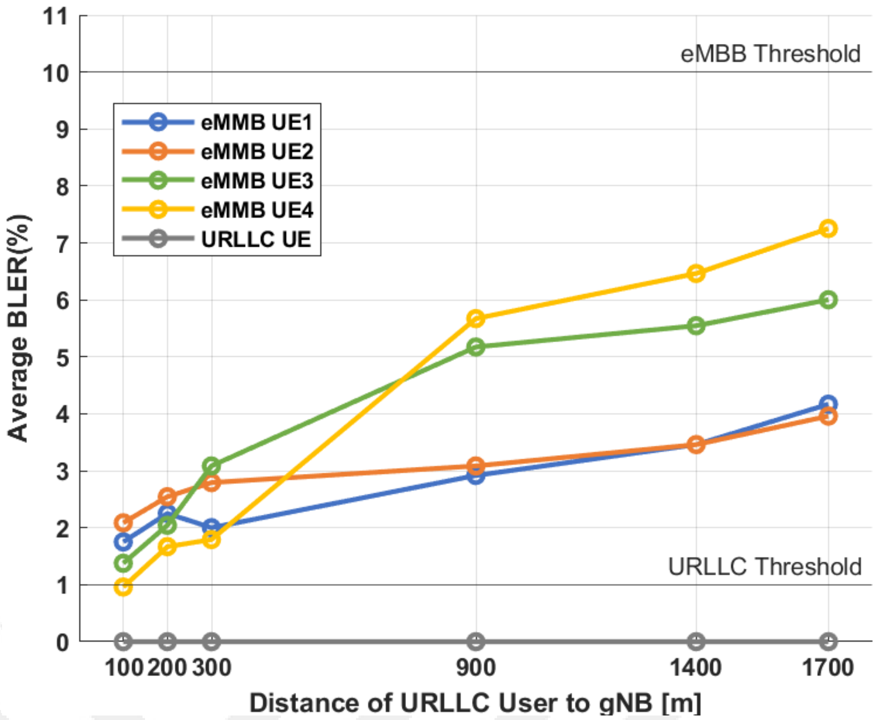

Figure 5.10: BLER(%) vs distance of the URLLC user to the gNB for the preemptive scheduler

Fig. 5.10 illustrates URLLC user distance impact on average BLER for both eMBB and URLLC users, showing a trade-off between maintaining URLLC reliability and eMBB performance. As the URLLC user distance increases, the BLER for all eMBB users progressively rises. This is expected due to increased resource allocation for URLLC users at larger distances, compensating for poorer signal quality. eMBB UE4, farthest from the gNB, has the poorest channel condition. The preemptive scheduling searches the lowest CQI resources to preempt, causing the highest BLER at all distances for UE4. Even with a 1700m distance, all eMBB and URLLC users remain below BLER limits.

5.4 Discussion

The proposed preemption-based scheduling mechanism offers a scalable solution for managing eMBB and URLLC slices in 5G networks across various real-world deployments such as autonomous vehicles, remote healthcare, and industrial automation. This approach dynamically reallocates resources with minimal throughput degradation, maintains low computational complexity with O(m·n) time complexity, and requires minimal control channel overhead compared to more complex scheduling methods.

The mechanism's strengths include its inherent scalability for high-density networks, effectiveness during network congestion, adaptability to multiple slice types beyond eMBB and URLLC, and energy efficiency through reduced retransmissions. While effective, the increasing complexity of network demands highlights the growing need for artificial intelligence to further optimize resource allocation decisions beyond what static algorithms can achieve. The current design allows network operators to meet stringent latency requirements without sacrificing throughput, yet future networks will require more intelligent, self-adapting solutions capable of predicting and responding to complex traffic patterns in real-time.

Building upon these foundations and addressing the need for more intelligent and adaptive resource allocation strategies, Chapter 6 introduces an advanced Context-Aware Multi-DQN-Driven Preemptive Scheduling paradigm for network slicing. This innovative approach leverages deep reinforcement learning with multiple Deep Q-Networks (DQNs) that simultaneously optimize various network performance metrics while dynamically adapting to changing network conditions. The AI-driven system continuously learns from network behaviors to make increasingly sophisticated preemption decisions, moving beyond static rule-based approaches to truly cognitive resource management. The chapter presents comprehensive system-level evaluations and discusses real-world deployment implications, examining how this AI-enhanced framework can support efficient network slicing and service differentiation by autonomously balancing competing slice requirements in 5G and beyond.

6. 3GPP-CONFORMANT RAN-LEVEL EVALUATION OF DQN-DRIVEN PREEMPTIVE SCHEDULING ALGORITHM

In 5G NR, CQI and packet size directly influence scheduling performance. CQI determines modulation and coding rate, impacting spectral efficiency, while URLLC packet size dictates latency constraints. Traditional single-model RL approaches fail to generalize across these variations, necessitating multiple DQN models specialized for different conditions.

Let $\mathrm{UE}^{\mathrm{eMBB}}$ and $\mathrm{UE}^{\mathrm{URLLC}}$ be the sets of eMBB and URLLC users. The throughput for an eMBB user i can be computed as:

$$\mathscr{T}_i^{\mathrm{eMBB}} = P_{\mathrm{success},i}(a) \cdot N_{\mathrm{RB},i} \cdot N_{\mathrm{sc}} \cdot N_{\mathrm{sym},i} \cdot Q_{m,i} \cdot C_i \cdot \frac{1}{T_{\mathrm{slot}}} \cdot (1 - OH_i) \tag{6.1}$$

where $P_{\mathrm{success},i}(a)$ represents the probability of successful decoding for the eMBB user i.

The resource allocation problem can be formulated as:

$$\begin{aligned} \underset{a \in \mathscr{A}}{\text{maximize}} \quad & \sum_{i \in \mathrm{UE}^{\mathrm{eMBB}}} \mathscr{T}_i^{\mathrm{eMBB}}, \\ \text{subject to} \quad & \tau_u^{\mathrm{URLLC}}(a) \leq \tau_{\text{u, max}}^{\mathrm{URLLC}}, \\ & \mathbb{P}_u^{\mathrm{URLLC}}(\text{success} \mid a) \geq (P_{\text{u, success}}^{\mathrm{URLLC}}), u \in \mathrm{UE}^{\mathrm{URLLC}}, \\ & a = (p,d) \in \mathscr{A} \\ & \mathscr{A} = \{(0,0)\} \cup \left\{ (p,d) \,\middle|\, p \in \mathscr{P}^{\mathrm{URLLC}}, d \in \mathscr{D}^{\mathrm{URLLC}} \right\} \end{aligned} \tag{6.2}$$

The set $\mathscr{P}^{\mathrm{URLLC}}$ represents all feasible preemption patterns:

$$\mathscr{P}^{\mathrm{URLLC}} = \bigcup_{n \in \{2,4,7\}} \bigcup_{0 \leq s \leq N_{\mathrm{SYM}} - n} (s, s+1, \ldots, s+n) \tag{6.3}$$

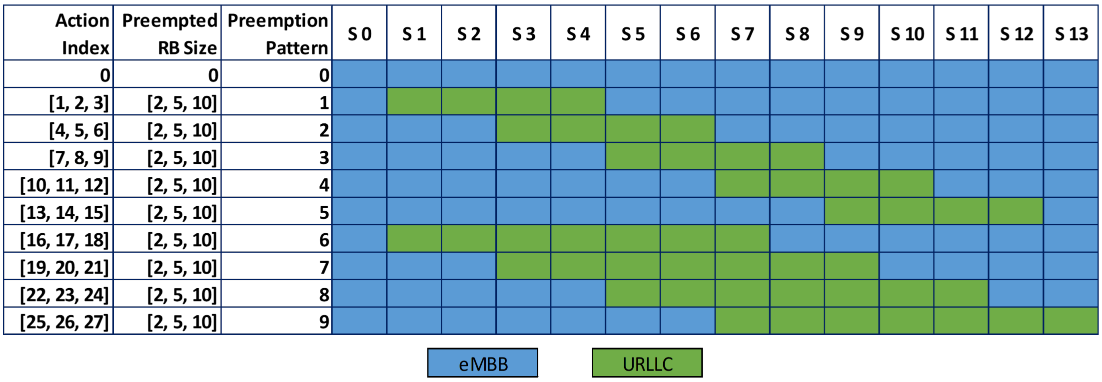

Figure 6.1: Simplified version of 3GPP-defined URLLC preemption patterns

where $N_{\text{SYM}} = 14$ denotes the total number of OFDM symbols in a slot. The total number of feasible preemption patterns is:

$$|\mathscr{P}^{\text{URLLC}}| = \sum_{n \in \{2,4,7\}} (N_{\text{SYM}} - n + 1) \tag{6.4}$$

The total number of possible subsets of resource blocks is:

$$|\mathscr{D}^{\text{URLLC}}| = \sum_{k=1}^{N_{\text{RB}}} \binom{N_{\text{RB}}}{k} = 2^{N_{\text{RB}}} - 1 \tag{6.5}$$

For a 10 MHz bandwidth ($N_{\text{RB}} = 52$), the action space size becomes approximately 1.44×10^{17}, making exhaustive search intractable. To simplify, we constrain the number of preempted RBs to $\{2, 5, 10\}$ and propose a streamlined action space:

$$\mathscr{A} = \underbrace{\{(0,0)\}}_{\text{No preemption}} \cup \left\{ (p,d) \,\middle|\, \underbrace{p \in \{1,\ldots,9\}}_{\text{Preemption patterns}}, \underbrace{d \in \{2,5,10\}}_{\text{\# preempted RB}} \right\} \tag{6.6}$$

This formulation encapsulates 28 discrete preemption strategies compatible with mini-slot-based scheduling mechanisms in 5G NR, enabling efficient multiplexing of URLLC traffic with ongoing eMBB transmissions while minimizing disruption.

Our method uses multiple DQN models trained for specific CQI levels and URLLC packet sizes, enabling adaptive and fine-grained preemptive scheduling—unlike single-model RL approaches that struggle with CQI variability. This design optimizes spectral efficiency while balancing URLLC latency and eMBB throughput, grounded

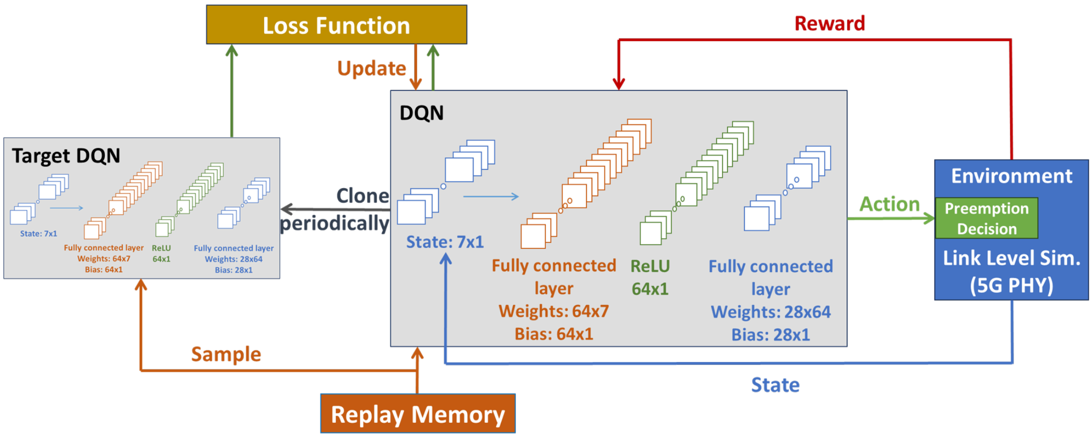

Figure 6.2: DQN Architecture for Training Phase

in 3GPP findings showing quasi-deterministic URLLC packet sizes (6–24 bytes) in industrial TSN use cases [88,95].

Training is performed using a MATLAB-based 5G NR link-level simulator, with DQN models trained per wideband CQI and packet size. These models are integrated into a full-stack MATLAB NR simulator, combining accurate physical-layer modeling with network-level validation.

The DQN architecture employs dual networks for stabilized learning: an online network refines policies via gradient descent, while a target network ensures stability through soft updates. The hidden layers, using ReLU activations, process a 7-dimensional state space—capturing URLLC buffer status, latency deadlines, CQI, and eMBB user allocations and channel qualities.

The agent outputs a 28-dimensional hybrid resource allocation action, adhering to 3GPP standards, and updates its Q-function via:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \beta \left[r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t) \right] \tag{6.7}$$

The reward function balances URLLC latency and eMBB throughput:

$$R = \min\left(1, \frac{\tau_{\text{remaining}}}{\tau_{\max}^{\text{URLLC}}}\right) - \frac{\Delta T^{\text{eMBB}}}{T_{\max}^{\text{eMBB}}} \tag{6.8}$$

This formulation encourages latency compliance while minimizing eMBB degradation, with reward clipping in $[-1,1]$ to ensure training stability. The proposed framework trains Context-Aware DQN policies across all 15 standardized CQI levels and four URLLC packet sizes (6–24 bytes), then validates them in a full-stack 5G NR simulator featuring PHY/MAC/RLC/PDCP modeling, including HARQ, RLC segmentation, and OFDM waveform generation.

Each state $s_t = [B_t, L_t, \mathrm{CQI}_t, \mathrm{RB}^{\mathrm{eMBB}}]$ captures real-time URLLC buffer levels, latency budgets, CQI, and eMBB allocations. The training employs 3GPP-compliant channel models (TS 38.214), maps SINR to CQI, and uses an ε-greedy policy with linear decay to balance exploration.

Actions yield URLLC transport block sizes via link-level simulation, and rewards are computed using the latency-throughput tradeoff in (6.8). Double DQN with Adam optimization minimizes the TD loss:

$$\mathscr{L}(\theta) = \mathbb{E}(s,a,r,s') \sim \mathscr{B}\left[\left(r + \gamma Q\theta^{-}(s', \arg\max_{a'} Q_\theta(s',a')) - Q_\theta(s,a)\right)^2\right] \tag{6.9}$$

with periodic target network updates every $C = 1000$ episodes. This setup ensures stable training and joint URLLC-eMBB optimization under realistic 5G conditions.

Operational deployment uses a dynamic model selection approach, where the gNB's cross-layer controller selects the appropriate DQN model in real time based on reported CQI and URLLC packet size. This allows swift adaptation to changing channel conditions while remaining compatible with existing scheduling schemes.

The proposed preemption strategy schedules URLLC packets within ongoing eMBB transmissions by selectively preempting only essential symbols and resource blocks, minimizing eMBB throughput loss while meeting URLLC's strict latency demands. Fine-grained symbol patterns and adjustable RB densities enable a flexible trade-off between spectral efficiency and transmission reliability. This adaptive approach ensures efficient resource use while addressing diverse QoS needs in 5G networks. The study focuses on single-cell scenarios without inter-cell interference or coordination.

Algorithm 1 5GB PHY-Aware CA-M-DQN Training Protocol

Parameters: $T_{\max} = 30{,}000$, $M = 32$, $\gamma = 0.99$, $C = 1000$, $\alpha = 0.0001$, $\varepsilon = 1.0$
Parameters: Replay buffer $\mathscr{B}[10^5]$, DQN with weights θ, target network $\theta^- \leftarrow \theta$

Sample initial URLLC CQI from $\mathcal{N}(\mathrm{CQI}_0, \sigma^2)$ truncated to $\{1, \ldots, 15\}$
Generate URLLC traffic bursts
for episode $= 1$ to $T_{\max}$ **do**
  Sample eMBB CQI from $\mathscr{U}\{1, 15\}$
  Map CQI $\rightarrow$ SINR $\rightarrow$ MCS/modulation
  Construct state $s_t = [B_t, L_t, \mathrm{CQI}_t, \mathrm{RB}^{\mathrm{eMBB}}, \mathrm{MCS}]$
  if rand() $< \varepsilon$ **then**
    Select random action a_t
  else
    $a_t \leftarrow \arg\max_a Q(s_t, a; \theta)$
  end if
  $\varepsilon \leftarrow \max(0.1, 1 - \frac{\text{episode}}{0.8 T_{\max}})$
  Compute TBS η_t using Eq. (6)
  Execute a_t in simulator $\rightarrow$ observe s_{t+1}, r_t
  Store (s_t, a_t, r_t, s_{t+1}) in $\mathscr{B}$
  if $|\mathscr{B}| \geq M$ **then**
    Sample M transitions from $\mathscr{B}$
    $y_j \leftarrow r_j + \gamma \max_{a'} Q(s_{j+1}, a'; \theta^-)$
    Update θ using:
  $\mathscr{L} = \frac{1}{M} \sum_j (y_j - Q(s_j, a_j; \theta))^2$
  end if
  if episode mod $C = 0$ **then**
    $\theta^- \leftarrow \theta$
    Evaluate policy on validation set
  end if
end for
return Trained weights θ

Although real-time deployment is out of scope, the framework is compatible with commercial gNBs, and future validation on SDR or FPGA platforms will explore hardware constraints.

The proposed scheduling strategy is implemented in a MATLAB-based 5G system-level simulator supporting mini-slot configurations (2, 4, 7 symbols), preemption indication, a preemption-aware HARQ with soft buffer flushing, and a DQN-based preemption policy. The simulator models a full 5G RAN stack with 3GPP-compliant signaling. While hardware testing is future work, the setup reflects practical gNB constraints. We evaluate URLLC and eMBB throughput and BLER against two baselines: Preemptive [83], which uses heuristic decisions, and

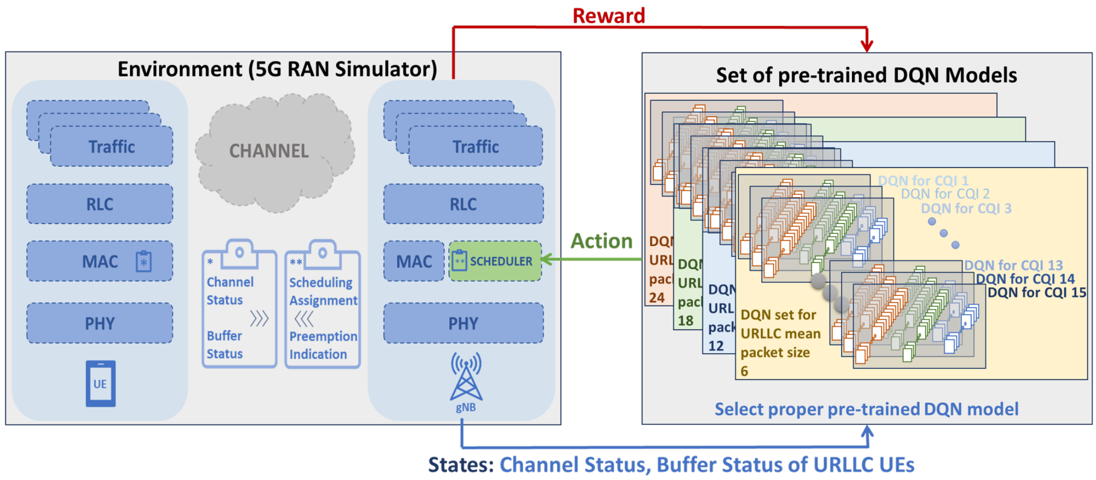

Figure 6.3: DQN Architecture for Test Phase

Priority-Aware, a round-robin variant that prioritizes URLLC. All strategies meet URLLC latency targets; however, CA-M-DQN achieves lower eMBB throughput loss, indicating improved spectral efficiency.

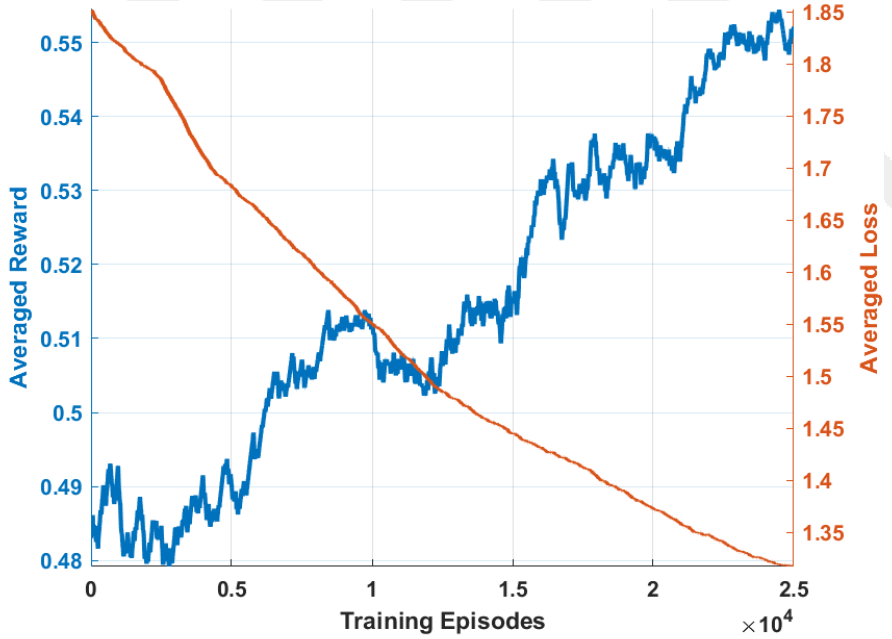

Figure 6.4: Averaged loss and reward history over training episodes

Fig. 6.4 shows training dynamics. The loss curve steadily declines, indicating convergence, while the reward increases over time, validating CA-M-DQN's policy improvement and alignment with prior DQN studies [96]–[98].

As shown in Fig. 6.5, CA-M-DQN achieves higher eMBB throughput compared to baseline methods. Unlike Preemptive and Priority-Aware strategies, which rely on

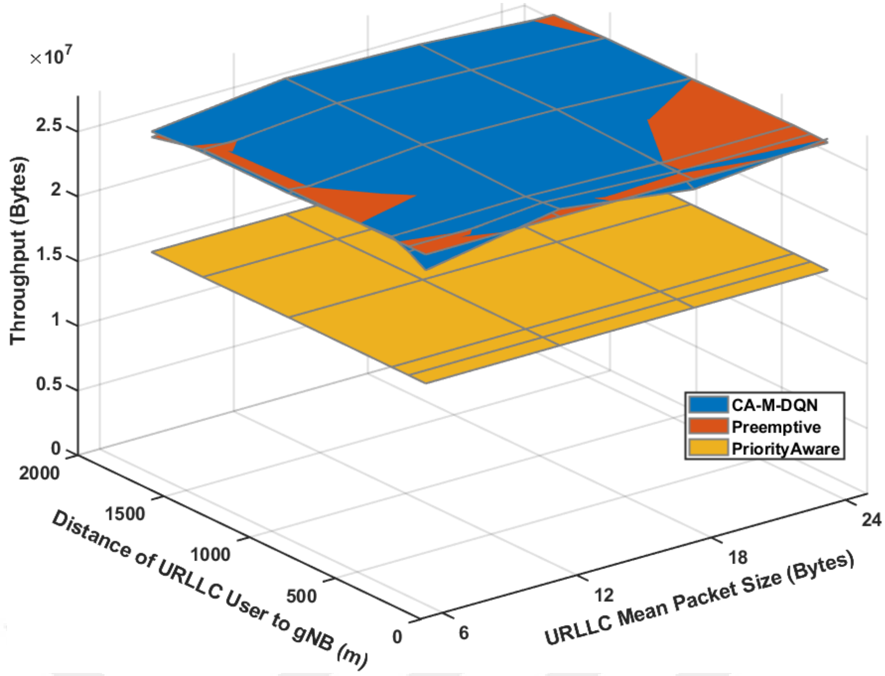

Figure 6.5: Comparison of eMBB throughput under different scheduling strategies

fixed rules, CA-M-DQN dynamically learns optimal preemption, effectively balancing URLLC reliability with eMBB performance.

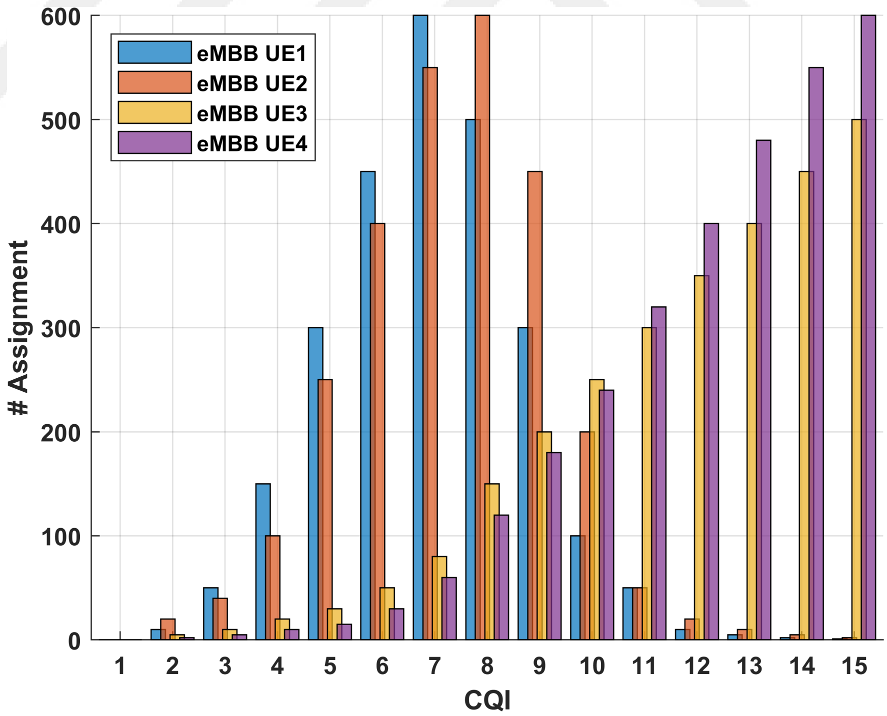

Figure 6.6: Histogram of the CQI values of eMBB users in simulations based on real-time observations [3]

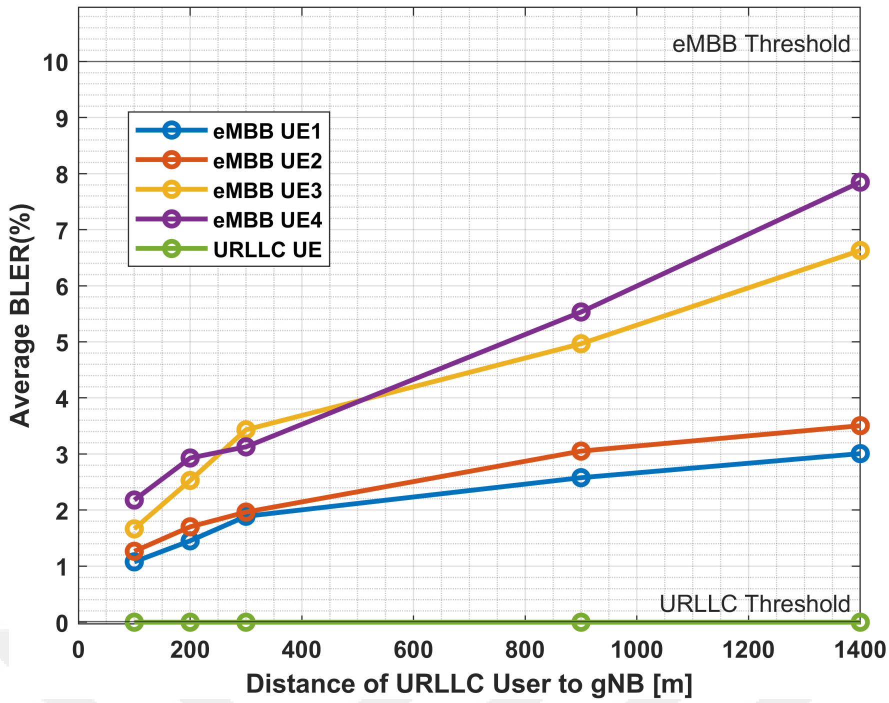

Figure 6.7: Relationship between distance of the URLLC user to the gNB and average BLER for eMBB and URLLC users

Fig. 6.6 shows the CQI distribution across eMBB users, while Fig. 6.7 illustrates increasing BLER with URLLC user distance. The system satisfies 3GPP TS 38.214 BLER targets ($\leq 1\%$ for URLLC, $\leq 10\%$ for eMBB), using distinct MCS tables for each service. The scheduler maintains URLLC reliability without significant degradation in eMBB performance, proving its robustness in dynamic environments.

6.1 Discussion

The proposed Context-Aware Multi-DQN Scheduling (CA-M-DQN) framework offers a novel approach to preemptive scheduling in 5G network slicing, balancing URLLC latency and eMBB throughput by dynamically selecting specialized DQN models based on CQI and packet sizes. This modular design improves decision accuracy and reduces training complexity, with potential for future extensions to mMTC. While maintaining multiple models increases computation, techniques such as pruning, FPGA acceleration, and quantization enable sub-10ms inference without major architectural changes.

Though this study focuses on single-cell scheduling, extending to multi-cell scenarios is essential for real-world deployment. Federated knowledge distillation can support scalable, efficient model sharing across gNBs. To reduce real-time overhead, the framework leverages dynamic model provisioning—caching key models locally and retrieving others on demand. Given URLLC's predictable traffic patterns, pre-trained models can be deployed edge-side, blending offline training with responsive, edge-compatible inference.

Energy consumption, though not evaluated, remains a key concern. Future work could explore energy-aware scheduling, integrating power-efficient reward terms and aligning with green AI principles for battery-constrained deployments. As 6G emerges, CA-M-DQN's adaptive, modular design can support advanced use cases like tactile internet, digital twins, and xURLLC by prioritizing low-latency traffic and adapting to bursty demands in dense environments.

7. CONCLUSIONS AND FUTURE WORK

Dynamic resource allocation methods in 5GB networks for scenarios where eMBB and URLLC slices coexist are explored in this thesis. These slices have distinct latency, throughput, and spectral efficiency requirements. The preemption mechanism, standardized by 3GPP, offers a solution to compensate for the trade-off between the low-latency needs of URLLC slices and the high-throughput requests of eMBB slices. Three preemptive scheduling approaches are introduced to answer these competing trade-offs of the two main slices efficiently throughout this thesis. While the first two investigations heuristically dealt with optimizing scheduling using preemption, the third investigation demonstrated the potential of AI-driven approaches in enhancing resource allocation efficiency. The first approach is validated through a physical layer simulator and highlights the need for a systematic verification. Consequently, the next two approaches are validated through a 5G RAN system simulator, extended with standard-compatible implementations such as priority definitions of slices, preemption mechanism with indication messages, physical layer adaptability, and HARQ compatibility. These validations demonstrate the readiness of the proposed algorithms to be used in real-world 5GB schedulers. Moreover, extensive simulations validate the effectiveness of such approaches, showing their potential for optimizing radio resource management in future wireless networks.

In the first approach, presented in Chapter 4, PIPE, a heuristic preemptive scheduler algorithm for coexisting eMBB and URLLC slices is introduced. This framework employs codeblock group-based transmission and a preemption mechanism compliant with 5G NR standards, ensuring compatibility with future communication systems. Unlike traditional heuristic approaches, PIPE leverages a puncture allowance rate to determine the maximum number of preemptions allowed per eMBB codeblock without triggering retransmissions, thereby balancing eMBB throughput preservation with URLLC latency guarantees. By prioritizing resource blocks with lower modulation orders (derived from real-time CQI observations) and assigning users based on empirically observed ratios, PIPE minimizes eMBB throughput degradation while satisfying URLLC deadlines. Simulation results demonstrate that this approach achieves superior performance in maintaining eMBB throughput with reductions of less than 5% under varying URLLC traffic loads, while fully meeting URLLC latency constraints.

In the second approach, detailed in Chapter 5, a 3GPP-compliant preemption mechanism for managing the coexistence of eMBB and URLLC slices in multi-service 5G and beyond networks is proposed. Unlike prior approaches, this framework employs a selective preemption strategy that dynamically reallocates low-priority eMBB resources, such as those with low modulation orders or high error-correction redundancy, to satisfy URLLC latency constraints while minimizing eMBB throughput degradation. This mechanism integrates the preemption indication to notify affected eMBB users, enabling efficient buffer management and reducing retransmissions. System-level simulations across PHY, MAC, and RLC layers demonstrated that our approach outperforms traditional schedulers, such as round-robin and priority-aware, by maintaining URLLC latency below strict thresholds and ensuring reliable eMBB performance under varying network conditions. The framework's low computational complexity and compatibility with diverse scheduling algorithms, including hybrid and grant-free approaches, support scalability in large-scale deployments. Furthermore, its design supports seamless extension to future slice types, such as mMTC and holographic communications, and congested environments through adaptive resource prioritization. Future work will focus on multi-cell coordination and energy efficiency optimizations to enhance real-world applicability in next-generation networks.

In the third approach, detailed in Chapter 6, the Context-Aware Multi-DQN Scheduling (CA-M-DQN) framework is described. This approach leverages multiple pre-trained Deep Q-Network (DQN) models, each specialized for distinct Channel Quality Indicator (CQI) levels and URLLC packet sizes, enabling dynamic adaptation to real-time network conditions. By integrating residual latency budgets and instantaneous channel quality metrics, the CA-M-DQN scheduler optimizes resource allocation decisions to minimize eMBB throughput degradation while strictly satisfying URLLC latency constraints. Simulation results demonstrate that this modular approach outperforms traditional preemption techniques, achieving a robust balance between URLLC reliability and eMBB performance, even under fluctuating traffic loads.

However, the framework currently faces limitations such as computational overhead from multi-agent coordination and reliance on simulation-based validation. Future work will extend the CA-M-DQN framework to support additional network slices (e.g., massive IoT, industrial automation), optimize its energy efficiency, and validate its efficacy in real-world deployments using Software-Defined Radio (SDR) testbeds and digital twin platforms. These advancements will further solidify its applicability in enabling scalable, intelligent resource management for emerging use cases such as autonomous vehicles and holographic communications.

Future research can expand upon this work by integrating multi-agent reinforcement learning (MARL) techniques, where multiple base stations collaboratively optimize resource allocation. Unlike single-agent approaches, MARL enables coordinated decision-making across distributed network entities, improving spectrum efficiency and reducing inter-cell interference. Additionally, investigating real-world deployment scenarios through hardware-in-the-loop (HIL) simulations or software-defined radio (SDR) testbeds would provide deeper insights into the practical feasibility of AI-driven scheduling mechanisms. Validating the proposed framework in live 5G test networks will also help identify system-level constraints such as inference latency, memory overhead, and implementation complexity on commercial baseband platforms.

Another direction of future interest is to use the framework to support additional types of slices, including massive Machine-Type Communications (mMTC) and extreme Ultra-Reliable Low-Latency Communications (xURLLC). These new slices present unique challenges with bursty traffic patterns and sub-millisecond latency requirements. Thus, they require more adaptive and more granular scheduling schemes. Using methods like neural architecture search (NAS) and meta-learning, AI models can adapt individually to the changing network conditions.

Finally, transitioning from simulation-based testing to real-world deployment will require optimizations both in the algorithmic and hardware domains. Techniques such as quantization-aware training, FPGA-based inference acceleration, and light-weight model pruning can potentially reduce computational overhead by orders of magnitude without compromising scheduling efficiency. Deploying the framework in commercial 5G hardware or digital twin environments will demonstrate its value and identify potential areas for future network slicing architectures to improve.

In conclusion, this thesis shows how heuristic and AI-driven resource management models have contributed to 5G scheduling optimization. The PIPE scheduler effectively minimizes eMBB throughput loss while ensuring URLLC latency requirements which is further enhanced with CQI awareness and tested with system-wide metrics. The CA-M-DQN approach further capitalizes on such advances using AI to learn dynamic scheduling choices in real-time based on up-to-date network conditions, maximizing spectral efficiency and QoS. While simulations prove the advantages of these approaches, research in the future should focus on validation in actual field scenarios using SDR testbeds and digital twin platforms to bridge the gap between theoretical advances and real-world deployment. These findings inform the innovation of future wireless technologies to support scalable, intelligent resource management in beyond-5G and 6G networks.

REFERENCES

[1] **3GPP** (2024). 5G; Service Requirements for the 5G System (Stage 1), Technical Report TS 122 261, ETSI, https://www.3gpp.org/ftp/Specs/archive/22_series/22.261/, version 18.13.0.

[2] **Zhang, H.**, **Feng, M.**, **Long, K.**, **Karagiannidis, G.K. and Nallanathan, A.** (2019). Artificial Intelligence-Based Resource Allocation in Ultradense Networks: Applying Event-Triggered Q-Learning Algorithms, *IEEE Vehicular Technology Magazine*, *14*(4), 56–63.

[3] **Sezgin, G.**, **Coskun, Y.**, **Basar, E. and Karabulut Kurt, G.** (2018). Performance Evaluation of a Live Multi-Site LTE Network, *IEEE Access*, *6*, 49690–49704.

[4] **3GPP** (2024). NR; Overall Description; Stage-2, Technical Report TS 38.300, 3GPP, https://www.3gpp.org/ftp/Specs/archive/38_series/38.300/.

[5] **3GPP** (2024). NR; Physical Channels and Modulation, Technical Report TS 38.211, 3rd Generation Partnership Project (3GPP), https://www.3gpp.org/ftp/Specs/archive/38_series/38.211/.

[6] **3GPP** (2024). NR; Physical Layer Procedures for Control, Technical Report TS 38.213, 3GPP, https://www.3gpp.org/ftp/Specs/archive/38_series/38.213/.

[7] **3GPP** (2024). NR; Physical Layer Procedures for Data, Technical Report TS 38.214, 3GPP.

[8] **3GPP** (2024). NR; Physical Layer Measurements, Technical Report TS 38.215, 3GPP, https://www.3gpp.org/ftp/Specs/archive/38_series/38.215/.

[9] **IEEE** (2004). IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems, *IEEE Std 802.16*.

[10] **Dahlman, E.**, **Parkvall, S. and Skold, J.** (2018). *5G NR: The Next Generation Wireless Access Technology*, 1 edition.

[11] **3GPP** (2024). NR; Multiplexing and Channel Coding, Technical Report TS 38.212, 3GPP, https://www.3gpp.org/ftp/Specs/archive/38_series/38.212/.

[12] **Lin, X.** *et al.* (2019). 5G New Radio: Unveiling the Essentials of the Next Generation Wireless Access Technology, *IEEE Communications Standards Magazine*, *3*(3), 30–37.

[13] **3GPP** (2024). NR; Radio Resource Control (RRC); Protocol Specification, Technical Report TS 38.331, 3GPP, `https://www.3gpp.org/ftp/Specs/archive/38_series/38.331/`.

[14] **Abdelsadek, M.Y., Gadallah, Y. and Ahmed, M.H.** (2020). Resource Allocation of URLLC and eMBB Mixed Traffic in 5G Networks: A Deep Learning Approach, *GLOBECOM 2020*, Taipei, Taiwan.

[15] **Zhou, Y., Tang, F., Kawamoto, Y. and Kato, N.** (2020). Reinforcement Learning-Based Radio Resource Control in 5G Vehicular Network, *IEEE Wireless Communications Letters*, *9*(5), 611–614.

[16] **Sutton, R.S. and Barto, A.G.** (2018). *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA, 2 edition.

[17] **Goodfellow, I., Bengio, Y. and Courville, A.** (2016). *Deep Learning*, MIT Press, Cambridge, MA.

[18] **Bishop, C.M.** (2006). *Pattern Recognition and Machine Learning*, Springer, New York.

[19] **Luong, N.C., Hoang, D.T., Gong, S., Niyato, D., Wang, P., Liang, Y.C. and Kim, D.I.** (2019). Applications of Deep Reinforcement Learning in Communications and Networking: A Survey, *IEEE Communications Surveys & Tutorials*, *21*(4), 3133–3174.

[20] **She, C., Sun, C., Gu, Z., Li, Y., Yang, C., Poor, H.V. and Vucetic, B.** (2021). Deep Reinforcement Learning for Mode Selection and Resource Allocation in Device-to-Device Communication, *IEEE Transactions on Wireless Communications*, *20*(3), 1865–1881.

[21] **Puterman, M.L.** (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, New York.

[22] **Watkins, C.J.C.H.** (1989). *Learning from Delayed Rewards*, King's College, Cambridge, UK, `http://www.cs.rhul.ac.uk/~chrisw/new_thesis.pdf`.

[23] **Bellman, R.** (1957). *Dynamic Programming*, Princeton University Press, Princeton, NJ.

[24] **Watkins, C.J.C.H. and Dayan, P.** (1992). Q-learning, *Machine Learning*, *8*(3), 279–292.

[25] **Sun, Y., Peng, M., Zhou, Y., Huang, Y. and Mao, S.** (2019). Application of Reinforcement Learning in Wireless Resource Allocation, *IEEE Wireless Communications*, *26*(4), 30–36.

[26] **Mnih, V.**, **Kavukcuoglu, K.**, **Silver, D.**, **Rusu, A.A.**, **Veness, J.**, **Bellemare, M.G.**, **Graves, A.**, **Riedmiller, M.**, **Fidjeland, A.K.**, **Ostrovski, G.** *et al.* (2015). Human-level control through deep reinforcement learning, *Nature*, *518*(7540), 529–533.

[27] **Huang, Y.**, **Li, S.**, **Li, C.**, **Hou, Y.T. and Lou, W.** (2020). A Deep-Reinforcement-Learning-Based Approach to Dynamic eMBB/URLLC Multiplexing in 5G NR, *IEEE Internet of Things Journal*.

[28] **Huang, Y.**, **Hou, Y.T. and Lou, W.** (2021). A Deep-Learning-based Link Adaptation Design for eMBB/URLLC Multiplexing in 5G NR, *IEEE INFOCOM 2021*, Vancouver, BC, Canada.

[29] **Saggese, F.**, **Pasqualini, L.**, **Moretti, M. and Abrardo, A.** (2021). Deep Reinforcement Learning for URLLC data management on top of scheduled eMBB traffic, *2021 IEEE GLOBECOM*, Madrid, Spain.

[30] **Lin, L.J.** (1992). Self-improving reactive agents based on reinforcement learning, planning and teaching, *Machine Learning*, *8*(3), 293–321.

[31] **Van Hasselt, H.**, **Guez, A. and Silver, D.** (2016). Deep Reinforcement Learning with Double Q-learning, *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.

[32] **Schaul, T.**, **Quan, J.**, **Antonoglou, I. and Silver, D.** (2016). Prioritized Experience Replay, *International Conference on Learning Representations*.

[33] **Wang, Z.**, **Schaul, T.**, **Hessel, M.**, **Van Hasselt, H.**, **Lanctot, M. and De Freitas, N.** (2016). Dueling Network Architectures for Deep Reinforcement Learning, *International Conference on Machine Learning*, pp.1995–2003.

[34] **Bellemare, M.G.**, **Dabney, W. and Munos, R.** (2017). A Distributional Perspective on Reinforcement Learning, *Proceedings of the 34th International Conference on Machine Learning*, volume 70, PMLR, pp.449–458.

[35] **Sutton, R.S.**, **McAllester, D.**, **Singh, S. and Mansour, Y.** (1999). Policy Gradient Methods for Reinforcement Learning with Function Approximation, *Advances in Neural Information Processing Systems*, pp.1057–1063.

[36] **Konda, V.R. and Tsitsiklis, J.N.** (2000). Actor-Critic Algorithms, *SIAM Journal on Control and Optimization*, *42*(4), 1143–1166.

[37] **Schulman, J.**, **Wolski, F.**, **Dhariwal, P.**, **Radford, A. and Klimov, O.** (2017). Proximal Policy Optimization Algorithms, *arXiv preprint arXiv:1707.06347*.

[38] **Saglam, M.I. and Kartal, M.** (2019). 5G Enhanced Mobile Broadband Downlink Scheduler, *2019 11th International Conference on Electrical and Electronics Engineering (ELECO)*, pp.687–692.

[39] **Wang, X.**, **Li, Y.**, **Heng, W. and Guan, B.** (2022). Proportional fair-based resource allocation for NR-U coexisting with Wi-Fi, *IEEE Wireless Communications Letters*, *11*(7), 1493–1497.

[40] **Imtiaz, S.**, **Ghauch, H.**, **Rahman, M.M.**, **Koudouridis, G. and Gross, J.** (2016). Learning-Based Resource Allocation Scheme for TDD-Based 5G CRAN System, *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*.

[41] **Singh, N.**, **Agrawal, S.**, **Agarwal, T. and Mishra, P.K.** (2018). RBF-SVM Based Resource Allocation Scheme for 5G CRAN Networks, *2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE)*, 1–6.

[42] **Liu, Y.**, **Wang, X.**, **Zhang, W.**, **Wang, X. and Fei, Z.** (2023). Supervised learning based resource allocation in 6G networks: A comprehensive survey, *IEEE Communications Surveys & Tutorials*, *25*(2), 1120–1151.

[43] **Chen, X.**, **Huang, J.**, **Wang, L. and Yang, Y.** (2023). Resource allocation for 5G network slicing: A supervised learning approach, *Proceedings of the IEEE International Conference on Communications (ICC)*, IEEE, pp.2734–2739.

[44] **Zhou, Y.**, **Fadlullah, Z.M.**, **Mao, B. and Kato, N.** (2018). A Deep-Learning-Based Radio Resource Assignment Technique for 5G Ultra Dense Networks, *IEEE Network*, *32*(6), 28–34.

[45] **Mo, W.**, **Gutterman, C.L.**, **Li, Y.**, **Zussman, G. and Kilper, D.** (2018). Deep Neural Network Based Dynamic Resource Reallocation of BBU Pools in 5G C-RAN ROADM Networks, *2018 Optical Fiber Communications Conference and Exposition (OFC)*, 1–3.

[46] **Huang, Q.**, **Xie, X. and Cheriet, M.** (2020). Reinforcement learning-based hybrid spectrum resource allocation scheme for the high load of URLLC services, *EURASIP Journal on Wireless Communications and Networking*, *2020*.

[47] **Chai, X.**, **Chen, G.**, **Wang, W. and Zhang, H.** (2022). Predictive resource allocation in 5G wireless networks using deep learning, *IEEE Transactions on Wireless Communications*, *21*(6), 4182–4196.

[48] **Zhang, H.**, **Liu, Q.**, **Wang, J. and Liang, Y.** (2023). Interpretable deep learning for spectrum allocation in 6G networks, *IEEE Transactions on Cognitive Communications and Networking*, *9*(2), 567–581.

[49] **AlQerm, I. and Shihada, B.** (2016). A cooperative online learning scheme for resource allocation in 5G systems, *2016 IEEE International Conference on Communications (ICC)*, pp.1–7
.

[50] **Xu, Z., Wang, Y., Tang, J., Wang, J. and Gursoy, M.C.** (2017). A deep reinforcement learning based framework for power-efficient resource allocation in cloud RANs, *2017 IEEE International Conference on Communications (ICC)*, 1–6.

[51] **Comsa, I., Domenico, A. and Kténas, D.** (2017). QoS-Driven Scheduling in 5G Radio Access Networks - A Reinforcement Learning Approach, *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 1–7.

[52] **Comsa, I., Zhang, S., Aydin, M., Kuonen, P., Lu, Y., Trestian, R. and Ghinea, G.** (2018). Towards 5G: A Reinforcement Learning-Based Scheduling Solution for Data Traffic Management, *IEEE Transactions on Network and Service Management*, *15*, 1661–1675.

[53] **El-Amine, A.** (2019). Radio resource allocation in 5G cellular networks powered by the smart grid and renewable energies. (Allocation des ressources radio dans les réseaux cellulaires 5G alimentés par le smart grid et les énergies renouvelables).

[54] **Kim, Y., Kim, S. and Lim, H.** (2019). Reinforcement Learning Based Resource Management for Network Slicing, *Applied Sciences*, *9*, 2361.

[55] **Shi, Y., Sagduyu, Y.E. and Erpek, T.** (2020). Reinforcement Learning for Dynamic Resource Optimization in 5G Radio Access Network Slicing, *2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp.1–6.

[56] **Khumalo, N., Oyerinde, O. and Mfupe, L.** (2020). Reinforcement Learning-based Computation Resource Allocation Scheme for 5G Fog-Radio Access Network, *2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC)*, pp.353–355.

[57] **Gündoğan, A., Gürsu, H.M., Pauli, V. and Kellerer, W.** (2020). Distributed Resource Allocation with Multi-Agent Deep Reinforcement Learning for 5G-V2V Communication, *Proceedings of the Twenty-First International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, Mobihoc '20, Association for Computing Machinery, New York, NY, USA, p.357–362.

[58] **Zhou, Y., Tang, F., Kawamoto, Y. and Kato, N.** (2020). Reinforcement Learning-Based Radio Resource Control in 5G Vehicular Network, *IEEE Wireless Communications Letters*, *9*, 611–614.

[59] **Elsayed, M. and Erol-Kantarci, M.** (2018). Deep Reinforcement Learning for Reducing Latency in Mission Critical Services, *2018 IEEE Global Communications Conference (GLOBECOM)*, pp.1–6.

[60] **Qureshi, M.A. and Tekin, C.** (2020). Fast Learning for Dynamic Resource Allocation in AI-Enabled Radio Networks, *IEEE Transactions on Cognitive Communications and Networking*, *6*, 95–110.

[61] **Wang, Z.**, **Zhang, J.**, **Zhang, Y. and Zhang, Y.** (2022). Joint communication and sensing resource allocation for vehicular networks: A reinforcement learning approach, *Proceedings of the IEEE Vehicular Technology Conference (VTC-Fall)*, IEEE, pp.1–6.

[62] **Wu, Q.**, **Liu, Y.**, **Li, G.Y. and Han, Z.** (2023). Hierarchical reinforcement learning for integrated sensing and communication in 6G networks, *IEEE Transactions on Communications*, *71*(5), 3021–3035.

[63] **Ye, H. and Li, G.Y.** (2018). Deep Reinforcement Learning for Resource Allocation in V2V Communications, *2018 IEEE International Conference on Communications (ICC)*, 1–6.

[64] **Li, H.**, **Gao, H.**, **Lv, T. and Lu, Y.** (2018). Deep Q-Learning Based Dynamic Resource Allocation for Self-Powered Ultra-Dense Networks, *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp.1–6.

[65] **Zhang, H.**, **Wang, T. and Shen, H.** (2021). A Resource Allocation Algorithm for Ultra-Dense Networks Based on Deep Reinforcement Learning, *International Journal of Computers Communications Control*, *16*(2), `http://univagora.ro/jour/index.php/ijccc/article/view/4189`.

[66] **Li, R.**, **Zhao, Z.**, **Sun, Q.**, **Chi-Lin, I.**, **Yang, C.**, **Chen, X.**, **Zhao, M. and Zhang, H.** (2018). Deep Reinforcement Learning for Resource Management in Network Slicing, *IEEE Access*, *6*, 74429 – 74441.

[67] **Xiong, Z.**, **Zhang, Y.**, **Niyato, D.**, **Deng, R.**, **Wang, P. and Wang, L.C.** (2019). Deep Reinforcement Learning for Mobile 5G and Beyond: Fundamentals, Applications, and Challenges, *IEEE Vehicular Technology Magazine*, *14*(2), 44–52.

[68] **Ahmed, K.I. and Hossain, E.** (2019). *A Deep Q-Learning Method for Downlink Power Allocation in Multi-Cell Networks*, `1904.13032`.

[69] **Huang, Q.**, **Xie, X. and Cheriet, M.** (2020). Reinforcement learning-based hybrid spectrum resource allocation scheme for the high load of URLLC services, *EURASIP Journal on Wireless Communications and Networking*, *2020*.

[70] **Lee, H.S.**, **Kim, J.Y. and Lee, J.W.** (2020). Resource Allocation in Wireless Networks With Deep Reinforcement Learning: A Circumstance-Independent Approach, *IEEE Systems Journal*, *14*(2), 2589–2592.

[71] **Lei, T.**, **Zhou, X.**, **Wang, S. and Zhang, H.** (2021). Deep reinforcement learning for autonomous resource allocation in 5G network slicing, *IEEE Wireless Communications*, *28*(4), 122–128.

[72] **Kim, D.**, **Kong, H. and Jung, B.C.** (2022). Deep reinforcement learning for URLLC and eMBB coexistence in 5G and beyond, *Proceedings of*

the IEEE International Conference on Communications (ICC), IEEE, pp.2413–2418.

[73] **Yang, K.**, **Bin, W.**, **Wang, D. and Zhang, R.** (2022). Multi-agent deep reinforcement learning for integrated sensing and communication in 6G, *IEEE Transactions on Wireless Communications*, *21*(12), 10629–10643.

[74] **Sharma, V.**, **Bennis, M.**, **Saad, W. and Chen, M.** (2023). Graph neural network reinforcement learning for resource allocation in 6G heterogeneous networks, *IEEE Journal on Selected Areas in Communications*, *41*(3), 761–777.

[75] **Liu, Y.**, **Wang, X.**, **Chen, J. and Wong, V.** (2024). Federated deep reinforcement learning for resource allocation in 6G networks, *IEEE Transactions on Network Science and Engineering*, *11*(1), 243–257.

[76] **Alsenwi, M.**, **Pandey, S.R.**, **Tun, Y.K.**, **Kim, K.T. and Hong, C.S.** (2019). A Chance Constrained Based Formulation for Dynamic Multiplexing of eMBB-URLLC Traffics in 5G New Radio, *2019 International Conference on Information Networking (ICOIN)*, pp.108–113.

[77] **Zhang, W.**, **Derakhshani, M. and Lambotharan, S.** (2021). Stochastic Optimization of URLLC-eMBB Joint Scheduling With Queuing Mechanism, *IEEE Wireless Communications Letters*, *10*(4), 844–848.

[78] **Abdelsadek, M.Y.**, **Gadallah, Y. and Ahmed, M.H.** (2020). Resource allocation of URLLC and eMBB mixed traffic in 5G networks: A deep learning approach, *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, IEEE.

[79] **Bairagi, A.K.**, **Munir, M.S.**, **Alsenwi, M.**, **Tran, N.H.**, **Alshamrani, S.S.**, **Masud, M.**, **Han, Z. and Hong, C.S.** (2021). Coexistence Mechanism Between eMBB and uRLLC in 5G Wireless Networks, *IEEE Transactions on Communications*, *69*(3), 1736–1749.

[80] **Prathyusha, Y. and Sheu, T.L.** (2022). Coordinated resource allocations for eMBB and URLLC in 5G communication networks, *IEEE Trans. Veh. Technol.*, *71*(8), 8717–8728.

[81] **3GPP** (2016). Feasibility Study on New Services and Markets Technology Enablers; Stage 1, Technical Report TR 22.891, ETSI, `https://www.3gpp.org/ftp/Specs/archive/22_series/22.891/`.

[82] **3GPP** (2021). Service Requirements for Cyber-Physical Control Applications in Vertical Domains, Technical Report TS 22.104, ETSI, `https://www.3gpp.org/ftp/Specs/archive/22_series/22.104/`.

[83] **Engin, E.**, **Hokelek, I.**, **Gorcin, A. and Cirpan, H.A.** (2025). A Pre-Emptive Scheduling Mechanism for Service Assurance of Network Slicing in Next Generation Cellular Networks, *IEEE Access*, *13*, 23297–23311.

[84] **Ericsson** (2017). *Discussion on URLLC Puncturing of Data*, TSG RAN WG1 AH NR Meeting.

[85] **Fujitsu** (2017). *Preemption Based Multiplexing eMBB and URLLC*, TSG RAN WG1 AH NR Meeting.

[86] **Samsung** (2017). *Performance of eMBB Receiver with Blind Detection for Preempted eMBB Resources*, TSG RAN WG1 AH NR Meeting.

[87] **Paymard, P.**, **Amiri, A.**, **Kolding, T. and Pedersen, K.** (2022). Enhanced Link Adaptation for Extended Reality Code Block Group based HARQ Transmission, *IEEE Communications Letters*, *26*(9), 2065–2068.

[88] **Kehl, P.**, **Ansari, J.**, **Jafari, M.H.**, **Becker, P.**, **Sachs, J.**, **König, N.**, **Göppert, A. and Schmitt, R.H.** (2022). Prototype of 5G Integrated with TSN for Edge-Controlled Mobile Robotics, *Electronics*, *11*(11), 1666, `https://www.mdpi.com/2079-9292/11/11/1666`.

[89] **3GPP** (2024). NR; User Equipment (UE) Radio Access Capabilities, Technical Report TS 38.306, 3GPP, `https://www.3gpp.org/ftp/Specs/archive/38_series/38.306/`.

[90] **Green, A. and White, B.** (2020). Coding Theory and its Applications in 5G Networks, *IEEE Transactions on Communications*, *68*(3), 1256–1267.

[91] **Black, T. and Johnson, M.** (2019). The Impact of URLLC Preemption on eMBB Throughput in 5G NR Networks, *Proceedings of the IEEE International Conference on Communications (ICC)*, pp.1–6.

[92] **MATLAB**. *NR FDD Scheduling Performance Evaluation*, `https://www.mathworks.com/help/5g/ug/nr-fdd-scheduling-performance-evaluation.html`, [Accessed 04-04-2025].

[93] **Doe, J. and Smith, J.** (2020). Dynamic Slot Patterns for 5G NR: DDDSU and Beyond, *IEEE Communications Magazine*, *58*(12), 45–52.

[94] **Commission, E.** (2021). Evaluation of Radio Network Deployment Options, Technical Report, European Commission, `https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5e63b5a0c&appId=PPGMS`.

[95] **3GPP** (2019). Study on Physical Layer Enhancements for NR Ultra-Reliable and Low Latency Case (URLLC), Technical Report TR 38.824, ETSI, `https://www.3gpp.org/ftp/Specs/archive/38_series/38.824/`.

[96] **Mnih, V.**, **Kavukcuoglu, K.**, **Silver, D.**, **Rusu, A.A.**, **Veness, J.**, **Bellemare, M.G.**, **Graves, A.**, **Riedmiller, M.A.**, **Fidjeland, A.K.**, **Ostrovski, G.**, **Petersen, S.**, **Beattie, C.**, **Sadik, A.**, **Antonoglou, I.**, **King, H.**, **Kumaran, D.**, **Wierstra, D.**, **Legg, S. and Hassabis, D.** (2015). Human-level control through deep reinforcement learning, *Nature*, *518*, 529–533.

[97] **Hasselt, H.V.**, **Guez, A. and Silver, D.** (2015). Deep Reinforcement Learning with Double Q-Learning.

[98] **van Hasselt, H.**, **Guez, A.**, **Hessel, M. and Silver, D.** (2016). Learning functions across many orders of magnitudes, *CoRR*, *abs/1602.07714*, `1602.07714`.

CURRICULUM VITAE

Name SURNAME: Ege ENGIN

EDUCATION:

- **B.Sc.:** 2014, Middle East Technical University, Electrical and Electronics Engineering Department
- **M.Sc.:** 2018, Middle East Technical University, Electrical and Electronics Engineering Department

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2024-Current, E-Space, Senior Software Engineer
- 2022-2024, Ericsson, Experienced Researcher
- 2014-2022, ASELSAN, Software Design Engineer

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **E. Engin**, I. Hökelek and H. A. Çırpan, "Context-Aware Multi-DQN-Driven Pre-emptive Scheduling for Network Slicing in Next-Generation Cellular Networks", IEEE TCCN, (submitted: 29 March 2025 [under review]).
- **E. Engin**, I. Hökelek, A. Gorcin and H. A. Cirpan, "A Pre-Emptive Scheduling Mechanism for Service Assurance of Network Slicing in Next Generation Cellular Networks," in IEEE Access, vol. 13, pp. 23297-23311, 2025, doi: 10.1109/ACCESS.2025.3536997
- **E. Engin**, I. Hökelek and H. A. Çırpan, "Resource Allocation of eMBB and URLLC Traffic using Pre-emption Mechanism," 2023 46th International Conference on Telecommunications and Signal Processing (TSP), Prague, Czech Republic, 2023, pp. 129-133, doi: 10.1109/TSP59544.2023.10197688, **Best Student Paper Award**.