Multi-user MIMO and Resource Allocation for Multi-Access Point Wi-Fi Systems

Mahboubeh Irannezhad Parizi

Doctor of Philosophy

University of York
School of Physics, Engineering and Technology

September 2025

Abstract

The increasing demands of emerging applications such as virtual and augmented reality, online gaming, and industrial wireless services have accelerated the evolution of next-generation Wi-Fi standards toward higher throughput, lower latency, and improved reliability. IEEE 802.11be (Wi-Fi 7) introduces Extremely High Throughput (EHT), while IEEE 802.11bn is being developed to support Ultra-High Reliability (UHR), where multi-access point (multi-AP) coordination has emerged as a key enabling technology. Among the four coordination schemes, this thesis focuses primarily on coordinated orthogonal frequency division multiple access (C-OFDMA) under practical Wi-Fi system assumptions.

To establish the research foundation, the CSMA/CA channel access mechanism is reconstructed and a fairness-based multi-agent reinforcement learning framework is reproduced to examine its ability to reduce channel access randomness and latency in dense environments. Building on this, a fixed resource unit (RU) allocation strategy is proposed for homogeneous coordinated Wi-Fi systems, together with a joint C-OFDMA and coordinated spatial reuse (C-SR) scheme for interference-aware throughput improvement. The study is then extended to heterogeneous systems where stations experience unequal channel gains. For this case, a deep reinforcement learning (DRL)-based variable RU allocation algorithm is developed under max-min fairness, demonstrating clear improvement in minimum throughput and Jain's fairness compared with benchmark allocation methods. Under the same system model, proportional fairness is investigated and a closed-form RU allocation solution is derived.

Finally, a realistic online scheduling problem is considered where both packet arrival rates and channel gains vary over time. To address the NP-hard queue-aware RU allocation problem, a DRL-based scheduling algorithm is proposed to minimise queueing delay, reduce packet drops, and shorten transmission opportunity duration. Simulation results confirm that the proposed methods improve fairness, reliability, and delay performance, while maintaining practical suitability for next-generation coordinated Wi-Fi systems.

To my beloved husband,
my inspiring father, and my endlessly supportive mother . . .

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References. Some of the work presented in this thesis has been planned to be submitted to IEEE journals, which are listed as follows:

1. M. I. Parizi, M. R. Ghourtani, F. Scahill and K. Cumanan, "Resource Unit Allocation in Coordinated OFDMA Multi-User Wi-Fi Systems," in IEEE Wireless Communications Letters, vol. 14, no. 6, pp. 1841-1845, June 2025, doi: 10.1109/LWC.2025.3559229.

2. M. I. Parizi, M. R. Ghourtani, D. Nunez, F. Wilhelmi, B. Bellalta, F. Scahill and K. Cumanan, "DRL-based Fair RU Allocation in C-OFDMA," Accepted to IEEE Open Journal of the Communications Society.

3. M. I. Parizi, M. R. Ghourtani, D. Nunez, F. Wilhelmi, B. Bellalta, F. Scahill and K. Cumanan, "Proportional Fair RU Allocation in Multi-AP coordination of Wi-Fi Systems" Submitted to IEEE Wireless Communications Letters.

Mahboubeh Irannezhad Parizi
September 2025

Acknowledgements

Firstly, I would like to express my deepest appreciation to my academic supervisor, Prof. Kanapathippillai Cumanan, for his unwavering support, insightful guidance, and patience throughout the development of this thesis. It is my distinct honour and privilege to have such a gracious supervisors. I would like to express my sincere gratitude to my industry supervisor, Dr. Frank Scahill, for providing a practical perspective, continuous support, and for bridging the gap between academic research and real-world applications.

I am also deeply grateful to my co-supervisor, Dr. Mostafa Rahmani Ghourtani, whose constructive feedback, thoughtful discussions, and continuous commitment have been instrumental in refining the quality of my work. His perspective and attention to detail greatly enriched both the research outcomes and my development as a researcher.

I am sincerely thankful for the opportunity to carry out a research visit at Universitat Pompeu Fabra, Barcelona, where I had the privilege to collaborate with Prof. Boris Bellalta and their outstanding team. Their expertise, collaboration, and generous support significantly contributed to broadening the scope and impact of my research.

I gratefully acknowledge the financial support from the U.K. EPSRC Industrial CASE Studentship Program under Grant EP/W522296/1(2599357), without which this work would not have been possible.

Finally, I owe a special debt of gratitude to my family for their unconditional love, patience, and encouragement throughout this journey. To my beloved husband, Mohammad, your constant support, understanding, and belief in me have been my greatest source of strength, I could not have accomplished this without you.

Table of contents

List of figures

List of tables

Abbreviation

Acronyms / Abbreviations

5G NR-U 5G New Radio - Unlicensed

AP Access Point

AR Augmented Reality

BSS Basic Service Set

C-BF Coordinated Beamforming

C-OFDMA Coordinated OFDMA

C-SR Coordinated Spatial Reuse

C-TDMA Coordinated TDMA

CDMA Code-Division Multiple Access

CSMA/CA Carrier Sense Multiple Access with Collision Avoidance

CSMA Carrier Sense Multiple Access

CTS Clear to Send

DIFS Distributed Inter Frame Space

DL Deep Learning

DPP Drift-plus-penalty

DRL Deep Reinforcement Learning

DSO Dynamic Sub-Channel Operation

DSSS Direct-Sequence Spread Spectrum

EDCA Enhanced Distributed Channel Access

EHT Extremely High Throughput

FHSS Frequency Hopping Spread Spectrum

IIoT industrial IoT

IoT Internet of Things

JTX Joint Transmission

LAA Licensed Assisted Access

MAC Layer Medium Access Control Layer

MAPC Multi-AP Coordination

MARL Multi Agent Reinforcement Learning

MCS Modulation and Coding Scheme

MDP Markov Decision Process

MIMO Multiple Input Multiple Output

MLD Multi Link Device

ML Machine Learning

MLO Multi Link Operation

MU-MIMO Multi User- Multiple Input Multiple Output

NOMA Non-Orthogonal Multiple Access

NPCA Non-Primary Channel Access

OBSS Overlapping Basic Service Set

OFDMA Orthogonal Frequency-Division Multiple Access

OFDM Orthogonal Frequency-Division Multiplexing

PAR Project Authorization Request

PHY Layer Physical Layer

PPDU Physical Layer Protocol Data Unit

QAM Quadrature Amplitude Modulation

QoS Quality of Service

RL Reinforcement Learning

RTS Request to Send

RU Resource Unit

SISO Signle Input Single Output

STA Station

STR simultaneous transmit and receive

SU PPDU Single User Physical Layer Protocol Data Unit

TDMA Time Division Multiple Access

TGax Task Group IEEE 802.11ax

TWT Target Waking Time

TXOP Transmission Opportunity

UHR Ultra High Reliability

VR Virtual Reality

WLAN Wireless Local Area Networks

WMM Weighted Max-min

Chapter 1

Introduction

1.1 Overview

Wi-Fi networks, as a fundamental component of modern communication infrastructures, play a crucial role in supporting a wide range of applications across industry, healthcare, and smart cities. This chapter provides a brief overview of Wi-Fi networks and highlights their importance. In addition, it addresses the open problems and challenges associated with the most recent Wi-Fi standards. Building upon this context, the key motivations and objectives underlying this study are presented. Finally, the primary contributions of the thesis are outlined, organized according to the structure of each chapter.

1.2 Wi-Fi Network

Wireless Local Area Networks (WLANs), commonly referred to as Wi-Fi, provide high-speed, cost-effective, and flexible connectivity across residential, enterprise, and public environments, and currently account for approximately 60% of global data traffic [12]. Wi-Fi, the commercial name for the IEEE 802.11 family of standards, was initially introduced in 1997 with link rates of $1-2$ Mbps. Since then, successive amendments have considerably enhanced its performance and capabilities. With the advent of emerging applications such as video streaming, cloud computing, augmented/virtual reality (AR/VR), and the Internet of Things (IoT), new standards have been introduced, each incorporating increasingly stringent performance requirements. These requirements initially emphasized the upper and lower bounds

of Wi-Fi 7 throughput and have more recently expanded to include latency and reliability considerations [1, 13, 2].

Across Wi-Fi generations, enhancements to the physical (PHY) and medium access control (MAC) layers have been pursued to address the requirements of diverse applications. IEEE 802.11a/b/g provided single-user connectivity with throughput up to 54 Mbps, while IEEE 802.11n (Wi-Fi 4) introduced multi-antenna techniques (MIMO) and channel bonding, raising achievable data rates up to 600 Mbps. IEEE 802.11ac (Wi-Fi 5) further enhanced MIMO with multi-user downlink transmission and wider channels of up to 160 MHz, enabling multi-gigabit throughput. More recently, IEEE 802.11ax (Wi-Fi 6/6E) marked a paradigm shift by incorporating orthogonal frequency division multiple access (OFDMA), uplink MU-MIMO, and operation in the 6 GHz band, targeting higher spectral efficiency, reduced latency, and improved performance in dense environments such as stadiums and offices [5, 1, 6, 7].

The forthcoming IEEE 802.11be/bn standards, referred to as Wi-Fi 7 and beyond, are projected to achieve peak throughputs exceeding 23 Gbps through the support of up to 16 spatial streams, and 320 MHz channel bandwidths. It also introduces advanced multi-AP coordination techniques, including coordinated spatial reuse (C-SR), coordinated OFDMA (C-OFDMA), coordinated beamforming (C-BF), and joint transmission (JTX). However, due to the high complexity of synchronization, signaling, and scheduling, multi-AP coordination has been postponed to the next Wi-Fi generation. In addition, the relevant research communities and standardization organizations, including IEEE, are actively investigating IEEE 802.11bn, with a stronger focus on ultra-high reliability to address the demands of industrial and mission-critical applications. Among the proposed technologies, multi-AP coordination is particularly promising for improving reliability and mitigating channel access contention [1, 7, 14].

Despite these advances, Wi-Fi continues to face fundamental challenges. In contrast to cellular systems, which operate over licensed spectrum with centralized scheduling, Wi-Fi relies on unlicensed bands (2.4 GHz, 5 GHz, and more recently 6 GHz). While this approach reduces deployment costs and enhances accessibility, it simultaneously increases the likelihood of interference, collisions, and failing to meet the quality of service (QoS) requirements of STAs. Channel access in Wi-Fi is governed by the decentralized carrier-sense multiple access with collision avoidance (CSMA/CA) mechanism, which is simple but often inefficient in dense environments. The adoption of OFDMA and multi-user MIMO has improved

spectrum utilization. However, practical scheduling and resource allocation methods remain underdeveloped relative to those in cellular systems. Given that Wi-Fi currently supports more than 60% of global mobile data traffic, the need for robust coordination strategies has become increasingly critical. Approaches such as multi-AP coordination, resource allocation optimization, and intelligent scheduling are expected to play a pivotal role in ensuring scalability and reliability in next-generation Wi-Fi networks [1, 7].

One of the key new technologies in the upcoming Wi-Fi generations is the introduction of multi-AP coordination. This mechanism is designed to mitigate the limitations of decentralized channel access, particularly in dense deployment scenarios [14]. In previous Wi-Fi generations, it was assumed that APs managed the QoS of their associated STAs within a basic service set (BSS) independently, without coordination across APs. However, this approach often results in inefficient resource utilization by some APs, while simultaneously generating interference and collisions for neighboring BSSs, especially in overlapping BSSs (OBSSs). Multi-AP coordination addresses these challenges by enabling adjacent (and potentially non-adjacent) APs to exchange control information and jointly manage spectrum, beamforming, and scheduling. Coordination can be realized through different approaches, such as C-OFDMA/C-TDMA, where APs schedule transmissions to mitigate collisions. Also, C-SR, which enhances frequency and time reuse efficiency, C-BF that allows the APs with joint beamforming to create nulls for non-intended STAs, and JTX, where multiple APs simultaneously serve a device to improve signal quality and reduce interference [2, 1].

This trade-off becomes particularly important in scenarios where STAs have different priorities or QoS requirements. For instance, assigning identical bandwidth resources to each AP with different primary channels may lead to inefficient resource utilization. Conversely, if APs operate on the same primary channel without coordinated RU allocation, collisions and significant performance degradation may occur [2, 15]. Moreover, many existing works primarily focus on RU allocation for throughput maximization [16, 17]. In heterogeneous environments with varying channel conditions, such approaches tend to favour STAs with better channel quality, leading to unfair resource distribution. These challenges highlight the need for intelligent RU allocation strategies that jointly consider fairness, efficiency, and QoS requirements.

The new features and technologies introduced for C-OFDMA also increase the complexity of RU allocation compared to the single-AP scenario. In coordinated

environments, APs must determine how to partition the available bandwidth among multiple STAs while accounting for heterogeneous channel conditions, QoS requirements, and coordination constraints among neighboring APs. Consequently, a trade-off arises between bandwidth allocation and transmission time. A STA may require a larger bandwidth allocation to transmit or receive data efficiently, whereas assigning a smaller bandwidth results in longer transmission durations. Therefore, it is essential to know the tradeoff between bandwidth allocation, time allocation, or, in general, the resource allocation in C-OFDMA. The reason is that a STA requires either a larger bandwidth allocation to transmit or receive data efficiently, or, if a smaller bandwidth is assigned, the data transmission will take more time. This shows the RU allocation tradeoff for different requirements of networks. For example, in a case with variable priority and QoS of STAs, assigning the same bandwidth for each AP with a different primary channel is not optimal, and this will lead to inefficient resource utilization. While if APs use the same primary channel for transmission without RU allocation, this can cause collisions and a huge loss of resources and data [2, 15]. Beyond this, most works seek RU allocation for throughput maximization [16, 17]. However, in scenarios with several STAs with different conditions, for example, for scenarios with variable channel conditions, STAs with better channel conditions will get the total bandwidth, which is not fair at all.

For the forthcoming IEEE 802.11bn amendment, which emphasizes UHR and low-latency communication, multi-AP coordination is anticipated to serve as a key enabler. By enabling APs to operate similarly to a distributed antenna system, multi-AP techniques enhance throughput, fairness, and reliability, particularly in high-density deployments and mission-critical scenarios. Joint transmission reduces packet error rates through spatial diversity, while coordinated scheduling lowers latency by mitigating random access delays.

Furthermore, scheduling in such spectrum-sharing contexts alleviates the inefficiencies of CSMA/CA, thereby moving Wi-Fi closer to the deterministic scheduling paradigm of cellular networks while maintaining its inherent cost efficiency and flexibility. These advantages position multi-AP coordination as a central component of IEEE 802.11bn to deliver reliability, high performance, and the ability to handle large-scale, mission-critical operations over unlicensed spectrum [2, 7, 1]. Further details will be provided in the following chapter to elaborate on the background of Wi-Fi networks and the emerging technologies envisioned for Wi-Fi 7 and beyond.

The evolution of Wi-Fi toward IEEE 802.11bn places increasing emphasis on ultra-high reliability and low-latency communication in dense and interference-limited

environments. Although advanced PHY and MAC features such as OFDMA and multi-AP coordination offer substantial performance gains, their effectiveness remains fundamentally constrained by the decentralized nature of CSMA/CA. Random channel access leads to unpredictable delay and reliability degradation, particularly in dense deployments where contention and interference dominate system performance. Consequently, improving channel access behavior is a prerequisite for enabling more efficient and reliable resource allocation in coordinated Wi-Fi networks. These challenges motivate the exploration of intelligent scheduling mechanisms capable of adapting channel access and resource allocation decisions according to dynamic network conditions.

Building upon this observation, this thesis investigates a sequence of interrelated problems spanning channel access, interference-aware resource allocation, fairness, and delay management in coordinated OFDMA systems. First, intelligent channel access mechanisms based on reinforcement learning are explored to reduce access randomness and improve reliability. On top of this enhanced access foundation, practical RU allocation strategies are developed for joint C-OFDMA and C-SR to mitigate inter-AP interference under realistic implementation constraints. The impact of channel gain on fairness-oriented RU allocation is then analyzed to understand trade-offs between throughput maximization and equitable resource distribution. Finally, to directly address the stringent latency and reliability requirements of IEEE 802.11bn, queue-aware RU scheduling is proposed to incorporate traffic dynamics and QoS constraints. Together, these objectives form a coherent framework for improving reliability, latency, and fairness in next-generation Wi-Fi networks through practical and scalable coordination mechanisms. Motivated by these challenges, this thesis investigates a sequence of interrelated problems spanning channel access, interference-aware resource allocation, fairness, and delay management in coordinated OFDMA Wi-Fi networks.

1.3 Motivations and Research Contributions

Although multi-AP coordination has the potential to significantly enhance the performance of next-generation Wi-Fi systems, several practical challenges remain unresolved. In dense deployments, conventional Wi-Fi channel access based on CSMA/CA suffers from excessive contention, random backoff delays, and inefficient channel utilisation. These issues become increasingly severe as the number of devices grows, leading to degraded reliability and increased latency. Such limitations

motivate the investigation of alternative mechanisms that can improve channel access efficiency while maintaining compatibility with existing Wi-Fi protocol operations.

Recent Wi-Fi standards, particularly IEEE 802.11be and the emerging IEEE 802.11bn, aiming to support extremely high throughput and ultra-high reliability (UHR). Achieving these goals requires more intelligent resource management mechanisms that can adapt to dynamic network conditions and dense user environments. Among the candidate technologies, C-OFDMA has emerged as a key mechanism for improving spectrum utilisation by allowing multiple STAs to transmit simultaneously using different RUs. However, efficient allocation of RUs remains a challenging problem because scheduling decisions must account for heterogeneous channel conditions, interference between neighbouring APs, and the QoS requirements of STAs.

In addition, the inherent randomness of CSMA/CA limits the ability of conventional scheduling strategies to achieve consistent latency and reliability performance. This has motivated the exploration of learning-based approaches capable of adapting channel access decisions based on observed network conditions. In particular, RL has recently gained attention as a promising tool for wireless resource management because it enables agents to learn optimal decision policies through interaction with the environment without requiring an explicit analytical model of the system. Therefore, RL-based approaches provide a potential mechanism for improving channel access behaviour in dense Wi-Fi environments.

Furthermore, while OFDMA enables parallel transmissions through RU partitioning, determining how these RUs should be allocated among STAs is a non-trivial scheduling problem. Practical RU allocation must consider multiple factors, including channel quality variations, fairness among users, interference between coordinated APs, and traffic dynamics. Despite the importance of these aspects, existing works often assume simplified allocation strategies that do not fully exploit available network information or fail to address realistic traffic conditions.

Motivated by these challenges, this thesis investigates intelligent resource allocation and channel access strategies for coordinated Wi-Fi networks. The research focuses on improving fairness, reliability, and delay performance while maintaining practical computational complexity for real deployments. These challenges motivate the investigation of practical scheduling and resource allocation mechanisms that can improve reliability, fairness, and delay performance in coordinated Wi-Fi networks. Accordingly, the thesis pursues four main research objectives:

1. **Learning-based channel access for reliability and latency improvement.** To investigate the application of RL algorithms for channel access in Wi-Fi

networks, with the objective of improving reliability and reducing latency compared with the conventional CSMA/CA mechanism.

2. **Practical RU allocation for interference management in coordinated Wi-Fi.** To design a computationally efficient fixed RU allocation strategy for a joint C-OFDMA and C-SR system model, enabling effective interference management between neighbouring APs.

3. **Fairness-aware RU allocation based on channel conditions.** To analyse the influence of channel gain on RU allocation and develop algorithms that achieve max–min and proportional throughput fairness among STAs in coordinated OFDMA networks.

4. **Queue-aware RU scheduling for delay-sensitive traffic.** To develop a queue-aware variable RU allocation algorithm capable of adapting scheduling decisions according to both channel conditions and traffic demand, with the aim of reducing queue latency and improving transmission reliability.

The main contributions and novelties of this thesis can be summarised as follows. First, the Fair-MARL channel access algorithm is evaluated and extended to heterogeneous variable-rate Wi-Fi scenarios, identifying key limitations and open problems including QoS-differentiated access aligned with EDCA. Second, a novel low-complexity joint C-OFDMA and C-SR RU allocation algorithm is proposed under practical semi-distributed coordination constraints, with each AP capable of acting as the master AP. Third, fairness-aware variable RU allocation under heterogeneous channel conditions is investigated, yielding both a DRL-based solution for the NP-hard max-min fairness problem and a closed-form solution with O(1) complexity for the proportional fairness case. Fourth, a queue-aware variable RU allocation framework is proposed for the first time in IEEE 802.11bn under stochastic, non-saturated traffic conditions, addressed through a DRL-based online scheduling algorithm. These contributions advance the state of the art in intelligent resource allocation for next-generation coordinated Wi-Fi systems, with results published or under review in IEEE Wireless Communications Letters and IEEE Access.

The non-triviality of these optimisation problems arises from constraints that are unique to the Wi-Fi protocol stack and the emerging C-OFDMA framework of IEEE 802.11bn. Unlike cellular OFDMA, where resource allocation is performed by a centralised scheduler with guaranteed channel access, Wi-Fi OFDMA operates on top of CSMA/CA which is a contention-based, decentralised mechanism, meaning

that scheduling decisions must jointly account for TXOP duration, probabilistic channel access, and signalling overhead including MU-RTS/CTS and trigger frames. Additionally, the C-OFDMA coordination procedure of IEEE 802.11bn remains under active standardisation, requiring problem formulations under practical assumptions with no direct precedent in the literature. Within this setting, the joint C-OFDMA and C-SR formulation in Chapter 5 introduces an OBSS interference structure requiring a graph-colouring approach not present in single-AP works; the fairness problems in Chapter 6 go beyond the fixed and equal-sized RU assumptions that dominate the existing Wi-Fi literature; and the problem in Chapter 7 addresses non-saturated stochastic traffic as a condition that prior Wi-Fi OFDMA works almost universally ignore by assuming fully saturated buffers.

To address the outlined objectives, this thesis first presents a comprehensive review of Wi-Fi networks, with a particular focus on Wi-Fi 7 and beyond. The advanced technologies introduced in IEEE 802.11be and IEEE 802.11bn, targeting extremely high throughput (EHT) and UHR, are examined in detail. Among these, multi-AP coordination is highlighted as a transformative approach due to its potential to deliver significant performance improvements. Nevertheless, each category of multi-AP coordination, C-SR, C-BF, C-OFDMA, and JTX, introduces specific challenges and limitations. In the initial stage of this work, C-OFDMA is selected as the primary case study, particularly under practical assumptions for Wi-Fi networks.

To develop a fundamental understanding of Wi-Fi network design and the CSMA/CA protocol, this work first implements a CSMA/CA model that enables the simulation and analysis of packet transmission procedures, signaling overhead, available information for scheduling, channel access challenges, and other essential fundamentals. Building upon this foundation and drawing inspiration from the literature, a channel access scheme based on multi-agent reinforcement learning (RL) is regenerated and analyzed. It is proposed to be in line with the primary objective of IEEE 802.11bn to achieve higher reliability with lower latency. RL offers the potential to overcome the inherent randomness of CSMA/CA. The principal motivation for this approach lies in the fact that conventional CSMA/CA can incur substantial delays in dense wireless environments, whereas RL provides a promising means of mitigating such latency by reducing access randomness.

Although conventional optimization methods remain suitable for several scheduling problems considered later in this thesis, they become increasingly difficult to apply when the scheduling environment evolves online and multiple system

variables must be jointly considered. In particular, latency-sensitive Wi-Fi scheduling requires rapid adaptation to stochastic queue evolution, variable channel quality, and discrete RU assignment constraints, which can lead to combinatorial complexity. Under such conditions, DRL offers an attractive framework because it can learn scheduling policies directly from observed system states without requiring repeated real-time solution of computationally expensive optimization problems.

In addition to the basic implementation and analysis of the built-in Wi-Fi protocol, this work investigates C-OFDMA scheduling with the objective of throughput maximization while satisfying STAs' QoS, and managing interference through joint consideration with C-SR. The analysis reveals that in practical scenarios, OFDMA scheduling requires further attention, as multiple factors influence RU allocation. Since OFDMA operates on top of CSMA/CA, it becomes essential to examine the factors that affect RU scheduling in this context, as they can either degrade scheduling performance or enhance key network metrics such as throughput, delay, and reliability. Furthermore, the necessity of coordination among APs provides additional motivation to explore what practical information can be leveraged for intelligent scheduling, particularly as system complexity grows rapidly. To address interference management through joint RU allocation in C-OFDMA and C-SR, this thesis proposes an algorithm developed under practical system assumptions. Furthermore, in order to account for implementation costs, the RU allocation is designed such that, rather than relying on a single centralized controller, each AP can act as the master AP within a group of coordinating APs.

In addition, this work investigates the impact of channel gain on max-min and proportional throughput fairness in a centralized C-OFDMA network. For the max-min fairness case, an algorithm is proposed and evaluated, while for proportional throughput fairness a closed-form solution for RU allocation is derived under practical network assumptions.

Finally, to fulfill the primary goal of IEEE 802.11bn of delay reduction and reliable transmission, this thesis proposes a queue-aware RU allocation algorithm aimed at minimizing queue latency and enhancing the reliability of packet delivery. The algorithm incorporates features of the practical QoS requirements of STAs that influence network performance with respect to latency and reliability. The subsequent section provides further details regarding the overall structure and outline of this thesis.

1.4 Thesis Outline

This thesis is structured into eight chapters, organized as follows,

- **Chapter 1- Introduction:** This chapter provides an introduction to Wi-Fi networks along with a brief background on their associated technologies. In addition, it outlines the motivations of this study and presents the primary scope, objectives, and goals addressed in each chapter.
- **Chapter 2- Wi-Fi Technology and Literature Review:** In this chapter, Wi-Fi networks are first reviewed against cellular networks in terms of their technical features, with the intention of providing a fundamental comparison rather than a discussion of service delivery. Subsequently, an overview of the previous amendments of Wi-Fi standards from the first generation is presented. The discussion then turns to the key technologies introduced in the most recent Wi-Fi, Wi-Fi 7, and beyond, focusing on both the PHY and MAC layers. Multi-AP coordination and its benefits are highlighted as one of the novel technologies envisioned for Wi-Fi 7 and beyond. Thereafter, the categories of multi-AP coordination are reviewed, with particular emphasis on C-OFDMA, which is selected as the primary focus of this study. To provide a deeper technical context, channel access mechanisms and the transmission signaling procedures of OFDMA/C-OFDMA are discussed. Finally, the chapter concludes with a summary of existing research on the fundamentals of C-OFDMA and channel access protocols.
- **Chapter 3- Mathematical Backgrounds:** In this chapter, the mathematical foundations relevant to the proposed problem formulations and algorithms in this thesis are presented. The discussion begins with fairness metrics, which are employed and further developed throughout the study. Subsequently, the fundamentals of convex optimization are introduced, followed by an overview of machine learning (ML) techniques. The ML algorithms, including supervised learning, unsupervised learning, reinforcement learning, and federated learning, are discussed, which is provided with additional clarification and context for their application in this work.
- **Chapter 4- Simulation-Based Validation: Fair-based MARL for Variable-Rate Wi-Fi Network:** This chapter presents the system implementation of the CSMA/CA protocol and the Fair-MARL algorithm for channel access. The

application of the Fair-MARL algorithm to mitigate the backoff time inherent in the CSMA/CA protocol is examined with regenerating numerical results. Simulation results demonstrate the effectiveness of Fair-MARL in achieving more efficient channel access within the same time duration, particularly in densely populated environments. In addition both CSMA/CA and the Fair-MARL algorithm are examined with in a heterogeneous system where variable data rates are available. Nevertheless, further analysis of the algorithm under more realistic scenarios is proposed, along with the development of an approach that enables proportional channel access in accordance with the QoS requirements of access categories, similar to the EDCA protocol.

- **Chapter 5- Resource Unit Allocation in Coordinated OFDMA Multi-User Wi-Fi Systems:** In this chapter, a joint C-OFDMA and C-SR RU allocation scheme is proposed. A low-complexity and practical algorithm is developed to optimize RU allocation, with the dual objectives of maximizing network throughput and satisfying the QoS requirements of each STA, while minimizing OBSS interference. The system model is formulated as a semi-distributed multi-AP coordination framework, with an emphasis on ensuring the practicality of implementation.

- **Chapter 6- Fair Variable RU Allocation in C-OFDMA:** This chapter begins by analyzing the effect of channel gain on variable RU allocation in C-OFDMA. Then, a max-min throughput formulation for STAs is presented to assign variable RU allocation based on the STAs channel gain. Owing to the NP-hard nature of the problem, a DRL-based algorithm is introduced to address the optimization. Subsequently, proportional throughput fairness is considered, with the objective of identifying variable RU allocation algorithm that ensure fairness across scheduling intervals. Later in the chapter, it is proved that the proposed problem formulation is convex, and a closed-form solution is derived under the assumption that all STAs are saturated.

- **Chapter 7- Online Variable RU Allocation in C-OFDMA with Buffer Emptying:** In this chapter, a more realistic scenario is considered in which not all STAs are saturated. Accordingly, variable RU allocation is investigated in an online network setting that accounts for both variable packet arrival rates and channel gains, with queue-awareness as the primary objective. The proposed algorithm aims to minimize queues waiting time delays while simultaneously

reducing packet drops and mitigating TXOP duration as effectively as possible, with the main goals of lower latency and higher reliability.

- **Chapter 8- Conclusions, Open Issues, and Future Works:** In the final chapter, the key contributions of this research are summarized, and the results presented in each chapter are concluded. In addition, potential directions for future work are discussed.

1.5 Summary

In this chapter, a brief introduction to Wi-Fi networks was presented, including their background and open research challenges. The motivations underlying this study were outlined, leading to the definition of the key objectives and contributions. In addition, the contributions of each chapter were summarized. The next chapter presents a literature review on Wi-Fi networks, highlighting their technical differences with cellular systems.

Chapter 2

Literature Review on Wi-Fi Technology

Wi-Fi technology, commercially known under the IEEE 802.11 standard for WLAN networks, has been developed since 1990 [13]. Initially, its primary function was to enable device connectivity through a wireless local area network or WLAN. Beginning with the first Wi-Fi standard in 1997, numerous amendments have been introduced to support emerging use cases in Wi-Fi networks. Each standard, starting with IEEE 802.11 which was offering $1-2$ Mbps link throughput to the most recent IEEE 802.11ax standard with 10 Gbps throughput, has been designed to address increasing demands such as higher throughput, higher reliability, and lower latency [18]. Wi-Fi networks carry more than half of today's network traffic, handling around 63% of global mobile data. The rest is delivered through cellular networks. This brings up a key question: what are the main technical differences between Wi-Fi and cellular networks, and how are they used differently? In the next sections, it is attempted to answer these questions with the fundamental concepts of WiFi technology while reviewing recent relevant work in the literature.

2.1 Wi-Fi vs Cellular

Wi-Fi was initially introduced with the primary objective of delivering services traditionally provided by wired networks, including high throughput, strong reliability, and consistent network connectivity [19, 13, 20]. The fundamental concept was to enable users to transmit locally using unlicensed frequency bands, ISM. Over time, both Wi-Fi and cellular networks have evolved to meet the increasing demands of users, emerging new use cases, and applications such as AR and VR with stringent delay requirements. Wi-Fi began with the IEEE 802.11 standard and has since

advanced to the current Wi-Fi 7 (IEEE 802.11be standard), which is being followed by IEEE 802.11bn enhancements (which could be commercialized as Wi-Fi 8) focused on ultra-reliability. In parallel, the cellular network has progressed from its first generation to the current 5G, and is now advancing toward 6G to provide growing, similar demanding requirements. The main technical distinctions between Wi-Fi and cellular networks lie in their spectrum usage, channel access methods, multiple access schemes, network architecture, and QoS capabilities, which are elaborated below [4, 18, 14, 21].

- **Spectrum Usage:** Wi-Fi networks operate over unlicensed spectrum bands, currently utilizing the 2.4, 5, and 6 GHz frequency channels [22], and it is probable to open 45 and 60 GHz links for the new IEEE 802.11bq amendment [1] for combining millimeter wave (mmWave) links with links lower that 7 GHz. The unlicensed nature of these bands allows all users in the network to attempt channel access without restriction [1]. In contrast, cellular networks traditionally operate on licensed spectrum. However, in certain scenarios and specific generations, such as 5G NR-U and LAA, unlicensed spectrum is also employed. The licensed cellular spectrum ranges from 700 MHz to 3.5 GHz, with additional mmWave bands in the 26–100 GHz range. Unlicensed spectrum permits channel access to all users, whereas licensed spectrum is exclusively reserved for operators who have obtained licenses for the specific licensed spectrum. This distinction represents the main difference between Wi-Fi and cellular networks. Because all users in the Wi-Fi network can attempt to access the available channels, there is a higher probability of interference, particularly in densely populated areas, which may result in collisions. Cellular networks, on the other hand, can mitigate such interference through centralized planning and scheduling mechanisms as well as through efficient resource allocation techniques[22, 4].

- **Channel Access Techniques:** In Wi-Fi networks, all users have the ability to access the communication link; however, this open channel access can lead to significant interference and packet collisions, especially in dense networks. To manage such interference within unlicensed bandwidth, the CSMA protocol has been employed for channel access. However, the CSMA can cause huge overhead and latency specially in overcrowded environments. This is due to the protocol channel access contention which later in Section. 2.6.1 is elaborated. Therefore, to improve latency the CSMA/CA protocol

was introduced, incorporating RTS/CTS signaling which significantly reduced latency. Under CSMA/CA, each STA that intends to transmit data must first sense the channel to determine whether it is idle or busy. If the channel is busy, the STA selects a random backoff interval and checks the channel status, if it is idle at each time slot, it will decrease backoff counter. Once the channel is found to be idle and the backoff counter reaches zero, the STA can access the channel [1, 4]. On the other hand, cellular networks use a scheduled channel access mechanism such as TDMA or OFDMA, which is centrally managed by the core network.

- **Multiple Access Techniques:** With respect to multiple access techniques, cellular networks have employed OFDMA, code division multiple access (CDMA) and non-orthogonal multiple access (NOMA) [23], along with massive MIMO, which provide extensive support for simultaneous channel access and efficient resource allocation. These technologies significantly enhance the ability of cellular networks to manage and coordinate multiple users. On the other hand, Wi-Fi 6 has only recently adopted OFDMA, and implemented OFDMA on top of the CSMA/CA protocol. The OFDMA feature in Wi-Fi 6 (IEEE 802.11ax) lacks practical solutions for resource allocation and scheduling algorithms. Therefore, the challenge is still underdeveloped in the Wi-Fi network to be capable of meeting the growing demands of emerging use-cases and applications [13]. In terms of spatial reuse, current Wi-Fi standards support up to 8×8 spatial streams in MIMO. In comparison, massive MIMO is deployed in 5G cellular networks, which can support up to 64 transmit antennas in theory [8, 1].

- **Network Architecture:** The architecture of Wi-Fi networks is designed around decentralized scheduling, eliminating the need for a centralized core as found in cellular networks. However, with increasing network density and the emergence of new technologies, there is growing consideration for integrating certain centralized functionalities at APs. These functionalities may include multi-AP coordination, seamless handovers, and limited scheduling algorithms. Despite the introduction of limited coordination and control functionalities at the AP level, the Wi-Fi network architecture remains fundamentally less centralized than that of cellular networks [1, 4].

- **QoS satisfaction:** Another key difference between Wi-Fi and cellular networks is QoS requirements. Cellular networks are designed to provide guaranteed

and application-specific QoS, to ensure different performance levels per service type. In contrast, Wi-Fi networks up to Wi-Fi 6 have primarily relied on a best-effort prioritization model. However, with the evolution of newer Wi-Fi generations, it is anticipated that advanced scheduling algorithms will enable better QoS delivery for STAs [4].

Table 2.1 Wi-Fi vs Cellular [1, 4].

Technical Distinctions	**Wi-Fi**	**Cellular**
Spectrum usage	Unlicensed band: 2.4,5,6 GHz, for mmWave 45,60 GHz	Licensed band: 700MHz-3.5 GHz, for mmWave 26 – 100 GHz
Channel access techniques	CSMA/CA	A scheduled channel access mechanism such as TDMA/OFDMA
Multiple access techniques	OFDMA on top of CSMA/CA, Up to 64 Spatial stream	OFDMA, CDMA and NOMA, Massive MIMO
Network architecture	Mostly focused on decentralized access	Centralized channel access
QoS satisfaction	A best-effort prioritization model	Guaranteed and application-specific QoS

2.2 Wi-Fi Generations

Since the initialization of the IEEE 802.11 standard, the standard has continuously evolved to meet the different demanding requirements of users over time. The main IEEE 802.11 amendments introduced throughout the years are elaborated in the following, along with their key features, typical use cases, and the maximum possible data rates that can be supported 2.2.

- IEEE 802.11 started on 2.4 GHz frequency band in unlicensed bandwidth, ISM band, with the primary goal of achieving a $1-2$ Mbps data rate. The main features of this standard included a 22 MHz bandwidth with one spatial stream using either direct sequence spread spectrum (DSSS) or frequency hopping spread spectrum (FHSS).

- In IEEE 802.11a, the allocated frequency band was either 5 GHz or 3.7 GHz, with a bandwidth of 20 MHz, and this amendment was released in September 1999. OFDM was used as its modulation scheme, and a SISO antenna technology was employed. This standard was capable of achieving a maximum throughput of 54 Mbps. The main use cases for IEEE 802.11a were faster enterprise connectivity and early stages of media streaming.

- The IEEE 802.11b standard was also released in September 1999 for use cases such as email, browsing, file sharing, and similar applications. Its frequency band was 2.4 GHz with a bandwidth of 22 MHz, while DSSS was used as its waveform. In addition, the same SISO antenna configuration was used, and the maximum provided throughput was 11 Mbps. In this case, the 2.4 GHz frequency range could become crowded, increasing the risk of interference.

- In 2003 the standard IEEE 802.11g was approved for WiFi network for applications such as Home media, and consumer electronics integration. The maximum network throughput for this standard was 54 Mbps. This standard used 2.4 GHz frequency with the bandwidth of 20 MHz. The OFDM or DSSS modulation was employed with the SISO antenna technology.

- The IEEE 802.11n standard, also known as Wi-Fi 4, was developed with the primary objective of enabling high throughput (HT). It reached a maximum data rate of 600 Mbps by introducing two major enhancements: the use of multiple input, multiple output (MIMO) technology, and the expansion of channel bandwidth from 20 MHz to 40 MHz. MIMO not only improves data throughput but also extends coverage range, allowing support for a greater number of users. Like the previous generations of IEEE 802.11g and IEEE 802.11a, the 802.11n standard employs OFDMA modulation. This standard was specifically designed to support applications such as HD video streaming, cloud service access, and connectivity in multi-device households.

- The IEEE 802.11ac standard, commercially known as Wi-Fi 5, was established in December 2013 with the goal of supporting very high throughput (VHT). This amendment operates in the 5 GHz frequency band and supports bandwidths of 20, 40, 80, and 160 MHz. The standard can achieve a peak data rate of up to 6.9 Gbps using up to 4×4 downlink MU-MIMO with 256-QAM. Therefore, as expected, IEEE 802.11ac, by employing multi-user MIMO technology, allowed users to communicate through one or more antennas, thereby improving

throughput. The main use cases for Wi-Fi 5 included 4K video streaming, online gaming, mobile offloading, and enterprise WLANs.

- The IEEE 802.11ax standard, also known as Wi-Fi 6, was designed to achieve high efficiency (HE) throughput and was finalized in September 2020. This standard was defined for applications such as IoT, dense environments (e.g., stadiums and campuses), and online conferencing. IEEE 802.11ax supported MIMO spatial streams with up to 8×8 antennas and utilized 1024-QAM. In this amendment for the first time proposed spatial reuse through power control and BSS coloring. Additionally, multi-user transmission in both downlink and uplink OFDMA, as well as target wake time (TWT), were introduced for the first time for multi-access and energy saving, respectively. Furthermore, both 2.4 GHz and 5 GHz frequency bands were proposed for use without any management and scheduling to use. The achievable peak data rate is 9.607 Gbps.

- IEEE 802.11be, also known as Wi-Fi 7, is the latest IEEE 802.11 standard and was commercialized in 2024. The standard was defined with the primary goal of supporting high-resolution applications such as 8K video, real-time collaboration, cloud gaming, AR/VR, and IIoT. Key features proposed in this new generation include support for 320 MHz bandwidth across the 2.4, 5, and 6 GHz frequency bands. Instead of relying on random frequency link selection, more intelligent link usage was proposed along with multi-link operation with the main objective of IEEE 802.11be for EHT. Additionally, this standard introduced up to 4096-QAM modulation and multi-RU assignment. Further details on this standard are provided in Sections 2.3 and 2.4.

  IEEE 802.11bn, expected to be released in 2028 under the commercial name Wi-Fi 8, extends the capabilities of previous amendments to support UHR for applications such as robotic surgery preparation, industrial automation, holography, and ultra-reliable closed control loops. Several key features have been proposed, including distributed RU allocation, low-density parity-check (LDPC) enhancement, unequal modulation, seamless roaming, dynamic power save, non-primary channel access (NPCA), dynamic sub-channel operation (DSO), and multi-AP coordination (MAPC). Although these represent the initial set of features, further details are still under development [5, 1, 6, 7].

Table 2.2 Comparison of IEEE 802.11 Amendments [5, 1, 6, 7]

IEEE 802.11 Amendments	Generation	Release Date	Frequency Band (GHz)	Channel Width (MHz)	Key Features	Peak Rate
802.11	—	1997	2.4	22	DSSS, FHSS, SISO	1 – 2 Mbps
802.11a	—	1999	3.7/5	20	OFDM, SISO	54 Mbps
802.11b	—	1999	2.4	22	DSSS,SISO	11 Mbps
802.11g	—	2003	2.4	20	OFDM, DSSS, SISO	54 Mbps
802.11n	Wi-Fi 4	2009	2.4/5	Up to 40	OFDM, MIMO	600 Mbps
802.11ac	Wi-Fi 5	2013	5	Up to 160	OFDM, MU-MIMO, 256-QAM	6.9 Gbps
802.11ax	Wi-Fi 6/6E	2020	2.4/5	Up to 160	DL/UL OFDMA, MU-MIMO, TWT, 1024-QAM	9.6 Gbps
802.11be	Wi-Fi 7	2024	2.4/5/6	Up to 320	MLO, multi-RU, 4096-QAM, R-TWT, enhanced QoS	23 Gbps
802.11bn	Expected Wi-Fi 8	2028+	2.4/5/6 maybe greater than 7	Up to 320	MAPC, distributed RU, dynamic power saving, and etc	TBD

2.3 Wi-Fi 7 and beyond

Due to the growing demand for high-throughput and low-latency networks—driven by emerging technologies such as VR/AR, online gaming, 4K–8K streaming, and wide-area coverage for industrial IoT—the IEEE 802.11 task group has proposed new amendments to address these evolving network requirements. IEEE 802.11be, developed by task group TGbe with the primary goal of EHT, aims to deliver at least 30 Gbps of network throughput. As a result, new and improved PHY and MAC technologies have been proposed and studied to meet this objective and its associated demands. IEEE 802.11be has been standardized and commercialized under the name Wi-Fi 7, although much of the ongoing discussion has already shifted toward the next amendment, IEEE 802.11bn (Expected Wi-Fi 8) [8, 1].

IEEE 802.11bn, introduced by task group TGbn in November 2023, has the primary goal of achieving UHR with a 25% throughput enhancement and a 25% reduction in the 95^{th} percentile latency. Additionally, it aims for at least 25% more efficient medium usage compared to high-throughput MAC/PHY operations. The IEEE 802.11bn standardization is expected to be completed by 2028 [24, 14]. In the following section, some of the main PHY and MAC layer enhancements in IEEE 802.11be and IEEE 802.11bn are discussed, along with potential open problems in IEEE 802.11bn.

2.4 PHY and MAC Layers Enhancement in Wi-Fi 7 and beyond

2.4.1 PHY Layer Enhancements

In IEEE 802.11be, aiming to enhance network throughput and reduce latency, several new PHY and MAC layer technologies and enhancements have been proposed for

applications such as VR/AR and online streaming. One of the PHY layer upgrades is the introduction of a new channelization and tone plan to support a wider 320 MHz bandwidth, allowing for configurations such as continuous 240 MHz, continuous 320 MHz, and non-continuous 160 + 160 MHz. In addition, since the 2.4 GHz and 5 GHz unlicensed bands are already saturated with legacy users, the 6 GHz band has been opened for the new amendment. Furthermore, for the upcoming IEEE 802.11bn generation, the use of 45 GHz and 60 GHz bands is being explored to support even higher data rates, Fig. 2.1 [25, 8].

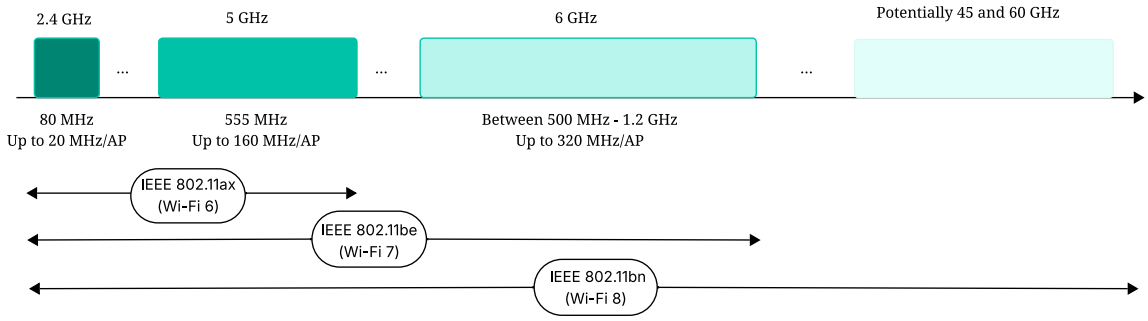
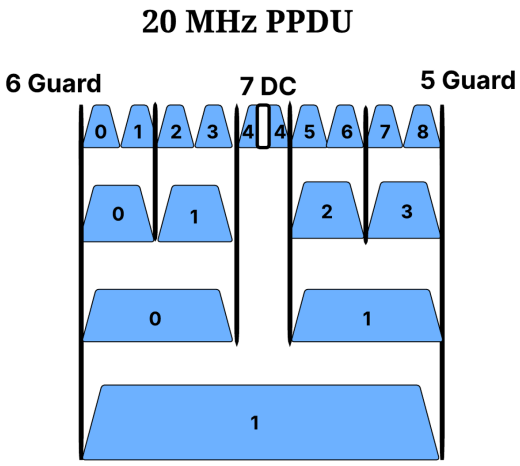


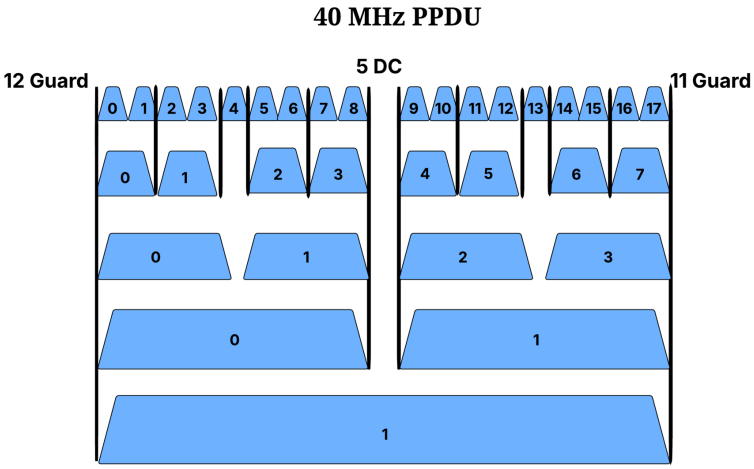
Fig. 2.1 Available transmission bandwidth for IEEE 802.11ax, IEEE 802.11be and IEEE 802.11bn standards [8, 1].

With the primary goal of increasing peak data rates, a higher modulation scheme than that of Wi-Fi 6 (IEEE 802.11ax) has been proposed at the PHY layer. Specifically, 4096-QAM is introduced, which provides approximately a 20% increase in data rate under the same coding rate compared to the 1024-QAM used in Wi-Fi 6. Another key PHY feature in IEEE 802.11bn, as outlined in the project authorization request (PAR), is the introduction of efficient preamble formats and puncturing mechanisms. Preamble formats have been defined across various IEEE 802.11 amendments and serve multiple purposes, including channel estimation, synchronization, signaling, frequency assignment, and resource allocation. Therefore, in Wi-Fi 7 for preamble format, two main requirements are emphasized: first, the novel preamble formats must maintain compatibility and coexistence with legacy preambles across the 2.4, 5, and 6 GHz bands. Second, they must be compatible with future preamble design. Additionally, preamble puncturing is explored, for example, since SU PPDU lacks a SIG-B field, IEEE 802.11be considers preamble puncturing for SU PPDU STAs. Consequently, Wi-Fi 7 preamble design encompasses legacy, EHT, and future preambles [8, 25].

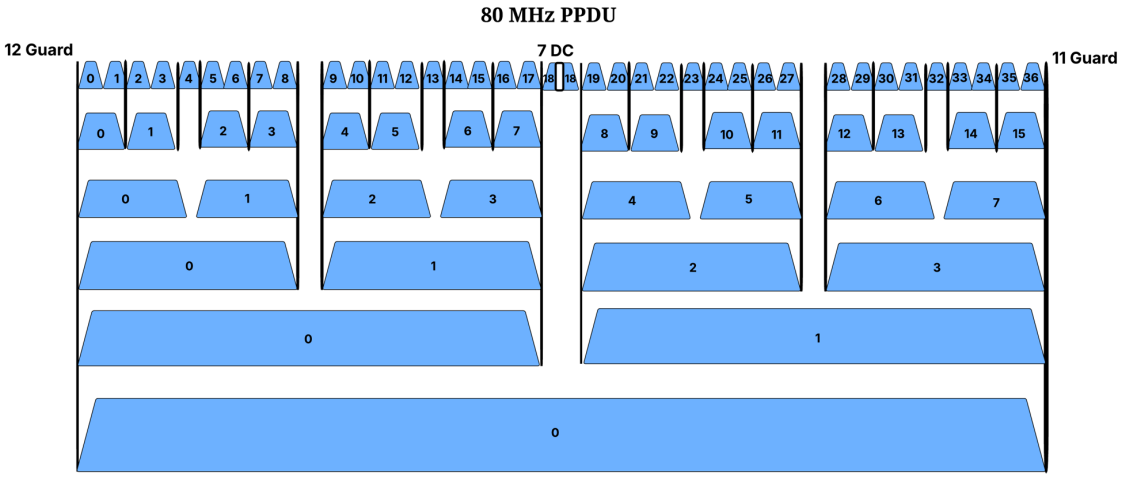
In IEEE 802.11ax, it is assumed that each STA can be allocated only one RU, which may lead to inefficient spectrum utilization and sub-optimal network throughput. To



(a) RUs combination in 20 MHz bandwidth



(b) RUs combination in 40 MHz bandwidth



(c) RUs combination in 80 MHz bandwidth

Fig. 2.2 Illustration of RUs through variable bandwidth [1, 2].

address this limitation, IEEE 802.11be introduces support for multi-RU assignment, allowing each STA to be allocated up to 3 RUs according to standardized RU combination sets [8, 2]. While wider bandwidth allocation increases peak throughput, OFDMA enables efficient support for devices with short and frequent transmissions [2]. Direct conversion (DC) subcarriers are used for purposes such as pilot tones and signaling. For instance, within a 20 MHz bandwidth and 78.125 kHz subcarrier spacing, a total of 242 data tones and 11 guard tones are available. The guard tones serve to separate the frequency channels from its adjacent frequency bands. The variable RU tone sizes that are practically supported for a maximum bandwidth of 80 MHz are summarized in Table 2.3, while the supported RU combinations across bandwidths of 20, 40, and 80 MHz are illustrated in Fig. 2.2.

Table 2.3 Allowed RU combinations in IEEE 802.11be [8]

Types	**Definition**	**Allowed Combinations for IEEE 802.11be**
Small-size RUs	26-tone, 52-tone, and 106-tones	- 26-tone RU + 106-tone RU for 20/40 MHz - 26-tone RU + 52-tone RU for 20/40/80 MHz
Large-size RUs	242-tone, 484-tone, 996-tone, 2×996-tone, 3×996-tone	- 242-tone RU + 484-tone RU for 80 MHz - 484-tone + 996-tone for 160MHz, 242-tone RU + 484-tone RU + 996-tone RU for 160 MHz - 484-tone RU + 2×996-tone RU for 240 MHz, 2×996-tone RU for 240 MHz - 484-tone RU + 3×996-tone RU for 320 MHz, 3×996-tone RU for 320 MHz

The Table 2.3 presents all the applicable multi-RU combinations defined in IEEE 802.11be. In IEEE 802.11bn, most of these combinations are expected to remain the same. However, there are notable exceptions, including the 242-tone RU and the 484-tone RU, which are designated exclusively for full-bandwidth multi-user MIMO. Additionally, it is anticipated that the 3×996-tone RU combination will be retained for puncturing of full bandwidth in multi-user MIMO scenarios.

2.4.2 MAC Layer Enhancements

IEEE 802.11be and IEEE 802.11bn have aimed to introduce enhancements at the MAC layer for high throughput and improved reliability, respectively. Initially, IEEE 802.11be enhanced multi-link operation and improved QoS management features, such as R-TWT. However, several enhancements, including multi-AP coordination, were postponed for IEEE 802.11bn. Consequently, in the initial drafts of IEEE 802.11bn, proposals have been made for distributed multi-link operations, multi-AP coordination, AP power saving, enhanced link adaptation, and retransmission protocols (e.g., hybrid automatic repeat request (HARQ), etc.).

Spatial Stream

IEEE 802.11be introduced the use of 16 spatial streams as a MIMO feature of this amendment to increase simultaneous transmission data rate. However, this enhancement was postponed to IEEE 802.11bn, where several challenges arise, such as channel sounding and feedback reduction complexity, due to the increase in spatial streams from 8 in IEEE 802.11ax to 16 [13, 8].

AP Power Saving

Mobile devices operate on limited battery resources, making energy efficiency a critical design consideration. While advancements in battery technology remain important, it is equally essential to enhance power-saving mechanisms, especially with the proposed improvements for the next Wi-Fi generations. Therefore, for example, ensuring QoS satisfaction for STAs in a multi-AP network with effective power management becomes essential [1, 8, 26].

Multi-link Operations (MLO)

While the previous amendments utilize the 2.4, 5, and 6 GHz channels, their choice of links has been largely random. Therefore, the main objective of this enhancement is to optimize the use of multiple links within a single association. The advantage of this MLO is not only the potential increase in throughput with link aggregation, but also the reduction in latency due to a higher probability of channel access. Additionally, improved reliability can be achieved by enabling multiple link access and separating traffic between the control and data planes across different bands. The MLO-capable APs are referred to as AP multi-link devices (AP MLDs), while the

corresponding non-AP STAs that support MLO are called non-AP MLDs. Multi-link operation can be divided into three main steps: multi-link discovery, multi-link setup, and traffic-to-link mappings.

Multi-link discovery refers to the mechanisms by which an AP MLD detects nearby MLD STAs. This can be achieved through passive scanning (listening on each link for probe response frames), active scanning (transmitting probe request frames on each link), or broadcasting basic information for all APs within the same AP MLD. The next step, multi-link setup, establishes a primary link that facilitates configuration across all other links, where all necessary information is exchanged via request/response messaging among the STAs. While previously each link has its own association request and response. Final step of multi-link operation mapping traffic across links can be applied with associating different TIDs (Traffic Identifiers) with specific links to improve traffic separation and prioritization. Nonetheless, challenges such as non-STR and legacy device blindness, inefficient bandwidth allocation, and unfair channel access relative to legacy STAs remain open issues for this enhancement in IEEE 802.11be [27].

While the MLO in IEEE 802.11be brings high flexibility, in the next generation, the concept of distributed MLO has started to be considered probably for next amendment. This means that although the AP MLD, which consists of a set of adjacent APs, can support MLO, IEEE 802.11bn aims to enable MLO between APs that are located far from each other, which requires coordination between APs. In other words, a distributed virtual cell is proposed to be created in IEEE 802.11bn [14].

Multi-AP Coordination

Multi-AP coordination was suggested initially for IEEE 802.11be, however due to open challenges and its complexity for implementation it has been postponed for IEEE 802.11bn. Multi-AP coordination is a revolutionary MAC layer enhancement that agreed APs for collaboration can have a coordination and signaling to use spectrum more efficiently and mitigate interference effectively.

2.5 Multi-AP Coordination in Wi-Fi

Multi-AP coordination was initially proposed to address collisions in dense environments with high traffic load resulting from inter-BSS interactions. In addition, delays in such scenarios become unpredictable due to signaling among adjacent APs,

which can significantly reduce system reliability [1, 8]. Therefore, as introduced in IEEE 802.11be and further pursued in IEEE 802.11bn, it is suggested that adjacent APs collaborate by sharing information such as scheduling assignments and channel state information (CSI) to enhance network performance. Multi-AP coordination is expected to support features such as resource allocation, interference management, and overall performance improvement while maintaining flexibility for diverse deployment scenarios. Despite these benefits, Multi-AP coordination introduces increased complexity and signaling overhead compared to conventional generations. Hence, it is crucial to develop a practical framework and mechanisms with realistic assumptions to ensure its feasibility.

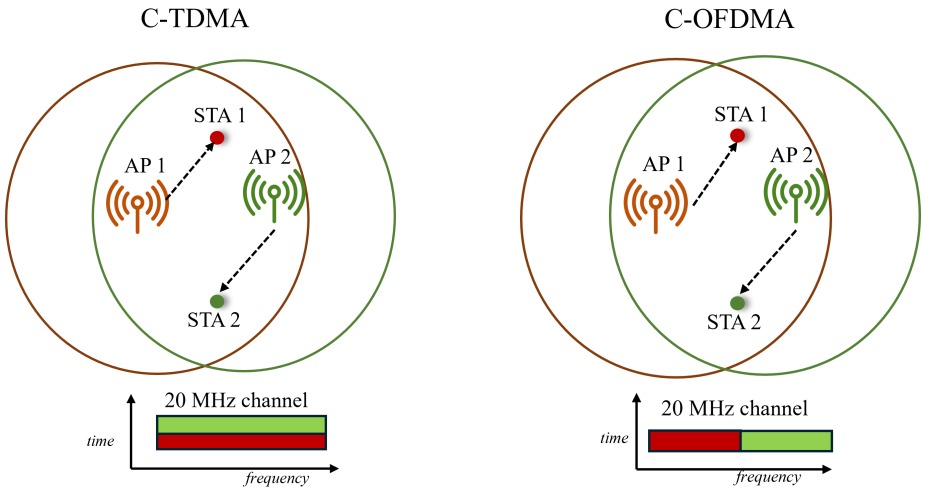
The proposed IEEE 802.11bn multi-AP coordination framework consists of four steps: multi-AP coordination discovery, negotiation, session initiation, and termination. The first step involves determining whether APs intend to collaborate with one another. This intent can be identified through beacons that APs continuously transmit or via probe request/response frames. The second step is the negotiation phase between available APs, during which parameters such as the designation of master and slave APs, maximum transmission opportunity (TXOP) duration, transmit power limits, and other coordination settings are established. This negotiation can occur using management frames, such as public action frames or newly defined individually-addressed action frames. The third step is the multi-AP coordination session, where the master AP, using coordination triggers, notifies the slave APs to initiate the joint coordination based on the previously agreed transmission parameters. The final step is the termination of the coordination session, which occurs when one of the APs transmits a tear-down message [1].

The MAPC mechanism is categorized into four primary schemes: C-SR, C-OFDMA/TDMA, C-BF, and JTX.

2.5.1 Coordinated OFDMA/TDMA (C-OFDMA/TDMA)

C-TDMA is a technique within MAPC that employs time-domain coordination to manage wireless transmissions. In this approach, the coordinator or master AP shares its TXOP with other participating APs in the MAPC. This mechanism enhances network efficiency through improved resource allocation and reduced contention overhead inherent in CSMA-based channel access. As illustrated in Fig. 2.3a, under a MAPC framework, neighbouring APs operating on the same 20 MHz channel coordinate their transmissions in the time domain. Instead of independently contending for channel access, the APs follow a coordinated scheduling decision

whereby transmission time is divided into non-overlapping intervals assigned to different APs. During a given interval, only the scheduled AP transmits while the others remain silent. This time-domain coordination mitigates co-channel interference in the overlapping coverage region.



(a) C-TDMA mechanism (b) C-OFDMA mechanism

Fig. 2.3 Illustration of C-TDMA and C-OFDMA coordination mechanisms in a MAPC framework with two neighbouring APs sharing a 20 MHz channel. (a) In C-TDMA, the master AP (which could be one of the APs based on lower backoff time) divides its TXOP into non-overlapping time slots, assigning each slot to one AP so that only one transmits at a time, eliminating co-channel interference through time-domain separation. (b) In C-OFDMA, the channel bandwidth is partitioned into distinct RU subsets assigned to each AP, allowing both APs to transmit simultaneously over orthogonal frequency segments within the same TXOP.

On the other hand, C-OFDMA enables the coordinator or master AP to share its allocated frequency resources with other coordinated APs. While the other way is a centralized controller to be employed to manage frequency resource allocation for each STA within the MAPC. As illustrated in Fig. 2.3b, neighbouring APs can share the same 20 MHz channel by partitioning it in the frequency domain. Rather than separating transmissions in time, coordinated APs are assigned distinct portions of the channel bandwidth, enabling simultaneous transmissions over different frequency segments. This coordinated bandwidth allocation reduces interference in the overlapping coverage area while maintaining concurrent operation. To facilitate this, essential information such as access category priorities and the traffic load of neighboring APs must be considered to optimize resource allocation. The specific

scheduling algorithms are implementation-dependent and may vary across vendors and system requirements.

However, a notable limitation of both C-TDMA and C-OFDMA is their lack of spatial reuse. Since time slots and frequency bands are taken directly from the sharing AP's TXOP and bandwidth, the overall airtime utilization does not inherently improve. Moreover, performance may degrade in scenarios with mismatched traffic profiles or uneven load distribution, such as in unbalanced OBSS environments. Therefore, integrating C-OFDMA and C-TDMA with C-SR is essential to enable simultaneous transmissions and enhance airtime efficiency.

2.5.2 Coordinated Spatial Reuse (C-SR)

As previously discussed, C-SR enables APs to perform simultaneous transmissions without limiting airtime or bandwidth usage. C-SR introduces a collaborative mechanism in which APs dynamically adjust their transmit power levels to manage inter-BSS interference, which is especially valuable capability in densely deployed wireless environments. Fig. 2.4 demonstrates with adjusting both APs' transmit power control in a shared TXOP, both AP1 and AP2 are enable for concurrent transmissions without compromising link quality. However, in the earlier Wi-Fi 6 standard, only one AP was permitted to transmit at maximum power, while the remaining APs were required to reduce their transmit power. This approach, in certain scenarios, resulted in some STAs being unable to transmit at all due to low power constraints.

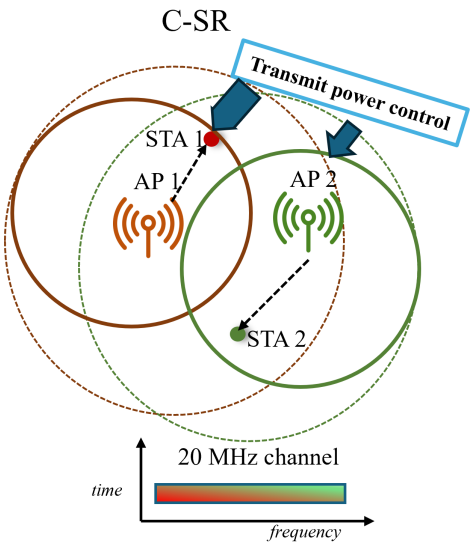


Fig. 2.4 Illustration of the C-SR mechanism showing AP1 and AP2 transmitting concurrently within the same TXOP to their respective STAs in overlapping BSSs. Each AP dynamically adjusts its transmit power based on inter-BSS RSSI, ensuring its signal does not cause harmful interference at the neighbouring AP's STAs. This cooperative power control enables simultaneous transmissions and bandwidth reuse across APs, improving airtime efficiency compared to the Wi-Fi 6 approach where only one AP could transmit at full power.

It is expected that by utilizing information such as the received signal strength indicator (RSSI) of inter-BSS links between the sharing AP and the STAs served by shared APs, the APs can negotiate an appropriate transmission power level. This cooperative adjustment is expected to enhance reliability by reducing packet loss and improve throughput, as it allows all APs to utilize the airtime more efficiently. When C-SR is employed in conjunction with C-OFDMA or C-TDMA, it is expected to scale up the number of STAs effectively served by their respective APs. As it is demonstrated in Fig. 2.4, with a joint C-OFDMA/C-TDMA and C-SR it is expected that the bandwidth be reused between APs.

Nonetheless, transmit power allocation remains an open challenge, particularly when addressing diverse use cases such as EHT and UHR. Key issues include optimal power and resource allocation, ensuring backward compatibility with legacy STAs to avoid significant performance degradation for legacy STAs, and minimizing coordination and signaling overhead [1, 8].

2.5.3 Coordinated Beamforming (C-BF)

C-BF is a spatial-domain coordination technique in which multi-antenna APs actively suppress interference to and from STAs associated with neighboring BSSs. By applying beamforming strategies that direct nulls toward interfering directions, C-BF reduces contention, supports concurrent transmissions, and improves spectral reuse and worst-case latency in dense network deployments.

Despite its benefits, a key challenge in C-BF lies in the channel sounding phase. This process introduces significant overhead, as it requires either all APs in the MAPC sequentially acquire CSI from neighboring APs or perform extensive listening procedures. In addition, null steering reduces the system's degrees of freedom and can limit scalability. Figure 2.5 illustrates the concept of C-BF, where each AP transmits to its associated STAs while simultaneously steering nulls toward STAs served by adjacent APs [1, 8, 14].

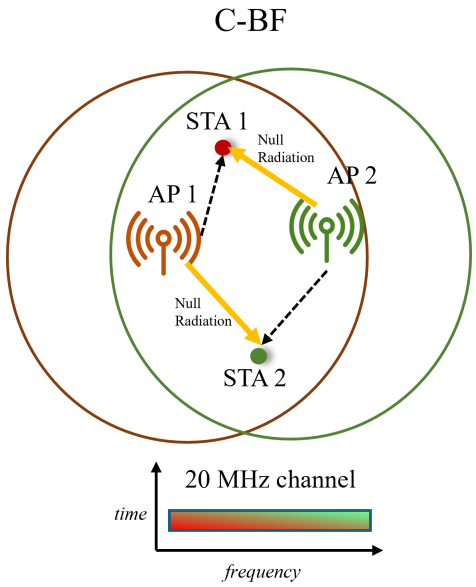


Fig. 2.5 Illustration of the C-BF mechanism showing two multi-antenna APs each directing their main beam toward their intended STA while steering a spatial null toward the STAs of the neighbouring AP. This cooperative null-steering suppresses inter-BSS interference and enables concurrent transmissions across OBSSs without requiring time or frequency separation, at the cost of channel sounding overhead and reduced beamforming degrees of freedom.

2.5.4 Joint Transmission (JTX)

JTX, also referred to as distributed MIMO, is a technique that allows multiple APs, even those that are not in each others neighborhood, to simultaneously transmit to a given STA. For instance, as illustrated in Figure 2.6, while STA1 is primarily served by AP1, AP2 can also retrieve STA1's data via the backhaul and participate in the downlink transmission to STA1 [8].

This approach enhances signal quality and reliability by leveraging spatial diversity. However, JTX introduces significant complexity, including the need for high-capacity, low-latency backhaul communication, distributed CSMA/CA coordination, and accurate synchronization in time, frequency, and phase across the transmitting APs. Due to these stringent requirements, it is likely that JTX will not be included in the IEEE 802.11bn standard [1].

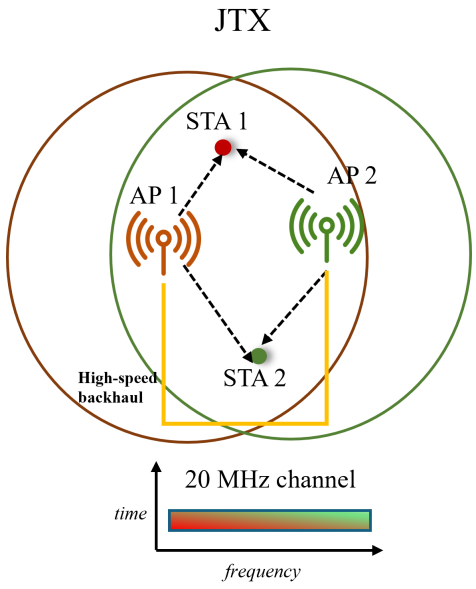


Fig. 2.6 Illustration of the JTX mechanism showing AP1 and AP2 simultaneously transmitting to STA1 within the same TXOP. AP2 retrieves STA1's data via the backhaul and joins AP1's transmission, so that the combined signals reach STA1 with spatial diversity, improving signal quality and reliability. This requires precise time, frequency, and phase synchronisation across both APs along with a high-capacity, low-latency backhaul connection.

2.6 Channel Access Mechanism: A Technical Overview

2.6.1 CSMA/CA

CSMA is the protocol proposed for IEEE 802.11 standard to give all the users in the Wi-Fi network ability to channel access with contention on unlicensed bandwidth without any controller. In this network users normally are called stations (STAs) which are the end nodes that can access the channel directly. CSMA which can also be called distributed coordination function (DCF) protocol, is a channel access technique with a random backoff countdown and an automatic repeat request protocol (ARQ) to handle packet errors and retransmissions.

To explain CSMA/CA transmission setup, each STA in the network needs to check whether the channel is free for at least a distributed inter frame space (DIFS) duration. If it has been free for DIFS, this means the channel is idle. There are two main methods to check whether the medium channel is free, either physical carrier sense or virtual carrier sense. With physical carrier sense STAs listen to the channel and if it detects the overall signal strength on the channel above −62 dBm it detect that the channel is busy. The virtual carrier sense is another way for clear channel assessment (CCA) with the aim of power savings. So with this carrier sensing the STA decodes the PHY header of STAs that have signal strength on the channel above −82 dBm, and the STA sets up a network allocation vector (NAV) [28, 13]. NAV setup is based on what information the STA reads from the PHY header of the other STAs, so it will put the NAV as the total duration of the channel that is busy as the other STA is using the medium. Therefore, NAV can help for physical carrier sensing and energy saving as it does not need to check whether the channel is idle or busy during the NAV setup. In this case the STAs can detect whether the channel is idle or not and based on this attempt to access the channel. However there is a downside for this case, and it is when any two STAs in the network sense the channel is idle and transmit, which can cause collision and packet drop for both STAs.

Therefore, as part of the protocol it is considered to have a random backoff counter for each STA, which is called the collision avoidance strategy with binary exponential backoff (BEB). It is after DIFS and checked during DIFS that the channel is idle, that a STA waits for a random backoff time or counter. The backoff counter is selected with uniform distribution any integer number between 0 and the minimum contention window (CW_{min}). Then the STA checks the channel, if the channel is idle, the backoff counter will be decreased until it is zero, where the STA accesses the channel. In any other backoff time the STA sense the channel is busy it will freeze the backoff

counter until for about DIFS duration the channel could become free again then it will again count down its backoff counter from what it was left. In any case that despite the backoff selection there is any collision, the maximum contention window will be increased by $CW_{max} = 2^m CW_{min}$, where m represents the times that the STA have occurred collision with the largest M time. Consequently, the actual backoff duration is uncertain, not only due to the initial random value, but also because of potential interruptions caused by transmissions from other STAs [1, 28].

2.6.2 RTS/CTS

CSMA/CA despite its high usage as the Wi-Fi protocol, it has a problem with its high delay especially in dense area. This is because once there is a collision, all the TXOP is lost and the STAs need to try again for channel access. In addition, there are hidden STAs in the network that not all the STAs are aware of them. Therefore, it is suggested a new signaling between STAs and APs in the network that after finish-up the backoff timer, then STA instead of whole packet transmission, transmits the request to send (RTS) signal and wait for clear to send (CTS) response from the AP. Once the CTS message is received, it means the duration of the subsequent data transmission (or TXOP) is reserved for that specific STA. This can allow all the STAs in the network that received either RTS or CTS or both to set up their NAV and stop meddling with the winner STA. This NAV can also help other hidden STAs to be informed that the channel is in use, thereby preventing them to further try to access the medium until the current transmission ends [1, 28].

2.6.3 Traffic Differentiation and Prioritization

Another protocol that is available in IEEE 802.11 is the EDCA as a replacement of DCF protocol. It is proposed that we consider four different access categories (AC), including voice (VO), video (VI), best effort (BE), and background (BK), ordered from highest to lowest priority. The difference with the conventional CSMA/CA protocol is that the arrived packets are categorized into the four ACs based on information in the IP headers, in which the DIFS is extended as an arbitrary inter-frame space (AIFS). The AIFS is a distinct value for each AC as well as the $CW_{\min}$ and $CW_{\max}$. In addition, a maximum TXOP duration is specified for each AC with the main goal of limiting the duration of channel occupancy. Higher-priority traffics for example, VO and VI benefit from shorter AIFS and smaller CWs, enabling reduced access delays and improved performance for time-sensitive applications. In scenarios where two

or more ACs finish their backoff period simultaneously, an internal collision occurs. In these cases, the protocol allows the highest priority AC involved in the collision to proceed with transmission. The default access categories mapping based on users' priorities is illustrated in Table.2.4. In addition, the EDCA default parameters are in the Table. 2.5.

Table 2.4 The user priority to AC mapping [9, 10]

Priority	**User Priority**	**AC**	**Traffic Types**
Lowest	1	0 - (AC_BK)	BK
	2	0 - (AC_BK)	BK
	0	1 - (AC_BE)	BE
	3	1 - (AC_BE)	BE
	4	2 - (AC_VI)	VI
	5	2 - (AC_VI)	VI
	6	3 - (AC_VO)	VO
Highest	7	3 - (AC_VO)	VO

Table 2.5 EDCA default parameters, where $CW_{\min} = 31$, $CW_{\max} = 1023$ [9, 11]

AC	$CW_{\min}$	$CW_{\max}$	**AIFS**	**TXOP limit**
0	$CW_{\min}$	$CW_{\max}$	7	-
1	$CW_{\min}$	$CW_{\max}$	3	-
2	$(CW_{\min}-1)/2$	$CW_{\min}$	2	3.0 ms
3	$(CW_{\min}-1)/4$	$(CW_{\min}-1)/2$	2	1.5 ms

2.7 Resource Allocation: A Technical Overview

As mentioned earlier in Wi-Fi two main technologies, MU-MIMO and OFDMA have been proposed for multi-user technologies. In this section the focus is on OFDMA which is also based on the main topic of thesis on OFDMA and C-OFDMA resource allocation. OFDMA helps with the contention time and consequently the spectrum efficiency can increase drastically.

2.7.1 Downlink OFDMA Transmission Procedure

In the downlink OFDMA, the AP after accessing the channel could transmit data simultaneously to multiple STAs. OFDMA is added to the CSMA/CA protocol, therefore, it is important to understand its process.

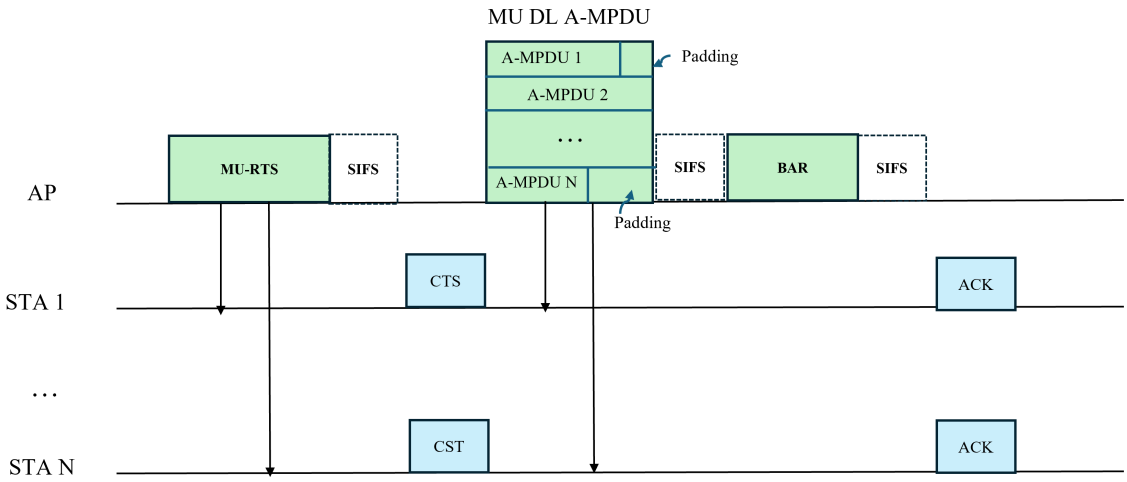


Fig. 2.7 Downlink OFDMA Transmission

As illustrated in Fig. 2.7, a downlink OFDMA process is as follows.

- Initially, the AP, after accessing the channel, the AP sends the MU-RTS signal to all the STAs that are open for OFDMA, and the AP waits for the STAs' response. The MU-RTS set-up the NAV to notify other STAs about that the upcoming TXOP that the channel is busy.
- Then STAs send their CTS signals while all should send simultaneously.
- The next step is that the STAs based on the information on the SIG-B in the preamble know their assigned RUs. Therefore the AP transmits each STA aggregated MAC protocol data unit (A-MPDU). Each STA based on its traffic and channel gain are able to transmit a sp and other traffic and physical layer features are able to transmit a specific A-MPDU, however as all the STAs transmit data should be transmitted simultaneously, the size of MU DL A-MPDU is the maximum A-MPDU of all the STAs. While if there are STAs with lower size that the maximum A-MPDU, it will add zero padding to make it the same size as the MU DL A-MPDU.
- After the MU DL A-MPDU, the AP sends the block acknowledgment request (BAR) to confirm that whether all the STAs received the data.
- Finally, the STAs in case of successful transmission will send back the acknowledgment (ACK) signal [29, 1].

2.7.2 Uplink OFDMA Transmission Procedure

Uplink transmission is not as straightforward as downlink OFDMA. The reason is that as the same synchronization is needed for uplink one, however, as it is uplink transmission the AP is not able to apply synchronization as the downlink OFDMA. Which means more signaling required so the STAs can be synchronized to send their data. It is required that all the STAs' preambles arrive at the AP at the same time, which makes the uplink OFDMA transmission procedure complex. Therefore, it is required that before the preamble transmission of STAs, the AP provides a careful scheduling to specify the exact timing, RU allocation, and other transmission parameters for the STAs.

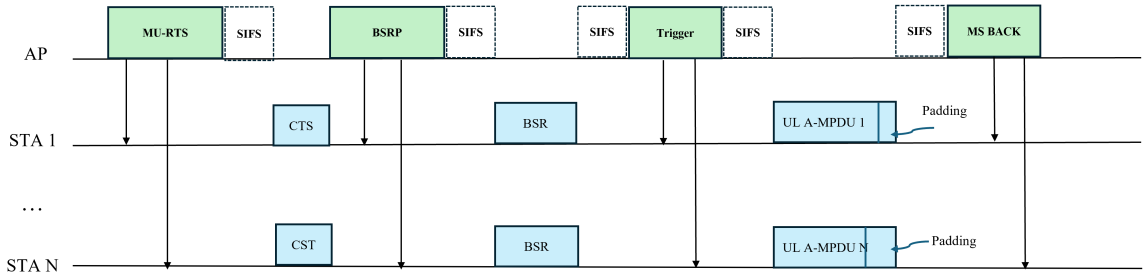


Fig. 2.8 Uplink OFDMA Transmission

As demonstrated in Fig. 2.8 the uplink OFDMA transmission procedure is as follows.

- Initially, the AP sends the MU-RTS trigger frame to the STAs agreed for uplink transmission. The STAs respond with the CTS frame simultaneously.
- Then the AP sends the buffer status report poll (BSRP) to indicate how much data is available in each STA's buffer. The AP waits for STAs' response, which is carried in a Buffer Status Report (BSR), and will help the AP for RU allocation and scheduling planning.
- After receiving the BSR, the AP transmits a Basic Trigger frame to coordinate the uplink transmission. Specifically, the Trigger frame establishes a common transmission start time to ensure time synchronization and prevent symbol overlap. It also enables carrier frequency alignment among participating STAs to minimize inter-user interference across adjacent RUs. In addition, the frame informs each STA of its assigned RU and RU size. Finally, it facilitates transmit power adjustment to ensure proper signal strength alignment at the AP.

- Based on assigned RUs and synchronized timing, the STAs transmit their UL A-MPDU to the AP simultaneously.
- Finally, the AP with multi-STA block acknowledgment, informs the STAs about whether their transmission has been successful [29, 1].

2.7.3 C-OFDMA Transmission Procedure

C-OFDMA in IEEE 802.11bn is still under investigation. The technology has not been standardized yet due to its complexity and probable higher latency on the backhaul [1, 30]. However, there are several proposed transmission procedures in the literature, that we would like to mention in this section.

In MAPC, there are two main blocks, one is the Sharing-AP or called Coordinator-AP, and the other one is the Shared-AP or Coordinated-AP. Sharing-AP in a semi-distributed network architecture is the AP that either wins the TXOP based on CSMA/CA protocol or it is allocated the TXOP based on round robin algorithm for a fair channel access. While the rest of adjacent APs agreed to be in the coordination are called either shared or coordinated-APs. Therefore in the C-OFDMA, once the Sharing-AP has the channel access it could either allocate each STA's RUs in the network or it could assign the bandwidth to the Shared-AP and each Shared-AP. Each set-up needs to have its own transmission procedure and solve the challenges in C-OFDMA. The main question is to propose a transmission procedure for C-OFDMA that while first to be able to provide all the information for coordination, second provide the information to increase the spectrum efficiency with and saving power while improve the simultaneous transmission for the STAs in the network. And finally and the most important that it is reliable with the low latency [2, 1]. It is mentioned earlier even in IEEE 802.11ax that the spectrum usage is not efficient in this standard, therefore with higher complexity in C-OFDMA spectrum usage needs to be improved and give a better RU allocation to improve spectrum efficiency [13].

In [31], the authors proposed a transmission procedure for C-OFDMA in order to meet the IEEE 802.11bn latency requirement. In this paper, it is suggested at the start of each TXOP, the Sharing-AP sends the airtime query report poll (AQRP) to request from the shared-APs some information about the buffer, channel (busy/idle) status of them and the shared-APs access categories (AC) that defines their airtime duration. Then the Shared-APs send back the airtime query report (AQR). For the next step, the Sharing-AP with the announcement trigger frame (ATF) transmission to the Shared-APs, informs the Shared-APs about their assignment resources whether part

of TXOP and their assigned RU. After the ATF, the APs will use the assigned RUs based on the scheduling algorithm and transmit the data with their STAs. In this paper [31], authors are attempting to solve the signaling between the Sharing-AP and the Shared-APs. In this network the Sharing-AP allocates the bandwidth to each Shared-AP and the rest of RU assignment is based on the Shared-APs scheduling algorithms. It should also be noted that this paper uses a semi-distributed network architecture that there is not a general master controller for resource allocation and scheduling.

Authors in [32] proposed a practical system model where a mixture of wired and wireless APs backhaul coordination in C-OFDMA is available. The proposed transmission procedure is quite as [31]. Where in the first step, the AP that wins the channel, transmits the beacon signal to the other APs, and then other APs send back the response beacon signal which is included information about multi-AP group identifier, their channel, STAs in the overlapping area and any other information that can help with APs in a better assessment of the transmission. In the next step the Coordinating-AP, will transmit a trigger frame (TF), for control data exchange. With TF, the information such as the TXOP duration, assigned RU and other coordination information will be exchanged between Coordinating-AP and Coordinated-APs. And finally the data transmission happens. It should be mentioned that in [32], a probabilistic algorithm is proposed where APs get the bandwidth randomly compared to [31] that a scheduling algorithm for RU allocation is suggested.

C-OFDMA transmission procedure still needs to be proven, which is in real time with a complex scheduling algorithm and the synchronization problem is solved, so the problem is still an open problem. Beside transmission procedure of C-OFDMA, there are several challenges in the C-OFDMA needed to be solved, including whether APs can access the channel during multi-AP coordination and whether they are willing to participate in the coordinated AP set. Furthermore, it is critical to determine the appropriate group of STAs to be scheduled together in each TXOP [33, 34]. In [34], an uplink polling problem is introduced, and an algorithm is proposed to efficiently and scalably address polling in high-density OFDMA-based, time-sensitive Wi-Fi networks. Furthermore, the authors in [31] proposed a transmission scheme for the C-OFDMA network to mitigate network latency while reducing overhead and ensuring throughput stability, an essential consideration in the context of C-OFDMA networks.

Another important challenge in C-OFDMA is the absence of a general, optimized RU allocation algorithm. While the RU allocation issue has been extensively

addressed in cellular networks, the problem in Wi-Fi systems is notably different. This distinction arises because OFDMA in Wi-Fi operates on top of the CSMA/CA protocol, which requires an AP to first compete for channel access. Once access is granted, the APs allocate RUs for different AP-STA data exchanges. In contrast, cellular networks have a fixed, traffic-independent time-frequency structure, allowing operators to manage and allocate resources in both time and frequency [35]. This means that a general scheduling algorithm is required with as much as low signalling and overhead, while any complexity might be avoided. In addition, a scheduling algorithm in Wi-Fi network to propose to do the scheduling based on STAs QoS is essential, especially for IEEE 802.11bn due to the new use cases and applications.

2.7.4 Resource Allocation in OFDMA/C-OFDMA

RU allocation in Wi-Fi initially proposed with IEEE 802.11ax, where the OFDMA for the first time used. The main reason was that with increasing the channel width such as 80 MHz, 80 + 80 MHz and 160 MHz, the channel suffers from frequency selective interference. The key concept behind OFDMA is the allocation of adjacent subcarriers into resource units (RUs) for multi-user transmission. By assigning non-overlapping frequency segments to different users, inter-user interference is reduced, which improves the received SINR. The enhanced SINR allows the use of higher MCS levels, thereby increasing the achievable throughput. For IEEE 802.11ax a very strict RU allocation conditions where set based on the standard and it was that only one RU can be assigned to each STA and assigned RUs cannot be reused, which can cause optimized RU allocation non-convex and hard to implement. The used scheduling algorithms are very simple that the vendors either use a random RU allocation or a round robin RU scheduling, that are not optimum way for spectrum usage and causes even lower performance compared to not using OFDMA [13].

IEEE 802.11be addresses the single RU allocation limitation in Wi-Fi 6 with suggesting multiple RU allocation for STAs, which increases the spectrum usage. It is attempted for compatibility assurance and overhead minimization, to propose a predefined RUs combination in IEEE 802.11be/bn standards. Therefore, small-sized RUs 2.3 can only be aggregated with other contiguous small-sized RUs to construct a small-sized MRU. In contrast, large-sized MRUs can be combined with various large-sized RUs, which are not necessarily contiguous. Then as there is an option to use multi-RUs there is the question how to allocated the RUs to the STAs to satisfy STAs throughput and QoS requirement. In addition it is critical that the proposed algorithm can achieve the required QoS while the required information

for the AP's decision for assigned RU. Which is essential in Wi-Fi that to make sure the information is in either in preamble's header and/or in the previous signaling between the AP and its STAs. Another key factor is as all the STAs are transmitting or receiving data simultaneously, it is essential to allocated RUs in an optimum way that the zeros padding for the STAs could be as smaller as possible with the main goal of not wasting airtime which is another resource [1, 15, 8].

C-OFDMA as introduced earlier is one of the MAPC technologies, with respect to the resource allocation in time and frequency. The initial reason for coordination between APs has been to allocated resources more efficiently, however this is not possible without an improved scheduler which has lower overhead and is optimizing time. Therefore the RU scheduling algorithm importance especially with multi-AP is clear. In addition it is essential to have scheduling in scenarios to be able to have as much as possible simultaneous transmission. This could be categorizes as joint C-OFDMA and C-SR [2, 8].

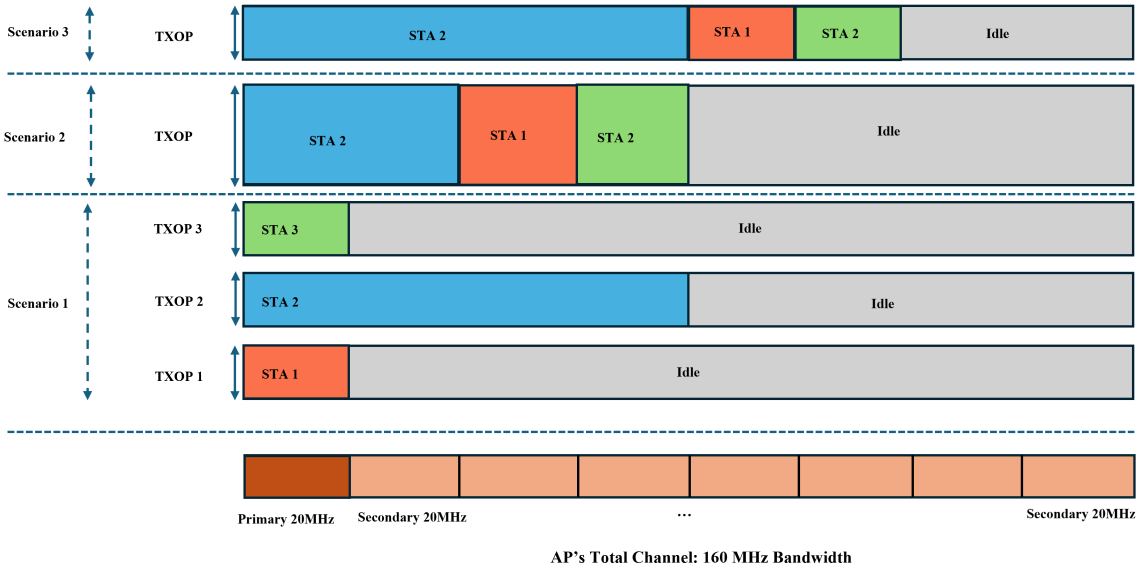


Fig. 2.9 Different RU allocation scenarios.

Fig. 2.9 demonstrates the different scenarios of RU allocation, either in C-OFDMA or in general. With consideration of three different scenarios, the problem is tried to be explained. Let's consider that there are three STAs in the network, STA1, STA2, and STA3 with the potential 20MHz, 80MHz and 20MHz bandwidth usage. Which is also another challenge that either it is the STAs limited bandwidth or whether based on its QoS and its features the AP can define how much RU it needs.

1. First scenario is once there is no OFDMA in the network. As it is demonstrated in Fig. 2.9, the AP allocated 3 TXOPs, each TXOP for one of the STAs. As the STAs capability is limited, in each TXOP only partial of bandwidth will be used. This a huge waste of time and frequency, especially with the increased bandwidth up to 320 MHz per an AP.

2. Then in the second scenario, it is tried to allocate RUs with limitation of maximum usage of 80 MHz of the channel, which can cause the TXOP duration larger than before. Then the rest could be allocated to other APs in C-OFDMA.

3. For the third scenario, it is used OFDMA with considering allocate bandwidth more to STAs while the TXOP duration can be the same as before. Where the rest could be allocated to the other APs' STAs.

These scenarios, demonstrates the importance of scheduling for different applications in the network and different use cases for STAs.

2.8 Background and Prior Studies

IEEE 802.11bn with the main goal of high reliability and lower latency compared to previous standardization especially with the new applications has proposed a variety of technologies and enhanced ones to reach to its goal. One of the new enhancements is to be able to optimize the back off time duration for channel access, which in the section 2.8.1 will be discussed. In addition multi-user resource allocation in multi-AP coordination especially C-OFDMA is an important challenge which is discussed in 2.8.2.

2.8.1 Channel Access

As mentioned earlier in section 2.6.1, CW or the backoff time in CSMA/CA protocol has an essential role for collision avoidance. However in overcrowded network, the network ability to give access to each and every STA in the network is not possible and therefore the network might collide. In addition to the new technologies such as AR/VR the requirement of lower latency and higher reliability compared to other conventional applications is essential. So it is important to look further with the main goal of decreasing backoff time or propose a new set of protocols for channel access to degrade the wasting time for channel access. Authors in [36] propose a

MARL to optimize the CW within the existing CSMA/CA framework. The authors claimed reducing tail latency for UHR, while maintaining throughput and MPDU loss. The work in [37] proposed to replace the conventional binary exponential backoff mechanism with MARL. The main addressed problem was the channel access of different types of RL agents, specifically, value-based or policy-based methods with the main objective of throughput maximization with fair allocation. It is proposed that the AP manages the centralized training phase, while individual stations perform decentralized execution.

Authors in [38] proposed a novel MAC protocol named QMIX-advanced Listen-Before-Talk (QLBT), which is built upon MARL. The proposed MARL algorithm outperformed CSMA/CA in terms of throughput and jitter with robustness in dynamic environment which can also have favorable channel access competence with legacy STAs. While in [28], a MARL algorithm for the APs in a multi-AP coordination network is proposed, where each agent can decide when to transmit. To mitigate the problem arising when only one AP accesses the channel, a centralized controller is proposed, which first, enhances learning performance, and second, a greedy algorithm for proportional fairness is proposed. Authors in [39] proposed both a centralized DRL and a distributed federated DRL approach to optimize the energy detection (ED) thresholds in a heterogeneous NR-U/Wi-Fi network, with the goal of maximizing uplink system throughput while ensuring fairness. The authors considered the channel access of both Wi-Fi and NR-U networks according to the already existing protocols in their networks. However, with DRL, it is expected that both networks with ED thresholds can dynamically configure a key parameter within these existing channel access schemes. Also, the fairness is applied as a reward or punishment if Wi-Fi throughput drops below a certain threshold, and FL is used as a solution for data privacy concerns.

2.8.2 Resource Allocation

According to the IEEE 802.11bn task group, MAPC is a feature to be included in IEEE 802.11bn to enable APs coordination across different BSSs. Compared to previous standards, this innovation significantly enhances the interference mitigation and leads to better spectral efficiency and higher reliability [40–43, 2]. MAPC identifies four advanced techniques pivotal to MAPC: C-SR, C-BF [40],C-OFDMA, and JTX [40, 18, 44]. These techniques play crucial roles in enhancing the performance of networks [45, 44].

C-OFDMA from the MAPC categories can provide simultaneous channel access with allocating bandwidth to all available STAs at different APs with coordination based on frequency separation, thus the interference will be managed efficiently compared to the conventional non-MAPC models [40]. C-OFDMA can improve the potential interactions between devices in an OBSS, so that the devices can coordinate to leverage the same TXOPs. In other words, C-OFDMA is introduced to improve spectral efficiency. On the other hand, C-SR provides parallel transmission with interference management [18] in the same frequency and time with coordination between APs and transmission power management of APs/STAs. Using C-OFDMA with C-SR can highly increase the spectrum efficiency. By coordinating RU allocation, multiple APs can transmit concurrently on the same frequency channels without causing significant interference, leading to more efficient use of the available spectrum [2]. Interference management through RU assignment tackles a fundamental spatial reuse issue. It enables concurrent transmissions in overlapping regions by employing efficient spectrum partitioning and interference mitigation strategies.

In addition, given the dynamic nature of wireless channels and the limited spectrum availability, RU allocation plays a crucial role, particularly in multi-AP scenarios. Without an efficient RU allocation strategy, interference can significantly degrade network performance [2]. Under C-OFDMA, APs and STAs are permitted to share OFDMA resources within each TXOP, thereby improving spectrum efficiency. In addition, with the proposed ability of multi-RU assignment in IEEE 802.11bn multi-users resource allocation with multi user opens up new possibilities and challenges to be addressed. Consequently, RU allocation and scheduling algorithms for RU allocation in Wi-Fi systems are important compared to all the previous amendments and standard. It is critical to be able to use spectrum efficiently and manage interference in a C-OFDMA resource allocation.

The authors in [46] propose the variable RU allocation for the IEEE 802.11ax standard with the main objective of maximizing throughput with delay degradation. The proposed algorithm is based on traffic backlog of STAs and the congestion of STAs in a single AP scenario. In [32], a multi-AP C-OFDMA network is proposed for industrial IoT networking environments. In which a probabilistic algorithm for RU allocation is proposed to mitigate latency mitigation, however, as the APs and STAs increase the network jitter degraded and the latency is increased, which shows the importance of a more intelligent algorithm. In [47] a fixed-RU allocation in joint C-OFDMA and C-SR is suggested based on graph coloring and Weighted Max-min algorithm. Authors in [48] proposed a fine-grained RU allocation to bring low latency

and ultra reliability of IEEE 802.11bn. However, the authors due to high overhead for a fine-grained RU allocation and also due to the synchronization proposed to use backhaul wire instead of over-the-air coordination of APs. The authors in [49] also proposed fine-grained RU allocation in C-OFDMA for industrial Wireless time sensitive network (TSN) scenarios. In order to reduce overheads a virtual sounding approach in which, rather than acquiring CSI for every link in each cycle is proposed. Measurements are taken for only a subset of links, while the remaining CSI values are estimated through interpolation based on previous measurements. In [50], the authors analyze the impact of different distributions of random and scheduled RU allocation on the performance of the MAC layer. Additionally, in [16], the fixed-size RU allocation problem is addressed by re-framing it as a channel access problem, with the primary goal of maximizing throughput while ensuring fairness. Moreover, authors in [17] proposed a random RU selection based ϵ-greedy, where each STA is an agent and tries to search for the random RU selection between all the available RUs. However, the random RU allocation specially in a high congested network can lead to collision which is the lost of resources. On the other hand, scheduled RU allocation scenario with needs to consider the for RU assignment in order to optimizes spectrum efficiency and the transmission parameters. As without no effective RU scheduling the network's performance can be degraded [51].

Therefore it is essential to understand the tradeoff between bandwidth allocation, time allocation or, in general, the resource allocation in the C-OFDMA. For example, in a case with variable buffer size for each AP assigning the same bandwidth with different primary channel is not optimum and this will be the loss of resources. While, if APs use the same primary channel for transmission without RU allocation, this can cause collisions and huge loss of resources and data [2, 15].

Considering allocating the total bandwidth to the STAs with the higher data rate has promised to maximize the total network throughput. However, in this case the system is not fair. Fairness can be defined in some classical definition as proportional fair (PF), modified largest weighted delay first (MLWDF), Max-min fairness and etc [52]. Authors in [52] propose FairSplit algorithm as a lightweight and near-optimal bandwidth splitting strategy with the following of [53]. It achieves significantly improved throughput and fairness compared to existing methods, while remaining practical and deployable in real-world systems. Considering real scenarios the STAs have different factors that can effect the network performance and it is essential to be considered during resource allocation consequently RU allocation. Therefore proportional fairness can be defined as a factor to be able to allocate resources to STAs

proportionally fair with consideration of STAs related factor on resource allocation [54].

Proportional fairness is defined as a metric to ensure a balance of efficiency and fairness. It is defined as a resource allocation principle that strikes a balance between efficiency (maximizing throughput) and fairness (ensuring equitable access) among multiple users or devices. Proportional fairness allows efficient trade-offs between users and is more throughput-friendly, while max-min fairness ensures strict equality at the expense of performance [55]. Moreover with the new IEEE 802.11bn assumptions the optimal fair multi-RU allocation with STAs' heterogeneity in a C-OFDMA system model is an open problem. This challenge must account for the realistic information required to enable effective multi-AP coordination [46]. Also the interference with the same RU allocation in a TXOP caused in the multi-AP coordination it is important to delve into the scheduling algorithm. Moreover with the use-cases such as extended reality bring the specific use cases for Wi-Fi network and the fairness definition with network throughput optimization needs essential RU management based on their use cases and QoS requirements [56].

One of the proposed algorithms for RU scheduling and random RU allocation that has been recently employed is the use of machine learning and deep learning techniques [16, 17]. Among these approaches, reinforcement learning (RL) is regarded as particularly effective for adapting to dynamically changing and uncertain environments of wireless communications [56–58]. In this regard, RL has been shown to be efficient for intelligently allocating resources based on STA requirements and QoS. In [17], a distributed RU selection method utilizing convolutional neural network (CNN)-based Multi-agent deep reinforcement learning (DRL) is proposed to enhance the throughput and reduce the latency of wireless networks. The authors have not considered the variable RU scheduling scheduling algorithm, nor any differences between STAs QoS. Authors in [16] drawing inspiration from mean-field game theory proposed, a novel mean-field-based MARL algorithm for uplink OFDMA in IEEE 802.11ax networks. This algorithm aims to learn the optimal channel access strategy, thereby improving network performance in high-density environments. However, RU scheduling algorithm with the new IEEE 802.11bn that each STA can get more than one RU once again brings the challenges of RU allocation specifically with diversity factors such as different channel gain of STAs, a C-OFDMA. In which also the coordination brings the flexibly to manage RU allocation in more intelligent way and increase the spectrum allocation optimized.

2.9 Summary

This chapter has critically examined the evolution of Wi-Fi technology with the objective of identifying the technical limitations that justify the research problems addressed in this thesis. The review has exposed a structural gap between the increasing PHY-layer capabilities of modern IEEE 802.11 amendments and the relatively underdeveloped MAC-layer coordination and resource allocation mechanisms required to fully exploit them.

While IEEE 802.11be and the emerging IEEE 802.11bn introduce advanced features such as wider bandwidths, multi-RU allocation, multi-link operation, and multi-AP coordination, the underlying channel access mechanism in Wi-Fi remains fundamentally contention-based. In dense and interference-limited deployments, CSMA/CA introduces randomness, unpredictable latency, and fairness degradation. Consequently, simply enhancing modulation schemes or bandwidth does not guarantee improved reliability or delay performance.

The review further reveals that existing studies on OFDMA and multi-AP coordination primarily emphasize throughput maximization under idealized assumptions, often neglecting heterogeneous channel conditions, fairness objectives, traffic dynamics, and queue-aware scheduling. In particular, practical and scalable RU allocation strategies for coordinated multi-AP environments remain insufficiently addressed. This disconnect between theoretical coordination gains and implementable scheduling solutions constitutes the central research gap identified in this chapter.

Therefore, the main purpose of this literature review has been to justify the need for intelligent channel access and resource allocation frameworks capable of meeting the ultra-high reliability and low-latency objectives of next-generation Wi-Fi systems. The limitations identified here directly motivate the contributions of this thesis: reinforcement learning-based channel access optimization, interference-aware and fairness-driven RU allocation in C-OFDMA systems, and queue-aware scheduling mechanisms for realistic traffic conditions.

In this context, learning-based scheduling has recently attracted attention because conventional MAC-layer decision mechanisms in Wi-Fi often rely on fixed heuristics that cannot fully react to rapidly changing traffic and interference conditions. In particular, DRL has emerged as a promising approach for wireless scheduling problems where resource decisions must be taken sequentially under uncertain network evolution. Its relevance to this thesis lies primarily in the later chapters,

where RU allocation must adapt online to heterogeneous channel conditions and queue dynamics while maintaining low scheduling latency.

In summary, this chapter establishes that achieving the performance targets of IEEE 802.11bn requires not only PHY-layer enhancements, but fundamentally improved coordination and scheduling strategies at the MAC layer. Addressing this requirement forms the core contribution of the subsequent chapters.

Chapter 3

Mathematical Background

This chapter provides an overview of the mathematical background for the resource allocation problems considered in this thesis. First, the fairness-aware resource allocation scheme and performance metrics used in the thesis are presented. Then, the mathematical techniques applied for developing resource allocation techniques, including convex optimization and ML are introduced. Finally different ML algorithms with the main goal of online resource allocation to achieve a better performance are presented.

3.1 Fairness Metric

The importance of RU allocation in the C-OFDMA is presented in the previous chapter, therefore the question arises as to which performance metrics should be used when proposing RU allocation and scheduling algorithms. One of the main objectives in the wireless network is the throughput enhancement. The outnumbered users and high demand over the decades still make the throughput enhancement one of the most significant challenges in the Wi-Fi network. In addition to throughput performance, users have diverse requirements, demanding different level of QoS satisfaction, such as distinction on reliability, and latency. Consequently, the system must incorporate fairness across various performance metrics while meeting the heterogeneous QoS requirements of users. Hence, scheduling algorithms need to jointly optimize multiple objectives, including throughput and fairness, during resource allocation.

In other words, whether it is throughput maximization and/or delay degradation or any other utility functions optimization, the simplest way for resource allocation

could be giving the more resources to the user with the fastest channel or the one that has the highest amount of data to transmit. So the throughput could be reached to the highest. However, in real scenarios where the network is overcrowded, severe dissatisfaction occurs among all users except the one who receives all available resources. In another case allocating more resources than needed to any user could be wasteful. Therefore, this confirms the importance of fairness for resource allocation. In simple systems, generally to achieve fairness random resource allocation might be used, however this is not an optimal approach for resource allocation. It also degrades the overall network-wide performance, thereby preventing the system from operating at its full potential. Another way of resource allocation could be the equally allocating resources between users, however this will cause collisions or excessive interference if the number of users exceeds available units [59–61].

Therefore, fairness is a critical performance metric that must be considered in the design of practical resource allocation mechanisms. However, fairness is inherently difficult to formulate and quantify. Its definition may vary depending on the network perspective, user requirements, or environmental conditions. In the following sections, the fairness metrics adopted throughout this thesis are presented and discussed [59, 60, 62–64].

3.1.1 Max-Min Fairness

Max-min fairness (MMF) is a fundamental concept to ensure that users in the network achieve a similar performance, aimed at ensuring that the user receiving the least resource is treated as fairly as possible. It seeks to balance efficiency with fairness, particularly in environments where resources are limited and multiple users compete for allocation.

Formally, let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N) \in \boldsymbol{\Lambda}$ represent a vector of assigned resources for N users. The vector $\boldsymbol{\lambda}$ is said to be max-min fair if it satisfies two conditions:

1. Feasibility: $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is the set of feasible allocations.

2. Fairness: For any alternative feasible allocation $\lambda' \in \boldsymbol{\Lambda}$, if $\lambda'_s > \lambda_s$ for some $s \in \{1, \ldots, N\}$, then there must exist at least one $t \in \{1, \ldots, N\}$ such that $\lambda_t \leq \lambda_s$ and $\lambda'_t < \lambda_t$.

This condition ensures that no user can increase its performance without decreasing the allocation of someone else with an equal or smaller share.

Now, consider a decision vector $\boldsymbol{x} \in \boldsymbol{S}$, where $\boldsymbol{S}$ is a set of feasible decisions. Here, the decision vector

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$$

represents the allocation of resources to the N users in the network, where each component x_i determines the resources assigned to user i. Each user $i \in I$ derives a utility $U_i(\boldsymbol{x})$ based on the decision $\boldsymbol{x}$, which is defined to quantify the user's utility. The utility vector is defined as $\boldsymbol{u} = (u_1, u_2, \ldots, u_N)$, where $u_i = U_i(\boldsymbol{x})$. In this setting, the max-min fairness problem can be formulated as the following optimization problem:

$$\max_{\boldsymbol{x} \in \boldsymbol{S}} \left(\min_{i \in I} \{U_i(\boldsymbol{x})\} \right). \tag{3.1}$$

This objective seeks the decision $\boldsymbol{x}$ that maximizes the utility of the user who receives the smallest utility, aligned directly with the max-min fairness condition described above.

A common approach to solving a max-min fairness problem is the progressive filling algorithm, also known as water-filling [65]. This algorithm incrementally allocates resources, prioritizing the users with the lowest current allocation. At each iteration, resources are distributed equally among all unsaturated demands, that is, users whose current allocation has not yet reached their maximum allowed value until one or more reach their maximum allocation. These saturated demands are then fixed, and the process continues with the remaining users [66].

3.1.2 Proportional Fairness

While MMF aims to ensure that the user with the lowest allocation is treated fairly, proportional fairness offers alternative fairness criteria that attempt to bring a balance between overall system efficiency and individual fairness. Rather than focusing only on maximizing the minimum utility, proportional fairness considers the relative gains and losses across all users.

Let $\boldsymbol{x} \in \boldsymbol{S}$ be a feasible decision vector, where $\boldsymbol{S}$ is the set of all feasible decisions, and let each user $i \in I$ derive a utility $U_i(\boldsymbol{x})$ from decision $\boldsymbol{x}$. A feasible allocation $\boldsymbol{x}$ is said to be proportionally fair if for any other feasible allocation $\boldsymbol{y} \in \boldsymbol{S}$, the aggregate proportional change in utility is non-positive:

$$\sum_{i \in I} \frac{U_i(\boldsymbol{y}) - U_i(\boldsymbol{x})}{U_i(\boldsymbol{x})} \leq 0. \tag{3.2}$$

This definition ensures that no feasible deviation can increase the total proportional utility gain. In other words, any attempt to improve the utility of one or more users must come at a cost to others.

A standard approach for achieving proportional fairness is to maximize the weighted sum of the logarithmic utilities:

$$\max_{x \in S} \sum_{i \in I} w_i \log \left(U_i(\boldsymbol{x}) \right), \tag{3.3}$$

where $w_i > 0$ is a weight that can reflect the relative priority or payment of i-th user. When all weights are equal, this objective is equivalent to maximizing the geometric mean of utilities and satisfies the proportional fairness condition [55].

This formulation is relevant in network systems, such as bandwidth allocation, where each user contributes an amount m_i, and receives service at a rate proportional to this contribution. If the utility function is taken as $U_i(x_i) = \log x_i$, and the feasible set S is defined by resource constraints such as $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{C}$ and $\boldsymbol{x} \geq 0$, then the optimization problem ensures a proportionally fair allocation:

$$\begin{aligned} \max_{x_i} & \sum_{i \in I} m_i \log x_i \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{C}, \\ & \boldsymbol{x} \geq 0, \end{aligned} \tag{3.4}$$

where $\boldsymbol{A}$ is a resource usage matrix, and $\boldsymbol{C}$ is a vector of resource capacities.

In contrast to max-min fairness, which may enforce equality at the expense of total utility, proportional fairness provides a more flexible trade-off between fairness and efficiency. This makes proportional fairness suitable for systems where users contribute unequally or where maximizing aggregate utility is a priority, as long as individual allocations remain reasonably equitable [59, 60, 55].

3.1.3 α-Fairness

α-fairness was introduced to integrate the trade-off between efficiency and fairness under a single framework, generalizing both proportional fairness and MMF [67]. It provides a flexible mechanism for tuning the level of fairness in resource allocation systems.

Let $\boldsymbol{x} \in \boldsymbol{S}$ be a feasible decision vector, where $\boldsymbol{S}$ is the set of feasible allocations, and let user $i \in I$ receive a utility based on their allocated resource x_i. The α-fairness criterion assigns a utility function $U_i(x_i)$ of the following form:

$$U_i(x_i) = f_\alpha^{(i)}(x_i) = \begin{cases} \log(x_i), & \text{if } \alpha = 1 \\ \frac{x_i^{1-\alpha}}{1-\alpha}, & \text{if } \alpha \neq 1 \end{cases} \tag{3.5}$$

Here, the parameter $\alpha \in [0, \infty)$ manages the trade-off between total system throughput and fairness across users:

- When $\alpha = 0$, the utility function reduces to $U_i(x_i) = x_i$, corresponding to the fairness, where the objective is to maximize the total throughput $\sum_{i \in I} x_i$.
- When $\alpha = 1$, the model regains proportional fairness, as maximizing $\sum_{i \in I} \log(x_i)$ ensures no feasible alternative can maintain the proportional gain.
- As $\alpha \to \infty$, the solution converges to MMF, in which the objective becomes to maximize the utility of the worst-off user.

Thus, α-fairness covers the gap between efficiency and fairness. Smaller values of α prioritize aggregate performance, while larger values emphasize equitable distribution.

The general optimization problem under α-fairness can be formulated as:

$$\max_{x \in S} \sum_{i \in I} f_\alpha^{(i)}(x_i), \tag{3.6}$$

subject to relevant constraints (e.g., $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{C}$, $\boldsymbol{x} \geq 0$), where $\boldsymbol{A}$ is the resource usage matrix and $\boldsymbol{C}$ denotes resource capacity vector. Specifically, each element a_{ji} of $\boldsymbol{A}$ quantifies the amount of resource j consumed per unit allocation to user i, while each component C_j of $\boldsymbol{C}$ specifies the maximum available amount of resource j in the system. The choice of α allows system designers to adjust the degree of fairness required by the application while maintaining efficiency and understandability.

3.1.4 Jain's Fairness Index

Beyond optimization-based fairness criteria such as max-min, proportional, and α-fairness, fairness can also be assessed through quantitative metrics. One widely used metric is Jain's fairness index [59], which provides a scalar measure of how equitably resources are distributed among users.

Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_N)$ be the vector of utilities assigned to N users, where each $u_i = U_i(x)$ denotes the utility obtained by user $i \in I$ under decision $\boldsymbol{x} \in S$. The Jain's fairness index J is defined as follows:

$$J(u) = \frac{\left(\sum_{i=1}^{N} u_i\right)^2}{N \sum_{i=1}^{N} u_i^2}, \tag{3.7}$$

where N is the total number of users. The index is bounded as follows,

$$\frac{1}{N} \leq J(u) \leq 1. \tag{3.8}$$

A value of $J = 1$ indicates perfect fairness, where all users receive equal utility. Conversely, $J = 1/N$ reflects the most unequal distribution, where only one user receives all resources. In general, if k out of N users are allocated equal utilities while others receive none, then $J = k/N$, making the metric both interpretable and intuitive.

Jain's index is closely related to the coefficient of variation (COV), which measures the relative dispersion of a dataset around the mean. Specifically, $J = 1/(1 + \mathrm{COV}^2)$, if the data is more spread out or inconsistent (high COV), fairness as measured by J drops. If the data is consistent (low COV), fairness J is higher [59]. Due to its bounded and normalized form, Jain's index is frequently used as a diagnostic tool to evaluate the fairness of outcomes in resource allocation problems, independent of the underlying optimization objective [59].

Fairness importance for practical system models is proven [68]. There is always the question about the resource allocation such as time, frequency, power, and any other resources, to provide the fairness that the system requires. Depending on the chosen fairness criterion, the system may either balance fairness and efficiency or enforce strict fairness across users, potentially sacrificing overall efficiency. In this thesis, an attempt is made to answer these questions for RU allocation, spatial reuse, and the channel access in the new C-OFDMA and C-SR technologies in Wi-Fi networks (IEEE 802.11bn). A summary of the main fairness criteria, including their principles, key characteristics, and weaknesses, is provided in Table. 3.1.

In this thesis, fairness is incorporated both as an optimization objective and as a performance evaluation metric. Specifically, proportional and max-min fairness principles are considered in the resource allocation design, depending on the system objective, while Jain's fairness index is employed to quantitatively assess and compare the fairness performance of the proposed algorithms. This approach

enables a comprehensive treatment of fairness from both theoretical and practical perspectives.

However, to propose fair resource allocation algorithms, there are several challenges, including the complexity of the problem, which makes it difficult to give an optimum answer. In order to provide an optimal solution, convex (CVX) optimization has been used for a long time. In addition, it is possible to have an approximation of optimization problem as a CVX problem, or some of the constraints are relaxed as a promising way to solve the NP-hard optimization problems. In the next section, an introduction of this optimization is explained, which is used throughout the thesis.

Table 3.1 Summary of Fairness Indices

Fairness Criterion	**Definition / Principle**	**Key Characteristics**	**Weaknesses**
MMF	Maximizes the allocation of the user with the smallest share; ensures the least-advantaged user is treated fairly.	Focuses on equity for the weakest user; guarantees a minimum level of service.	Can significantly reduce total system throughput; not efficient when resources are abundant.
PF	An allocation is proportionally fair if any change that increases one user's utility results in a proportionally greater decrease for others.	Balances fairness and efficiency; widely used in network bandwidth allocation.	Requires well-defined utility functions; may be complex in large-scale systems.
α-Fairness	A parameterized family of utility functions that generalizes MMF and PF, providing a tunable fairness-efficiency trade-off.	Flexible: $\alpha = 0$ corresponds to throughput maximization, $\alpha = 1$ to PF, and $\alpha \to \infty$ to MMF.	Choice of α is subjective; different values can lead to very different allocations.
Jain's Index	A scalar measure of fairness based on the distribution of resources across users; bounded between $1/N$ and 1.	Simple, normalized, and easy to compute; independent of optimization framework.	Only measures fairness; does not suggest an allocation strategy or account for efficiency directly.

3.2 Convex Optimization

A resource allocation problem can be formulated as an optimization problem. An optimization problem proposes to seek an optimal solution or the best possible action from a set of actions to give the best choice [3]. To solve an optimization problem, a solution method which is an algorithm based on an instance of a problem can be proposed. Since the 1940s, algorithms have been developed for various class of optimization problems. However, in order to solve these optimization problems, there are several factors to be considered. For example, the mathematical structure of the objective function (e.g., linear, convex, or non-convex), number of variables and constraints that are available, and special structure, such as sparsity.

Even with simple objective and constraints, the general optimization problem is difficult to solve. To solve these problems normally needs compromises, such as very long computation times or the possibility of not finding the actual global optimum. Convex optimization and non-convex optimization represent two major classes of optimization problems, distinguished by the properties of their objective and constraint functions. In non-convex optimization, a locally optimal point is not guaranteed to be globally optimal [3]. This is a key difference from convex problems. In the following sections, some fundamental concepts of convex optimization are provided.

3.2.1 Convex Sets

A convex set is a fundamental concept in mathematics and optimization. A set C is defined as convex if, for any two points $x_1 \in C$ and $x_2 \in C$, the entire line connecting these two points also lies within C. This can be expressed as,

$$\theta x_1 + (1-\theta)x_2 \in C, \forall x_1, x_2 \in C, \forall \theta \in [0,1]. \tag{3.9}$$

Geometrically, this means a convex set has a solid body without holes or inward curves. Fig. 3.1 illustrates this clearly, Fig. 3.1a is a hexagon (including its boundary) that is convex, whereas a kidney-shaped set or a square that includes some boundary points but not others in Fig. 3.1b would be considered non-convex because a line segment between two points within the set might fall outside its bounds [3].

(a) A convex set. (b) A non-convex set.

Fig. 3.1 Illustration of convex and non-convex sets [3].

Understanding how convexity is retained through various mathematical operations is crucial, especially in the context of optimization. Certain operations

performed on convex sets result in new sets that are also convex, thereby preserving this important property. These operations include:

1. Intersection: If the sets $C_1, C_2, \ldots, C_n$ are convex, then their intersection C is also convex, where
$$C = C_1 \cap C_2 \cap \cdots \cap C_n. \tag{3.10}$$

2. Affine Transform: If a set $C \subseteq \mathbb{R}^n$ is convex, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is a matrix, and $\boldsymbol{b} \in \mathbb{R}^m$ is a vector, then the affine transformation $\boldsymbol{A}C + \boldsymbol{b}$ is also convex, where
$$\boldsymbol{A}C + \boldsymbol{b} = \{\boldsymbol{A}x + \boldsymbol{b} \mid x \in \boldsymbol{C}\}. \tag{3.11}$$

3. Perspective Transform: If the set $C \subseteq \mathbb{R}^{n-1} \times \mathbb{R}_+$ is convex, then the perspective transformation $f(C)$ is also convex, where the mapping $f : \mathbb{R}^n \to \mathbb{R}^{n-1}$ is defined by
$$f(C) = \left\{ \left[\frac{x_1}{x_n}, \frac{x_2}{x_n}, \ldots, \frac{x_{n-1}}{x_n} \right]^T \mid (x_1, \ldots, x_n) \in C \right\}. \tag{3.12}$$

Here, x_i denotes the i-th element of a vector in C and $\mathbb{R}_+$ represents the set of positive real numbers. These preservation properties form the basis for constructing complex convex sets from simpler ones and support the design of convex feasible regions in optimization problems [3].

3.2.2 Convex Cones

Another key concept in convex optimization is the convex cone, which represents a particular subclass of convex sets. A set C is referred to as a convex cone if, for any x_1, $x_2 \in C$ and any non-negative scalars $\theta_1, \theta_2 \geq 0$, the linear combination $\theta_1 x_1 + \theta_2 x_2$ also belongs to C. Convex cones are foundational in numerous optimization problems, particularly in wireless communication systems. Two commonly used convex cones are briefly described below [3]:

1. Second-Order Cone: A second-order (or quadratic) cone in $\mathbb{R}^{n+1}$ is defined as
$$C = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\| \leq t\}. \tag{3.13}$$

2. Semidefinite Cone: In the space of $n \times n$ symmetric matrices, denoted by $\mathbb{S}^n$, the semidefinite cone is defined as

$$C = \{\boldsymbol{X} \in \mathbb{S}_+^n \mid \boldsymbol{X} \succeq 0\}, \tag{3.14}$$

where $\boldsymbol{X} \succeq 0$ implies that $\boldsymbol{X}$ is positive semidefinite; Otherwise, the eigenvalues of $\boldsymbol{X}$ are all positive or zeros.

3.2.3 Convex Functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be convex if its domain, denoted dom f, is a convex set and, for all $x, y \in \text{dom } f$ and any $\theta \in [0, 1]$, the following inequality is satisfied [3, 69],

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y). \tag{3.15}$$

In this expression, the right-hand side represents the value along the straight line segment connecting the points $(x, f(x))$ and $(y, f(y))$, whereas the left-hand side corresponds to the function value at the convex combination of x and y. Geometrically, this implies that the graph of a convex function always lies below or on the chord joining any two points on the curve. This concept is illustrated in Figure 3.2, where the curve represents a convex function $f(x)$, and the line segment denotes the linear interpolation between two function values.

Furthermore, there are two key conditions that can be used to verify the convexity of a function f defined over a convex domain:

1. First-order condition: If f is differentiable, then it is convex if and only if

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T (\boldsymbol{y} - \boldsymbol{x}). \tag{3.16}$$

This condition implies that the function lies above its first-order (linear) approximation at every point.

2. Second-order condition: If f is twice differentiable, then it is convex if and only if its Hessian matrix is positive semidefinite, i.e.,

$$\nabla^2 f(\boldsymbol{x}) \succeq 0. \tag{3.17}$$

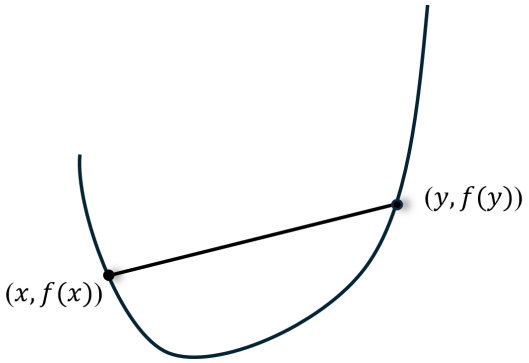


Fig. 3.2 Example of a convex function [3].

3.2.4 Convex Optimization Problems

A convex optimization problem is formulated as following [3]:

$$\begin{aligned} \min_{x} \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & g_i(\boldsymbol{x}) \leq 0, \quad i = 1,2,\ldots,m, \\ & h_i(\boldsymbol{x}) = 0, \quad i = 1,2,\ldots,p, \end{aligned} \tag{3.18}$$

where $x \in \mathbb{R}^n$ represents the optimization variables. $f, g_1, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ are convex functions, and $h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are affine functions. In this context, $f(\boldsymbol{x})$ is referred to as the objective function, while $g_i(\boldsymbol{x})$ and $h_i(\boldsymbol{x})$ denote the inequality and equality constraint functions, respectively.

The aim of this problem is to find an optimal x that minimizes $f(x)$ subject to all constraints within the feasible region, defined over a convex domain:

$$D = \operatorname{dom} f \bigcap_{i=1}^{m} \operatorname{dom} g_i \bigcap_{i=1}^{p} \operatorname{dom} h_i. \tag{3.19}$$

Moreover, for a differentiable objective function f, the optimality condition is given by the following criterion [3]. Let the feasible set be defined as:

$$X = \left\{ x \,\middle|\, g_i(x) \leq 0,\ \forall i = 1,\ldots,m;\ h_i(x) = 0,\ \forall i = 1,\ldots,p \right\}. \tag{3.20}$$

Then, a point $x^* \in X$ is said to be optimal if and only if it satisfies the following condition:

$$\nabla f(x^*)^T (y - x^*) \geq 0, \quad \forall y \in \mathcal{X}. \tag{3.21}$$

Based on these fundamental concepts of convex optimization, several common encountered standard problem types can be introduced, as follows,

1. Linear programming (LP): In LP problems, both the objective function and all constraint functions are affine, resulting in a convex optimization problem that can be efficiently solved using standard algorithms. It is formulated as following [3],

$$\begin{aligned} \min_{x} \quad & \boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{d} \\ \text{s.t.} \quad & \boldsymbol{G}\boldsymbol{x} \preceq \boldsymbol{h}, \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \end{aligned} \tag{3.22}$$

   where $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{G} \in \mathbb{R}^{m \times n}$, $\boldsymbol{A} \in \mathbb{R}^{p \times n}$, $\boldsymbol{h} \in \mathbb{R}^m$, and $\boldsymbol{b} \in \mathbb{R}^p$.

2. Quadratic programming (QP): A quadratic programming problem refers to a convex optimization problem in which the objective function is quadratic and the constraints are linear. It is formulated as [3],

$$\begin{aligned} \min_{x} \quad & \frac{1}{2}\boldsymbol{x}^T \boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}^T \boldsymbol{x} + r \\ \text{s.t.} \quad & Gx \preceq h, \\ & Ax = b, \end{aligned} \tag{3.23}$$

   where $\boldsymbol{P} \in \mathbb{S}_+^n$ (the set of positive semidefinite matrices), and $q \in \mathbb{R}^n$. It should be noted that the LP problem is a special case of QP when $P = 0$.

3. Quadratically constrained quadratic programming (QCQP): A QCQP problem extends QP by allowing both the objective and constraint functions to be quadratic. It is defined as [3],

$$\begin{aligned} \min_{x} \quad & \frac{1}{2}\boldsymbol{x}^T \boldsymbol{P}_0 x + \boldsymbol{q}_0^T \boldsymbol{x} + r_0 \\ \text{s.t.} \quad & \frac{1}{2}\boldsymbol{x}^T \boldsymbol{P}_i x + \boldsymbol{q}_i^T \boldsymbol{x} + r_i \leq 0, \quad i = 1, 2, \ldots, m, \\ & \boldsymbol{A}\boldsymbol{x} = b, \end{aligned} \tag{3.24}$$

   where $\boldsymbol{P}_i \in \mathbb{S}_+^n$ for all $i = 0, 1, \ldots, m$. Setting all $\boldsymbol{P}_i = 0$ reduces the QCQP to a QP problem.

4. Second-order cone programming (SOCP): An SOCP problem includes second-order (or quadratic) cone constraints and is formulated as [3]:

$$\begin{aligned} \min_{x} \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} \quad & \|\boldsymbol{A}_i \boldsymbol{x} + \boldsymbol{b}_i\| \leq \boldsymbol{c}_i^T \boldsymbol{x} + d_i, \quad i = 1, 2, \ldots, m, \\ & \boldsymbol{F}\boldsymbol{x} = g, \end{aligned} \tag{3.25}$$

   where $\boldsymbol{A}_i \in \mathbb{R}^{k_i \times n}$, $\boldsymbol{b}_i \in \mathbb{R}^{k_i}$, $\boldsymbol{c}_i \in \mathbb{R}^n$, $\boldsymbol{d}_i \in \mathbb{R}$, $\boldsymbol{F} \in \mathbb{R}^{p \times n}$, and $\boldsymbol{g} \in \mathbb{R}^p$. The first constraint in the formulation of problem (3.25) is referred to as a second-order cone constraint.

5. Semi-definite programming (SDP):

   In semi-definite programming, the problem involves matrix inequality constraints defined over the cone of symmetric positive semi-definite matrices. Given symmetric matrices $\boldsymbol{G}, \boldsymbol{F}_1, \ldots, \boldsymbol{F}_n \in \mathbb{S}^n$, an SDP problem is typically expressed as [3],

$$\begin{aligned} \min_{x} \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{F}_1 \boldsymbol{x}_1 + \boldsymbol{F}_2 x_2 + \cdots + \boldsymbol{F}_n \boldsymbol{x}_n + \boldsymbol{G} \leq 0, \\ & \boldsymbol{A}\boldsymbol{x} = b, \end{aligned} \tag{3.26}$$

   where $x = [x_1, x_2, \ldots, x_n]^T$ is the vector of decision variables.

   Another common formulation of SDP uses matrix variables and incorporates linear equality constraints and a positive semi-definiteness condition. It is written as [3],

$$\begin{aligned} \min_{X} \quad & \text{tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} \quad & \text{tr}(\boldsymbol{A}_i \boldsymbol{X}) = b_i, \quad i = 1, 2, \ldots, m, \\ & \boldsymbol{X} \succeq 0, \end{aligned} \tag{3.27}$$

   where $\boldsymbol{X}, \boldsymbol{C}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \in \mathbb{S}^n$ and $\text{tr}(\cdot)$ denotes the trace operator.

In wireless communications, resource allocation problems are common to be formulated as convex optimization problems such as SOCP or SDP, in order to convert the non-convex problems to convex ones. Once the problem is reformulated in convex optimization problem then it is possible to solve it with toolboxes such as `CVX` [70] and `YALMIP` [71]. It should be noted that in convex optimization, any

local optimal solution is also a global optimum [3, 69]. Therefore, transforming non-convex problems into convex formulations is highly advantageous, as it could enable computational efficient, lower in complexity, and theoretically robust solutions.

3.3 Machine Learning (ML)

ML has proven useful for solving complex optimization problems in communications [72]. While convex-based algorithms offer the advantage of mathematical rigor, they are inherently inflexible, especially in scenarios with random and dynamic nature. In addition with high demands and increase the complexity of the wireless network, the resulting complexity is difficult to handle even with convex optimization. Therefore, ML is proven to be useful in wireless communication. Similarly, in Wi-Fi networks, emerging technologies introduced to meet increasing network demands require ML functionality is required to address the high complexities for the next generation of Wi-Fi. In the upcoming sections some fundamental ML algorithms are introduced.

The derived ML algorithms can be divided into three main categories, supervised learning, unsupervised learning, and reinforcement learning [73] as illustrated in Fig. 3.3. In addition, federated learning can be used with all the three main categories of ML for data privacy, security and decentralized training.

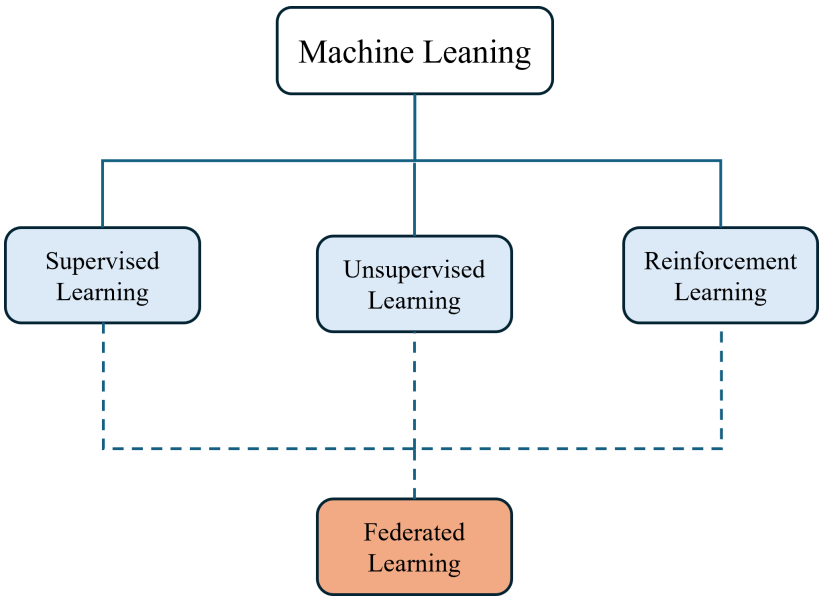


Fig. 3.3 ML techniques.

3.3.1 Supervised Learning

Supervised learning was among the earliest machine learning techniques to be formally developed [73]. The idea is that, based on a labeled dataset, which means each input data point is paired with a corresponding output or target value, the model could learn a mapping between the inputs and outputs. Therefore, once the mapping is set it is expected that the model could predict the output for a new set of inputs accurately, whether it has seen it before or not. This ML's technique is used in applications such as image classification, traffic classification, speech recognition, and spam detection, where historical data with known labels is available [74, 73]. Common supervised learning algorithms include linear regression, support vector machines, decision trees, and neural networks [73, 75].

3.3.2 Unsupervised Learning

Unsupervised learning is another type of ML where the model is trained on data without labeled outputs. Instead of learning a direct mapping from input to output, the algorithm explores the underlying structure or patterns within the data. This approach is commonly used for tasks such as clustering, dimensionality reduction, and anomaly detection. Without relying on labeled data, unsupervised learning is more data-efficient, while it offers less accuracy as lacking labeled data for guidance [73].

3.3.3 Reinforcement Learning (RL)

Unlike supervised and unsupervised learning, RL does not rely on a pre-collected training dataset. RL allows an agent to interact with an environment, receive feedback in the form of a reward, and select actions accordingly. In other words, RL is a framework in which an agent must learn from interactions with its environment what to do based on its interaction with the environment and how to map situations to actions, in order to maximize a numerical reward. The agent to take the action should discover the optimal strategy through exploration. The RL is explicitly different from other ML models, where the agent can achieve its goal in an uncertain and dynamic environment [76–80].

Elements of RL

Without agent and environment, which are the basic elements of RL, there are four main elements of RL, including a policy, a reward signal, a value function and a model of environment.

- **Policy:** A policy defines how a learning agent behaves at any given time. Generally, a policy maps observed states of the environment to actions to be taken once it at those states. Policy could be a simple function or a lookup table or a complex search process. The policy is the core of a RL agent in the sense that it is able to determine the agent's behavior.
- **Reward:** A reward signal defines how good was the action the agent took. The agent's goal is to maximize its reward over the long run. The reward signal helps for altering the policy, it means once an action has low reward, then the policy might be changed for maximizing the reward.
- **Value function:** Although the reward provides feedback, this feedback is received immediately after the agent takes an action. In order to observe the long-term reward, which is the main goal of RL, the value function is defined. The reward is accumulated since the start of the initial state. In other words, the main goal of the agent is to search for actions to bring about states of the highest value.
- **Model:** A model simulates the behavior of the environment and estimates its behavior. For instance, with a set of states and actions, one is able to predict the next state and next reward.

Markov Decision Processes (MDPs) are introduced in RL as a fundamental mathematical framework [76] to define the interaction between an agent and its environment over time. An MDP includes a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P})$, these key elements are as following,

- $\mathcal{S}$ is state space that has all the available states in the environment. Each state illustrates the corresponding information for the agent to take an action
- $\mathcal{A}$ is the action space set which is all the possible actions an agent can take.
- $\mathcal{R}$ is the reward function that defines the immediate reward value after taking a action.

- $\mathcal{P}$ is the transition probability, that defines the probability of transforming from one state to another one with a given specific action.

RL can be divided into two main categories, on-policy and off-policy RL. In policy-based RL, an agent learns a policy directly that is useful especially for the cases where continuous action spaces are available or the environment is complex. This helps for more effective exploration compared to the value-based RL. In value-based RL, the agent estimates the long-term reward of the taken action through the states. This function is also referred to as the Q-value function. Instead of directly mapping states to actions, the agent uses the value function to estimate the expected reward of each action and selects the action with the highest estimated value.

Q-learning is one of the first RL algorithms that has been proposed. It is a value-based RL algorithm, which means that the agent learns the optimal policy based on calculating the Q-values where the main goal is to maximize the long-term cumulative reward of the agent. Q-learning is based on generating a Q-table where the rows of the table are all possible states and the columns are possible actions. Each element of the table is a Q-value of the specific action and state. The Q-values are calculated as follows, representing the expected long-term reward for each action in a given state [76].

$$q^{\pi}(s(t),a(t)) = \mathbb{E}_{\pi}\left[\sum_{k=0}^{\infty}\gamma^{k} r(t+k+1) \mid s = s(t), a = a(t)\right], \tag{3.28}$$

where $\gamma \in [0,1]$ is a discount parameter indicating the priority of future reward, and $r(t+k+1)$ is the reward at time $t+k+1$. Also π is the policy. The agent takes actions according to the policy $\pi(a|s)$, which maps from states to probabilities of selecting each action. The table is illustrated in Table 3.2.

	a_1	a_2
s_1	$Q(s_1,a_1)$	$Q(s_1,a_2)$
s_2	$Q(s_2,a_1)$	$Q(s_2,a_2)$
$\vdots$	$\vdots$	$\vdots$
s_N	$Q(s_N,a_1)$	$Q(s_N,a_2)$

Table 3.2 An example of Q-table with two actions and N states.

In each state the agent with observation of its current state, and with its policy π takes an action. Based on the taken action the agent receives a reward and the agent's next state will be updated. It is expected, a well-trained agent takes actions

with the highest return accumulated reward from each state. Using Q-values, the optimal policy can be expressed as selecting the action with the highest Q-value for each state, which is given by:

$$\pi^*(s(t)) = \arg\max_{a(t)\in A} Q(s(t), a(t)). \tag{3.29}$$

It means that at a given state, agent learning policy is to take the actions that lead to higher Q-values. The agent's policy for taking the action is based on the ϵ-greedy policy. Which means the agent takes a random action with probability ϵ to explore and takes an action according to the Q-table with probability $1-\epsilon$, where $\epsilon \in [0,1]$. As the agent takes actions, it observes the resulting rewards and transitions to new states. The Q-values are updated using the Bellman equation, expressed as [76]:

$$Q(s(t), a(t)) \leftarrow r(t) + \gamma \max_{a(t)\in A} Q(s(t+1), a(t)). \tag{3.30}$$

This Q-value updating will be continued until it reaches the optimal policy. Q-table despite its simplicity and efficiency, has a drawback related to the state space. It means that as the number of states increases the table will be drastically increased, so once the agent wants to take an action it needs to observe the whole table; consequently, it will not applicable with continuous state space and infinite space. This also requires extensive memory. Therefore, it is suggested instead of calculating direct Q-values, use neural network to approximate the Q-values. In the next section, an introduction to deep learning and deep reinforcement learning is presented.

Deep Learning

Deep learning is an advanced extension of traditional neural networks in ML, characterized by the presence of multiple hidden layers and the capacity to handle a larger number of input and output nodes. This layered architecture enables deep learning models to perform more complex and hierarchical feature extraction, making them particularly effective in processing high-dimensional data such as images or time series.

In the context of RL, this capability addresses the limitations of traditional methods like Q-learning, which rely on a tabular representation of state-action pairs (Q-tables). Q-tables become impractical in environments with large or continuous

state spaces due to their inability to generalize and the exponential growth in required memory.

The integration of deep learning into RL, which can be used as a prediction for Q-values, can be used. In other words, using deep learning in RL, such as in Deep Q-Networks (DQN), allows for function approximation of the Q-values, enabling agents to operate effectively in high-dimensional or continuous environments. This advancement has further motivated the development of more sophisticated RL algorithms, such as Proximal Policy Optimization (PPO) and Deep Deterministic Policy Gradient (DDPG), which are specifically designed to handle continuous action spaces and complex dynamics using deep neural networks [76].

Within wireless resource allocation problems, DRL becomes particularly relevant when the action space is discrete but large, and when the system state changes continuously over time. In the scheduling problems addressed later in this thesis, the scheduler must jointly observe queue states, channel gains, and RU availability before selecting resource assignments under strict transmission deadlines. Under such conditions, classical tabular RL becomes impractical, whereas DRL enables state generalization through neural networks and supports scalable decision-making in high-dimensional scheduling environments.

Deep Reinforcement Learning (DRL)

DRL is the combination of RL with neural networks for more complex environments and scenarios. DRL algorithms can be divided to very advanced new algorithms in RL, such as deep deterministic policy gradient (DDPG), but in this section, DQN is explained. DQN is the extension of Q-table to solve the state space continuity and infinity problem.

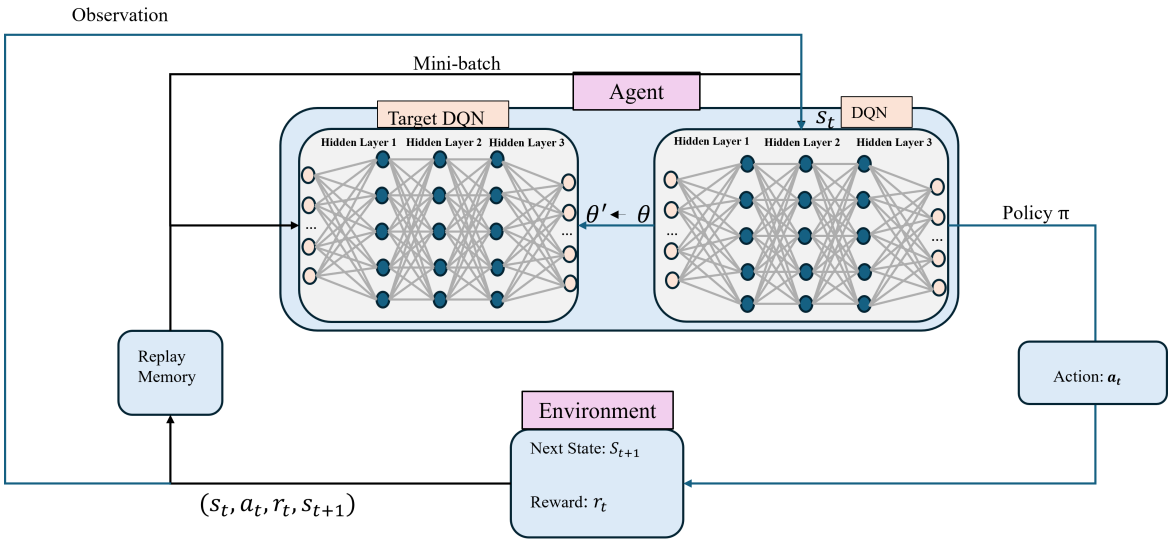


Fig. 3.4 DQN Algorithm

DQN has three main components, a main neural network, a target neural network, and an experience replay buffer. The main structure of the DQN is illustrated in Fig. 3.4. As depicted in Fig. 3.4, the agent initially has random neural network weights and with ϵ-greedy policy to balance exploitation and exploration takes action, where the ϵ-greedy policy is as follows,

$$a_t = \begin{cases} \text{random action} & \text{with probability } \epsilon \\ \arg\max_a Q(s_t, a; \theta) & \text{with probability } 1-\epsilon \end{cases}$$

During the interaction with the environment, the agent stores its experiences in a replay buffer $\mathcal{D}$. Each entry in the buffer is a tuple (s_t, a_t, r_t, s_{t+1}) where s_t is the current state, a_t the action taken, r_t the reward received, and s_{t+1} the next state. At each training step, a mini-batch of transitions is sampled randomly from $\mathcal{D}$ to break the temporal correlation of sequential data and improve data efficiency. In addition, to make the Q-learning stable, the use of two network one as the future Q-value estimation network, which is called target network and the other one to gather the tuple data and calculate the current Q-value is used [81]. For each transition in the batch, the target Q-value is computed using the target network, which remains fixed for a certain number of steps to provide stable targets. So the next reward is calculated as follows,

$$y_t = r_t + \gamma \max_{a' \in \mathcal{A}} Q(s_{t+1}, a'; \theta^{'}). \tag{3.31}$$

where θ' represents the parameters of the target network. Given this target, the main network estimated the current $Q(s_t, a; \theta)$ The learning objective is the loss function for the main network and is defined as the mean squared error between the predicted Q-value and the target value,

$$L(\theta) = \mathbb{E}_{(s_t, a_t, r_t, s_{t+1}) \sim \mathcal{D}} \left[\left(y_t - Q(s_t, a_t; \theta) \right)^2 \right] \tag{3.32}$$

The parameters θ of the main network are updated by minimizing this loss using stochastic gradient descent. The parameters of the target network $\theta^{'}$ are updated less frequently to stabilize learning. Specifically, after every fixed number of steps τ, the target network is synchronized with the main network,

$$\theta^{'} \leftarrow \theta \tag{3.33}$$

This delayed update prevents harmful feedback loops and allows the main network to learn toward a more stable target [81, 76].

Multi-agent Reinforcement Learning (MARL)

Single-agent RL algorithms need one agent that has control over the environment. However, there are scenarios with high complexity tasks that make agents' behaviors difficult to predict or even impossible to design. MARL is well-suited for such scenarios, as it allows multiple agents to make decisions while interacting with the same environment, rather than relying on a single RL agent.

MARL according to the type of targeted tasks in the system can be classified into fully cooperative, fully competitive, or mixed stochastic games (SGs) [82]. In fully cooperative SGs all agents share the same reward function, which means all agents have the same goal to maximize the common discounted return. In other words, each agent is learning its own policy while all the agents get the same reward, and each agent has its own observation from the environment. While on the other hand with the fully competitive SG, all agents get different reward and each agent is trying to maximize its own reward. The last classification of MARL is mixed SG. In mixed SG different cooperative groups while each group is competing with the other one are available and within each group, the agents in the group will have a shared reward. So the agents in one group trying to increase their shared reward while out of group each group is competing with the other one.

Each MARL classification has its own advantages and disadvantages. Fully cooperative SGs benefit from agents sharing the same reward function with the main goal of joint maximization of discounted reward. However, their primary disadvantage is in the coordination problem between independent agents, which can lead to suboptimal action selection with even considering that the coordination is not consistently broken. In contrast, fully competitive SGs, defined by agents having directly opposite goals, enable each agent to maximize its own benefit even though if it needs to go against an assumed optimal opponent, or even exploit a suboptimal one, rendering the concept of fairness irrelevant to their adversarial objectives. Lastly, mixed SGs offer the flexibility to model interested agents with unconstrained reward functions, where goals may align or conflict. However, this results significant challenges such as non-stationary due to simultaneously learning agents and the equilibrium selection problem when multiple solutions exist.

In mixed SGs classification where agents have heterogeneous reward structures, equilibrium selection problem, and private observations, FL offers a decentralized framework for learning shared or personalized models while preserving privacy. FL can facilitate more inclusive and robust learning in partially competitive settings.

3.3.4 Federated Learning

FL [83] has recently attracted significant attention due to its potential in emerging applications such as IoT, where reducing communication overhead is critical. In contrast to centralized training approaches where data is aggregated at a central server, FL enables multiple clients to collaboratively train a global model by sharing only model parameters or gradients, without sharing local data. This approach provides a privacy preserving, making it especial for sensitive data.

Another advantage of FL is its robustness in the presence of non-identically and independently distributed (non-IID) and unbalanced datasets, which is common in practical scenarios. FL is designed to operate under data heterogeneity and variable client participation.According to [83], the FL optimization problem can be formalized as a finite-sum objective:

$$\min_{w \in \mathbb{R}^d} f(w) \quad \text{where} \quad f(w) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^{n} f_k(w), \tag{3.34}$$

where $f_k(w)$ represents the utility function of local objective of the k-th client, commonly expressed as $f_k(w) = l_k(x_k, y_k; w)$, the loss of a neural network model with parameters w evaluated on local dataset (x_k, y_k).

To solve the optimization problem in (3.34) efficiently, the authors in [83] proposed several key algorithms, Federated Stochastic Gradient Descent (FedSGD) and Federated Averaging (FedAvg). In FedAvg, each client performs multiple steps of local SGD on its own data before sending its updated neural network model to a central server, which then with weighted averaging it updates the global model. This approach significantly reduces the number of communication rounds while maintaining competitive convergence behavior compared to centralized training.

However, FL also presents new challenges such as model divergence due to heterogeneous data, client drop-out, and fairness concerns. In particular, minimizing the average global loss may lead to suboptimal performance for underrepresented or low available data of clients [84], motivating ongoing research in fairness-aware and personalized federated learning strategies. However there are concerns in using the FL for fairness. This issue is discussed in detail in section 4.2.2.

3.4 Summary

This chapter has presented the mathematical foundations required for the development of the proposed algorithms in this thesis. While several fairness metrics, optimization tools, and machine learning techniques were introduced, only a subset of these methods are directly utilised in the subsequent chapters, based on their suitability for the considered problem settings.

Among the fairness metrics, max-min fairness and proportional fairness are adopted as the primary performance objectives in this work. Max-min fairness is employed in Chapter 6 to ensure equitable throughput distribution among STAs under heterogeneous channel conditions, particularly when protecting users with poor channel quality is critical. Proportional fairness is also utilised in Chapter 6 due to its ability to balance overall network throughput and fairness, and it enables the derivation of a closed-form RU allocation solution under convex problem formulations. In addition, Jain's fairness index is used throughout the thesis as an evaluation metric to quantify fairness performance in simulation results.

From an optimization perspective, convex optimization techniques are specifically applied in Chapter 6 to formulate and solve the proportional fairness problem. The convexity of the problem enables the derivation of efficient and computationally

tractable solutions, which are suitable for practical implementation in coordinated Wi-Fi systems.

With respect to machine learning, RL, and in particular DRL, is adopted as a key tool in this thesis. RL is utilised in Chapter 4 for improving channel access behaviour, and in Chapters 6 and 7 for solving complex RU allocation problems that are either NP-hard or dynamically evolving. The use of RL is motivated by its ability to learn adaptive policies in stochastic and time-varying environments, where conventional optimization approaches become intractable or require complete system knowledge.

Other techniques discussed in this chapter, such as supervised and unsupervised learning, are included for completeness but are not directly applied in this work, as they are less suitable for sequential decision-making problems inherent in wireless resource allocation.

Overall, this chapter establishes the theoretical tools that directly support the design of fairness-aware, optimization-driven, and learning-based resource allocation strategies developed in the remainder of this thesis.

Chapter 4

Simulation Study of Fairness-Based MARL in Variable-Rate Wi-Fi Networks

This chapter examines the application of the Fair-MARL algorithm, as proposed in [85], within Wi-Fi networks. The Fair-MARL algorithm is as a replacement for the conventional backoff mechanism of the DCF protocol, with the objective of improving network latency and spectral efficiency. In addition the Fair-MARL channel access has been tested and investigated for the scenarios with heterogeneous data rate. Finally, it is attempted to demonstrate the main concept learned through the initial system implementation and explore different relevant open research problems as the future work.

4.1 Introduction

As mentioned earlier in section 2.6.1 the channel contention in Wi-Fi network is critical especially due to the randomness of the network. With consideration of the CSMA/CA protocol, higher number of collisions in overcrowded environment are unavoidable. Throughout the literature several algorithms and protocols have been proposed to address this issue, specifically with the new ML algorithms. Among ML algorithms, DRL due to its capability for online decision making and the ability to learn in the dynamic wireless communication environment has attracted attention. Therefore for the channel access problem in Wi-Fi network, there are several papers for contention window (CW) optimization. One approach could be the backoff time

optimization of CSMA/CA protocol. In [86] authors recommended a DRL-based algorithm to select the best CW with the main objective of throughput optimization. While authors in [36] proposed a multi-agent DRL in a multi-AP coordination Wi-Fi network, where the agent aim is to select the best backoff time and clear channel assessment (CCA) threshold for reducing latency.

While backoff optimization in CSMA/CA with the RL algorithm is one aspect of conventional protocol optimization, another direction is to have a new protocol for channel access based on DRL. The authors of [87] were among the first to propose DRL as the new protocol for channel access. They proposed maximizing aggregate throughput and ensuring proportional fairness among all STAs in the network. Furthermore, the approach enables coexistence with heterogeneous protocols such as TDMA and ALOHA, without necessitating prior knowledge of their operational principles. In addition, authors with α-fairness objective proposed a semi-distributed MARL where the training process is performed centrally at the gateway node, while execution is carried out independently by the DRL nodes. However, the authors have not considered for channel contention among the DRL agents. Instead, they assumed that the central gateway assigns the channel using a round-robin algorithm whenever a DRL agent is expected to access the channel.

Following the approach in [87], the authors in [85] suggested the channel contention with a fully distributed multi-agent RL in a fixed physical data rate of multi-user single AP Wi-Fi network. Authors in [85] proposed a new algorithm for channel contention while within the proposed DRL algorithm and after training only one DRL STA could win the channel. Therefore, they proposed a federated learning (FL) for fairness between the DRL agents, instead of a centralized round-robin DRL nodes selection. However, it should be noted that the use of FL algorithm for fairness is debatable [38], especially since it has been attempted to minimize the network's latency while maintaining the throughput. In addition, the authors have not validated the algorithm in different scenarios with practical realistic assumptions. Therefore, in this chapter, the main goal is to validate the algorithm within realistic and practical scenarios such as the variable data rate of STAs, and identify new open problems to investigate in our future work.

4.2 System Model and Fair-MADRL Algorithm

4.2.1 System Model

The system model comprises a set of STAs associated with a single AP, where the STAs contend for channel access in order to perform uplink data transmission, as shown in Fig. 4.1. The set of STAs is represented as $\mathcal{N} = \{1, 2, \ldots, N\}$. Each STA is assumed to operate as an RL agent, capable of autonomously determining its transmission time.

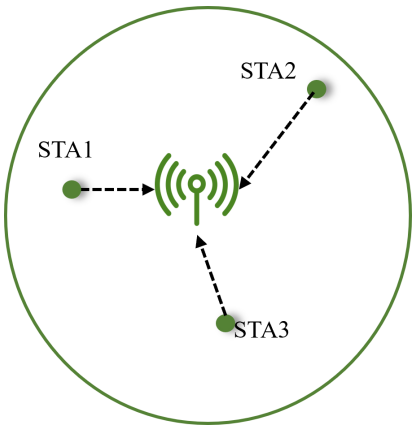


Fig. 4.1 Illustration of the proposed system model, where multiple STAs associated with a single AP contend for uplink channel access under the CSMA/CA mechanism, with each STA modeled as an RL agent.

As illustrated in Fig. 4.2 and discussed earlier in Section 2.6.1, the basic CSMA/CA (DCF) mechanism requires that, following a DIFS interval and while the channel remains idle, each STA randomly selects a contention window (CW). The contention window (CW) is selected with random uniform distribution within the range $[0, CW_{\text{min}}]$ and performs carrier sensing for the chosen backoff duration. If the channel is idle for the total backoff time, the STA will transmit its data. However, if any two STAs have the same backoff time, it means they transmit data at the same time so neither transmission is successful. If any two STAs have unsuccessful transmission, the contention window is doubled, up to $CW_{max} = 2^i CW_{min}$ and i is the number of unsuccessful transmission. The whole network performance will be degraded as the number of STAs increases so the backoff time will be increased drastically. So no one can transmit for a longer time to avoid collision and it leads to much higher latency.

Therefore with the DRL algorithm, it is expected to solve the high backoff time for overcrowded scenarios, with the goal of each STA as an agent accessing the channel with its local information. According to the proposed system model in [85], in each TXOP, all the stations contend to gain the channel access, which can be modeled as a multi-agent RL problem. Where each station is an agent that interacts with an unknown environment and based on the achieved experience, each agent derives a policy to attain the TXOP. In addition the channel access can be formulated as an MDP of DRL.

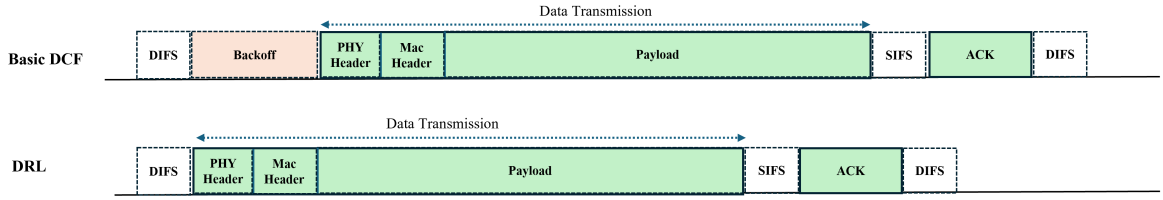


Fig. 4.2 Comparison of data transmission with basic DCF and DRL.

It is proposed that each STA's buffer state is always full, which means the STAs are in saturation mode. STAs as the DRL agents are trying to access the channel and contending for primary channel. Each STA agent receives the local observations and, using its trained DQN, selects an action. At t-th time step, the n-th station, given the current environment states s_t^n, takes an action a_t^n. The implementation part of the MARL is described in the following:

State Space

The station gets the reward, and the environment moves to the next state $\mathbf{s}_{t+1}^n$. Here in this model, the state is defined as the current channel state, the channel state for $M-1$ past time slots, and M previous actions. The channel state is denoted as $c = [idle, busy]$. Hence, the concatenated state is defined as

$$\mathbf{s}_{t+1}^n = \left[a_{t-M+1}^n, a_{t-M+2}^n, \ldots, a_t^n, c_{t+1}^i\right].$$

Action Space

In this model, since each station contends for the TXOP, the action space is identified as $a_t^n = [transmit, wait]$.

Reward

Reward design for our objective model can be established based on the station's action. So the reward is defined in the algorithm presented in Algorithm 1 [88],

Algorithm 1: Reward's Algorithm

Data: $\mathbf{s}_t^n$, feedback for all actions in $\mathbf{s}_t^n$
Result: Reward: $r\left(\mathbf{s}_t^n, a_t^n, \mathbf{s}_{t+1}^n\right)$

$r\left(\mathbf{s}_t^n, a_t^n, \mathbf{s}_{t+1}^n\right) \leftarrow 0$
if $a_t^n = \{1\}$ **then**
  $l \leftarrow L$
  while $l \geq 0$ **do**
    if $a_{t-l}^n = 1$ ***and*** *ACK received* **then**
      $r\left(\mathbf{s}_t^n, a_t^n, \mathbf{s}_{t+1}^n\right) \leftarrow \eta \cdot r\left(\mathbf{s}_t^n, a_t^n, \mathbf{s}_{t+1}^n\right) + 1$
    end
    else
      $r\left(\mathbf{s}_t^n, a_t^n, \mathbf{s}_{t+1}^n\right) \leftarrow \eta \cdot r\left(\mathbf{s}_t^n, a_t^n, \mathbf{s}_{t+1}^n\right) - 1$
    end
    $l \leftarrow l - 1$
  end
else
  if $c_{t+1}^n = \{1\}$ **then**
    $r\left(\mathbf{s}_t^n, a_t^n, \mathbf{s}_{t+1}^n\right) \leftarrow 1$
  end
  else
    $r\left(\mathbf{s}_t^n, a_t^n, \mathbf{s}_{t+1}^n\right) \leftarrow -1$
  end
end

Where l represents the previous $M-1$ time slots used by the agent to observe the channel state. η denotes a reward decay factor that controls the influence of past transmission outcomes when computing the current reward. L denotes the maximum number of previous time slots considered in the reward computation. With the proposed DRL formulation, it is typically observed that a single STA behaves more aggressively than the others and consequently gains most of the channel access opportunities. To address this issue, an algorithm is introduced to enable fair channel access among all DRL agents. In [85], the authors proposed the FedAvg-based FL algorithm for this purpose. However, the notion of fairness in

FL remains an open research question and continues to be actively debated in the literature [38]. Therefore, in the next section, the Fair-MARL algorithm proposed in [89] is implemented and evaluated under different network scenarios.

4.2.2 Fair-MARL Algorithm

Authors in [85] claimed the FedAvg has been proposed to address the fairness issue. However, the fairness, especially with the vanilla FL is debatable [38]. Therefore, the proposed Fair-MARL algorithm in [89] is considered throughout the rest of this chapter. The Fair-MARL algorithm is based on the Population-Based Training (PBT) framework proposed in [90], which jointly performs exploration and exploitation by training a population of neural networks and periodically adapting their parameters and hyperparameters.

Algorithm 2: Fair-MARL Algorithm

for $t = 1, 2, \ldots, T_{total}$ **do**
  if $t \mod \tau_{\text{update}} = 0$ **then**
    Find $n^\star \in \arg\max_{n \in N} T_n$
    Choose $j \sim \mathcal{U}\big(N \setminus \{n^\star\}\big)$
    set $\omega_j \leftarrow \omega_{n^*}$
    Choose $q \sim \mathcal{U}\big(N \setminus \{j, n^\star\}\big)$
    set $\omega_{n^\star} \leftarrow \omega_q$
  end
end

In this algorithm once at least one agent wins the channel and at every update step, where $t \mod \tau_{\text{update}} = 0$, the STA with the highest number of successful packet transmissions is chosen. The STA is denoted as n^* and its trained weights are copied as ω_{n^*}. Then, ω_{n^*} are copied to the j-th random STA. The q-th STA is chosen randomly and its weight are copied to n^*-th STA. This will provide each STA at some certain time slots to access the channel and prevent any aggressive channel access from one STA [89, 91]. In the next section, the algorithm is compared with the basic DCF, evaluated its performance and identified open problems with some discussions.

4.3 Performance Evaluation and Simulation Results

It is assumed the network has only one AP where several number of STAs between $2-8$ STAs is available through different simulation setups. It is supposed all the STAs have the same packet length and are contending for channel access of 20 MHz bandwidth. The STAs' modulation and coding schemes (MCSs) vary between $0-13$ with physical data rate between $15-150$ Mbps. Throughout the simulation, the Fair-MARL is compared with basic DCF as the benchmark. The simulation parameters are provided in Table 4.1.

Table 4.1 Simulation parameters

Parameter	**Value**
Slot time	10 μs
SIFS	20 μs
DIFS	40 μs
PHY Header	20 μs
Headers	60 bytes
ACK	40 μs
Packet Length	1500 bytes
T	400 μs

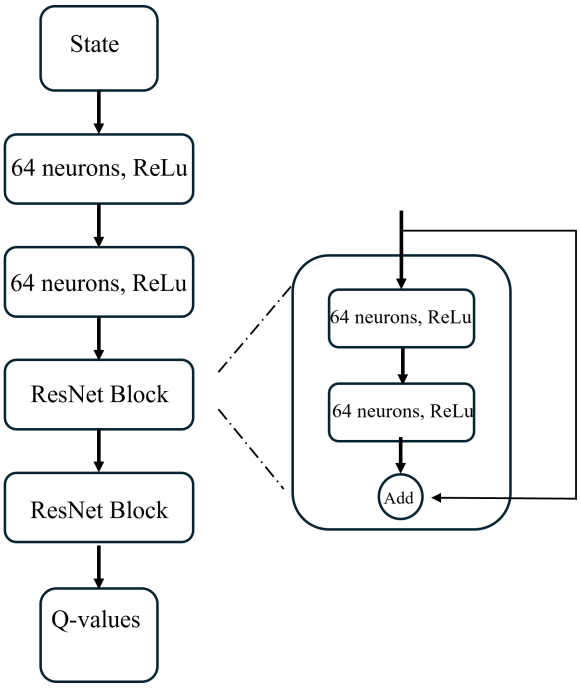


Fig. 4.3 Neural network model.

In terms of the configurations of the considered neural network, the deep learning component of the model employs a six-layer ResNet architecture, with each hidden layer consisting of 64 neurons [92], as illustrated in Fig. 4.3. The first two hidden layers are fully connected layers, and the two ResNet blocks is added. Each ResNet block is included with two fully connected hidden layers plus one shortcut from the input to the output of the ResNet block [92]. The deep learning neural network hyperparameters are presented in Table 4.2. It is assumed each station has a common knowledge of the channel status, and attempts to obtain the TXOP whenever the channel is sensed to be idle. However, each node does not have any information about the other STAs at the same time, therefore would lead to a collision.

Table 4.2 DQN hyper parameters.

Parameter	**Value**
Learning Rate	0.00001
State Size	40
Action Size	2
Exploration Rate ϵ	1
Min Exploration Rate ϵ	0.05
Exploration Rate Decay	0.995
Batch Size	32
Memory Size	1000
γ	0.9
Number of Epochs	80000

The throughput is defined in this chapter based on the aggregated successful transmitted bits over the total simulation time. However, another way to define throughput is Bianchi's throughput formula [93]. In this way, the throughput is derived using a discrete-time Markov chain that models the backoff procedure of each STA in the network. The probability that a station transmits in a randomly chosen slot, as well as the probability of a collision, are calculated from the steady-state distribution of the Markov chain. Using these probabilities, the throughput is expressed as the fraction of time the channel is used to successfully transmit payload bits, taking into account successful transmissions, collisions, and empty slots. This approach provides an analytical estimate of the network throughput under saturated traffic conditions. It is assumed that all STAs transmit packets of equal length and initially operate at a physical data rate of 60 Mbps. The transmitted data bits can be calculated based on the STAs' data rate, and their duration of transmissions in each TXOP. Therefore, the throughput can be defined as follows,

$$S = \frac{E[\text{payload information transmitted in a slot time}]}{E[\text{length of a slot time}]}. \tag{4.1}$$

Here, $E[\cdot]$ denotes the expected value, reflecting the average over all possible slot events, including successful transmissions, collisions, and idle slots, which allows the throughput to be expressed as the long-term fraction of time used to successfully transmit payload bits. The time slot is set equal to the time needed at any station to detect the transmission of a packet from any other station, and here it is assumed $10\mu s$ [93].

Fig. 4.4 illustrates the Fair-MARL and basic DCF averaged throughput per STAs in the network. As demonstrated by the basic DCF algorithm, with increasing the number of STAs, the averaged network throughput with this protocol is decreased. While Fair-MARL protocol is able to maintain the average network throughput. In addition, note that within the 60 Mbps physical data rate, the Fair-MARL protocol can achieve 75% of the physical data rate. While basic DCF can achieve near 55% of its data rate.

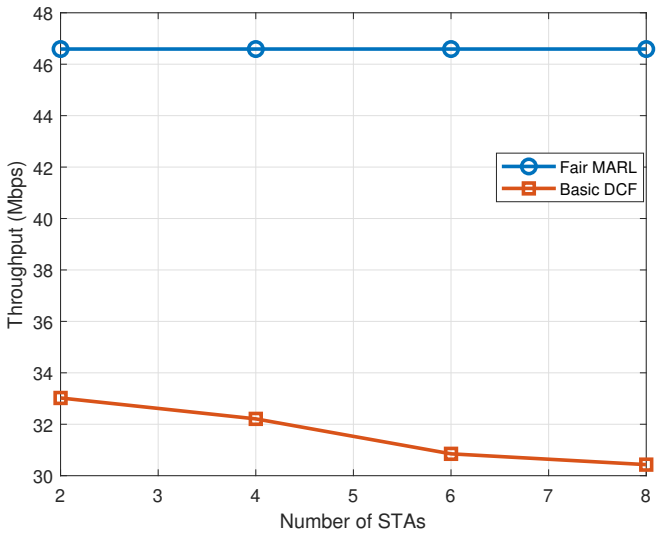
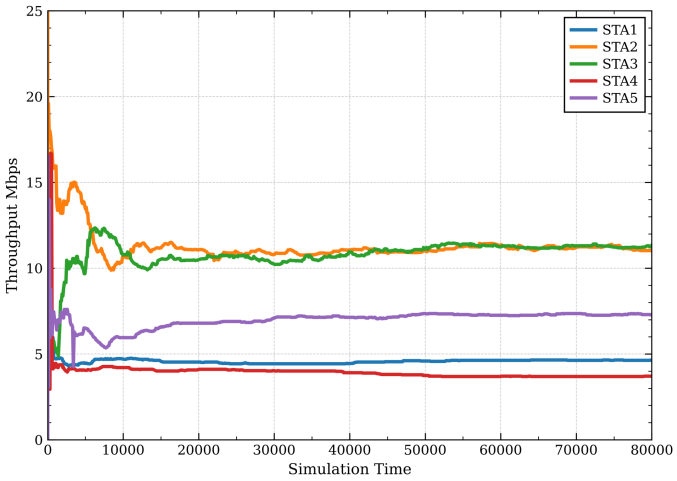


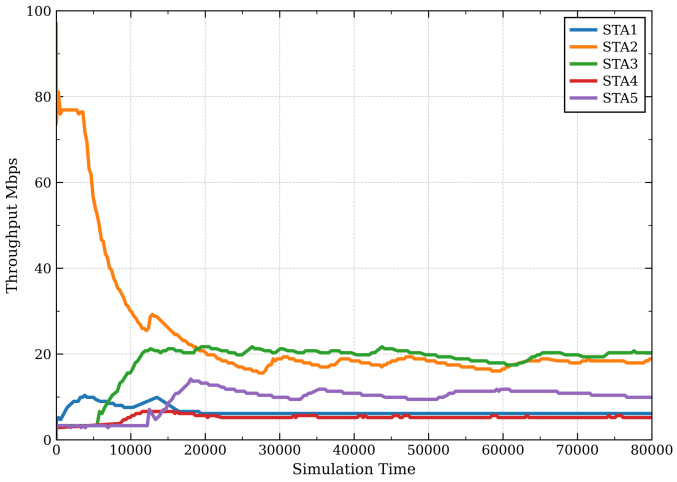
Fig. 4.4 Averaged network throughput in Mbps per number of STAs.

Fig. 4.5 shows the comparison between basic DCF and Fair-MARL algorithms for channel access. As it is illustrated, both Figs. 4.5a and 4.5b present the proportional fairness based on each STA data rate. In both figures the data rate of STAs is between $15-150$ Mbps, which is based on the variable channel condition of each STA with the same 20 MHz primary channel. The data rate of $n = [1,2,\ldots,5]$ are $[23,124,138,18,67]$ respectively. Therefore, it is presented that STA3 and STA2, with

data rate 124, 138 respectively, can access the channel more frequently compared to STA1 and STA4 with data rates 23, 18 Mbps. Note that it is assumed that all the STAs have airtime fairness. This is a valid assumption, as in practical implementation, it is assumed that each STA has equal airtime to avoid the slowest STAs in the system decreasing the system throughput and performance. In addition, authors in [94] have proved that airtime fairness provides proportional fairness in a Wi-Fi network.



(a) The throughput of the STAs over the course of the simulation under the basic DCF protocol.



(b) The throughput achieved by the STAs over the course of the simulation when employing the Fair-MARL protocol.

Fig. 4.5 STAs' throughput for different channel access protocols.

The MARL algorithm without fairness-aware allocation (Algorithm 2) results in highly unbalanced channel access, where a single STA occupies the channel for extended periods. In heterogeneous data rate scenarios, this behaviour tends

to favour the STA with the highest data rate, which can dominate channel access throughout the transmission duration. In contrast, in fixed data rate Wi-Fi networks, one STA may randomly gain persistent access to the channel, leading to similarly unfair outcomes. These observations highlight that, without incorporating fairness mechanisms, MARL-based channel access can result in significant performance imbalance among STAs.

Moreover, access categories discussed in section 2.6.3 for EDCA protocol are presented and simulated in Fig. 4.6. As presented in Table 4.3 , STAs 1-5 are assigned different access categories, which means they have different CW_{min} and CW_{max} with different TXOP duration, which was explained earlier in section 2.6.3. The differences in AC make the probability of accessing the channel different, whereas variations in data rates determine the duration of each channel access. As illustrated in Fig. 4.6 the STA4 is AC-VO, access the channel slightly more frequent compared to STA2 and STA3 with AC-VI. In addition, STA1 and STA5 are AC-BE that need less frequent access to the channel however, their TXOP duration is higher than AC-VO and AC-VI. This means it is expected them once they access the channel they could transmit for longer time. Which Fig. 4.6 is demonstrating this concept. Fig. 4.6 is based on the conventional channel access EDCA, however the Fair-MARL role in this is not discussed to the best of our knowledge. The Fair-MARL protocol has not considered this concept and this will be investigated in our future work. In other words, the main goal is to design a new algorithm or protocol that enables each agent to access the channel in accordance with its access category. While with the current Fair-MARL algorithm, each agent has the same frequency of accessing the channel.

STAID	**Access Category**	CW_{min}	CW_{max}	**Selected Data Rate**	**TXOP limit**
1	AC-BE	31	1023	60	5 ms
2	AC-VI	15	31	60	3 ms
3	AC-VI	15	31	60	3 ms
4	AC-VO	7	15	60	1.5 ms
5	AC-BE	15	1031	60	5 ms

Table 4.3 STA Information with Access Category and Data Rate

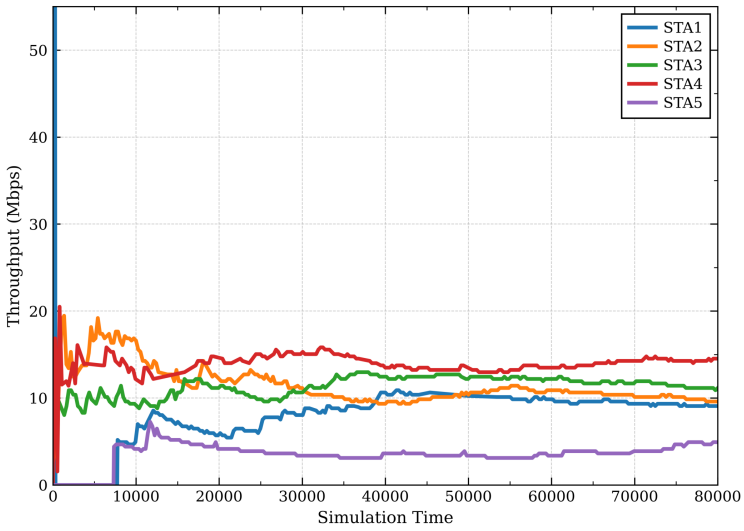


Fig. 4.6 The throughput of the STAs under the EDCA protocol is evaluated over the entire simulation duration, reflecting the average rate of successfully transmitted payload bits throughout the simulated network operation.

Although the proposed optimisation framework provides useful analytical insight into resource allocation under the considered system assumptions, several limitations should be noted. First, the optimisation relies on centralised decision-making and assumes accurate knowledge of long-term channel conditions for all users, which may become difficult to maintain in dense deployments with a large number of stations. Second, as the number of users increases, the computational burden associated with solving the optimisation problem grows, which may limit direct real-time implementation in practical systems with strict scheduling deadlines. In addition, the current formulation does not explicitly account for the signalling overhead required to collect channel information and distribute scheduling decisions. Future improvements may therefore consider reduced-complexity heuristics, distributed implementations, or approximation methods that preserve fairness while reducing computational cost and deployment complexity.

4.4 Summary

In this chapter, the fundamental concepts of channel access using the basic DCF and EDCA protocols were investigated and simulated. Furthermore, one of the protocols proposed in [85] was implemented and evaluated. The initial results demonstrated that the Fair-MARL algorithm improved throughput, as each agent determined its

transmission time based on local observations. This led to a saving of time and enabled the transmission of a larger number of packets compared to the basic DCF. In addition, the Fair-MARL protocol was tested in a more practical scenario where STAs operated at different data rates. The results showed that Fair-MARL achieves proportional fairness comparable to that of the basic DCF, while ensuring equal airtime allocation among STAs in the network. Finally, open challenges related to incorporating access categories into Fair-MARL were identified and discussed, which will be addressed in the next stage of this research. Following this in the next chapter, we propose a scheduling algorithm for a joint C-OFDMA and C-SR system model to improve total network throughput and satisfying STAs' QoS.

Chapter 5

Resource Unit Allocation in Coordinated OFDMA Multi-User Wi-Fi Systems

The scheduling of RUs in a multi-AP coordination network is considered vital for high spectrum efficiency and effective interference mitigation in future Wi-Fi Networks. In this chapter, we propose a RU allocation scheme to maximize the long-term average network throughput of STAs in a multi-AP coordinated Wi-Fi network within a joint coordinated orthogonal frequency-division multiple access (C-OFDMA) and coordinated special reuse (C-SR) framework.

The proposed scheme ensures that the average rates satisfy the STAs' QoS requirements and the interference in the overlapping basic service sets (OBSSs) is efficiently managed. The original RU allocation problem is formulated as binary integer programming, which is an NP hard. To address this, a heuristic graph coloring model is introduced for RU allocation in OBSSs, where the Sharing-AP allocates RUs with weighted max-min (WMM)-based graph coloring to overlapping STAs while simultaneously assigning RUs to its own STAs within the same frequency band using the WMM algorithm. Simulation results demonstrate that the proposed RU scheduling algorithm enhances overall mean network throughput by 30%. The simplicity and low computational complexity of the proposed algorithm confirm its effectiveness and practicality for implementation.

5.1 Introduction

MAPC network design is divided into two principal categories, 1) centralized MAPC (C-MAPC) incorporates a centralized primary controller that possesses comprehensive network information and requires high-performance processors with wired backhaul connections. 2) Distributed MAPC (D-MAPC), operates on a sharing/shared coordination principle without a central controller, allowing any of the APs to assume the sharing role. This architecture enhances network scalability and cost-effectiveness. Signal transmissions in MAPC can be facilitated over the air if the APs are within range of each other or through the wired backhaul. This chapter primarily explores the D-MAPC owing to its scalability, cost efficiency, ease of implementation, and the simplicity of integrating it with existing Wi-Fi amendments [2]. Regardless of the specific MAPC architecture, the literature identifies four advanced techniques pivotal to MAPC: C-SR, C-BF [40], C-OFDMA, and JTX [40, 18, 44]. These techniques play crucial roles in enhancing the performance of networks [45, 44].

C-OFDMA can provide simultaneous channel access at different APs with coordination based on frequency separation thus the interference will be managed efficiently compared to the conventional non-MAPC models [40]. For this feature, two types of resource allocation among different APs are available, one idea is that the agreed coordinated AP sets have the same primary channel while the RUs are divided between the agreed coordinated AP sets' requirement. However, failing to allocate the entire bandwidth to the coordinated set of APs leads to inefficiencies in resource allocation [2]. The other one uses different primary channels for the neighborhoods' APs which causes the inter-channel interference (ICI) [2]. This approach allocates entire frequency bands to all STAs and their respective APs, regardless of whether overlap or interference is present, potentially leading to suboptimal spectrum utilization. As a result, this approach may exhibit inefficiencies, particularly in real-time and industrial applications where optimal spectrum utilization is crucial for maintaining performance and reliability [2].

On the other hand, C-SR provides parallel transmission with interference management [18] in the same frequency and time with coordination between APs and transmission power management of APs/STAs. Using C-OFDMA with C-SR can highly increase the spectrum efficiency. By coordinated RU allocation, multiple APs can transmit with different RUs while with C-SR, STAs can transmit concurrently on the same RUs leading to more efficient use of the available spectrum and interference

management [2]. It enables concurrent transmissions in overlapping regions by employing efficient spectrum partitioning and interference mitigation strategies. Therefore, in this chapter it is proposed a RU allocation scheduling and interference management scheme within a joint C-OFDMA and C-SR framework.

Given the dynamic nature of wireless channels and the limited spectrum availability, RU allocation plays a crucial role, particularly in multi-AP scenarios. Without an efficient RU allocation strategy, interference can significantly degrade network performance [2]. In the literature, RU allocation is primarily based on real-time channel quality. For instance, the authors in [95, 96] propose a multi-cell OFDMA downlink channel assignment, considering the MAX k-CUT problem in graph theory. Authors [97] proposed RU allocation based on max-min fairness aiming to optimize the smallest ratio of achievable throughput to the minimum requested throughput.

The main contribution of this work is the development of a low-complex and practically implementable RU scheduling algorithm for a Wi-Fi network under realistic practical assumptions. The details are as follows. First, we propose an RU allocation in a joint C-OFDMA and C-SR D-MAPC system model. The main objective of the proposed scheduling algorithm is to maximize the total long-term average network throughput while satisfying STAs' required throughput The problem is formulated as a binary integer programming problem, which is an NP-hard problem. To solve the problem first a simple yet practical heuristic graph coloring model is proposed for RU allocation in the OBSS where no precise STAs' locations and no SINR information are needed. Subsequently, a WMM algorithm is introduced to allocate RUs for each non-overlapping STA, ensuring the proposed problem formulation is effectively addressed. The simulation results highlight the enhancement in total network throughput achieved by the proposed algorithm, showcasing its effectiveness despite being simple and practical.

5.2 System Model and Problem Formulation

The considered system model consists of a number of APs and their STAs operating under a D-MAPC based Wi-Fi network. It uses a joint C-OFDMA and C-SR model, where one AP in the MAPC group acts as the Sharing-AP, allocating resources for overlapping STAs to enable simultaneous frequency reuse in different locations. We assume an over-the-air connection between the Sharing-AP and Shared-APs, meaning they are within each other's coverage area, as shown in Fig. 5.1. The Sharing-AP

secures the TXOP based on the lowest backoff time. The uplink transmissions of STAs are considered.

The sets representing the STAs served by the m-th AP, the APs themselves, and the available RUs are denoted as follows: $n \in \mathcal{N} = \{1,2,\ldots,N\}$, $m \in \mathcal{M} = \{1,2,\ldots,M\}$ and $k \in \mathcal{K} = \{1,2,\ldots,K\}$, respectively. Additionally, the set of intersecting regions between OBSSs is represented as:

$$\mathcal{I} = \bigcup_{m,m' \in \mathcal{M}} \{C \mid C \in \mathcal{C}_m \cap \mathcal{C}_{m'}, m \neq m'\}.$$

where $\mathcal{C}_m$ and $\mathcal{C}_{m'}$ represent OBSS areas served by m-th and m'-th APs, respectively. C is the overlapping area for these two APs and $\mathcal{I}$ aggregates all common regions for any two OBSSs in the system model. Moreover the STAs located in $\mathcal{I}$ are $n' \in \mathcal{N}' = \{1,2,\ldots,N'\}$.

Each STA is limited to one RU per TXOP. All APs are assumed to use the same 20 MHz primary channel [98]. Therefore the co-channel interference needs to be mitigated. In this context, $h^m_{n,k}$ is the channel gain between the n-th STA and the m-th AP on RU k. This gain is assumed to be independent and identically distributed (IID) across each time step, emphasizing the stochastic nature of the channel conditions in this coordinated network model.

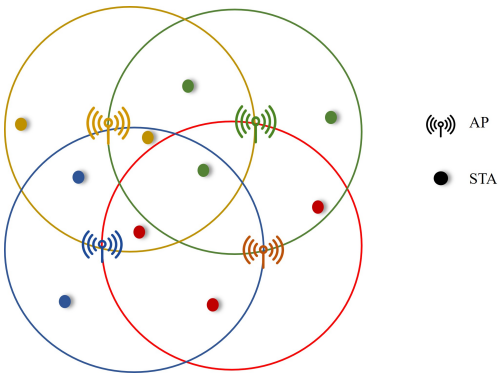


Fig. 5.1 System model of the considered D-MAPC-based Wi-Fi network. Multiple APs with their associated STAs form overlapping BSSs (OBSSs). A Sharing-AP coordinates resource allocation among neighboring APs to enable joint C-OFDMA and coordinated spatial reuse for STAs located in overlapping coverage regions.

Additionally, all the STAs transmit with a constant power and $\rho^m_{n,k}(t)$ is defined as a binary variable. When $\rho^m_{n,k}(t) = 1$, it indicates that the n-th STA is using the k-th RU; if it is zero, no resources are allocated for transmission of the n-th STA in the k-th RU. The data rate is defined as follows:

$$r_{n,k}^{m}(t) = r\left(h_{n,k}^{m}(t)\right)\frac{T_{\text{TXOP}}}{T_{\text{OFDM}}}, \tag{5.1}$$

where $h_{n,k}^{m} = \text{PL}_{n}^{m}(d) \times g_{n,k}^{m}$, in which $g_{n,k}^{m}$ is small scale fading of n-th STA and the m-th AP on RU k and $\text{PL}_{n}^{m}(d)$ is denoted as formula in (5.1). T_{TXOP} is the duration of uplink OFDMA for each TXOP transmission, and it is fixed during the simulation. T_{OFDM} denotes the duration of OFDM symbol, and $r(.)$ models the rate selection scheme for the n-th STA of the m-th AP on the k-th RU. As mentioned earlier, in this chapter, the main objective is to maximize the long-term average throughput of all OBSSs. We assume that each STA gets only one RU, therefore the formula can be rewritten as follows:

$$r_{n}^{m}(t) := \sum_{k=1}^{K} \rho_{n,k}^{m}(t) r_{n,k}^{m}(t). \tag{5.2}$$

The main goal here is to maximize the average data rate of STAs, therefore the long-term average data rate can be defined as,

$$\bar{r}_{n}^{m} := \lim_{T_{\text{total}} \to \infty} \sup \frac{1}{T_{\text{total}}} \sum_{t=0}^{T_{\text{total}}-1} \mathbb{E}[r_{n}^{m}(t)]. \tag{5.3}$$

The average long-term data rate is a concave, continuous, and entrywise non-decreasing function with respect to the STAs' data rates. This formulation is considered under the constraint that the average data rate of each STA satisfies a predefined minimum data rate threshold. Accordingly, the RU allocation problem in the considered system can be formulated as follows:

$$\underset{\rho_{n,k}^{m}}{\text{maximize}} \quad \sum_{m=1}^{M} \sum_{n=1}^{N} \bar{r}_{n}^{m} \tag{5.4a}$$

$$\text{subject to} \quad r_{n}^{\text{Tr}} \leq \bar{r}_{n}^{m}, \quad \forall n \in \mathcal{N}, \tag{5.4b}$$

$$\sum_{n=1}^{N} \rho_{n,k}^{m}(t) \leq 1, \forall k \in \mathcal{K}, \tag{5.4c}$$

$$\sum_{k=1}^{K} \rho_{n,k}^{m}(t) \leq 1, \forall n \in \mathcal{N}, \tag{5.4d}$$

$$\sum_{n'=1}^{N'} \rho_{n',K}^{m}(t) \leq 1, \forall k \in \mathcal{K}. \tag{5.4e}$$

The objective function (5.4a) is the long-term average of network data rate. Constraint (5.4b) ensures that the STA's average data rate is greater than r_n^{Tr}, which is the average data rate QoS threshold. The constraints in (5.4c) and (5.4d) confirm that for each BSS, RUs can only be allocated to at most one STA, and each STA can have a maximum of one RU. Constraint (5.4e) ensures that any two overlapping STAs in the set $\mathcal{I}$ cannot get the same RU.

In this model, it is assumed that STAs closer to the AP are able to utilize higher Modulation Coding Schemes (MCSs) for data transmission and reception, while those positioned further away use lower MCSs. APs select the appropriate MCS for each STA based on the received SNR of the STA Further details can be found in Section 5.4. Additionally, it is assumed that the Sharing-AP gathers information on the channel amplitude response of overlapping STAs, and the BSSs in which these overlapping STAs are located. The pathloss of n-th STA in m-th AP based on the Residential cases of the Task Group ax (TGax) of the IEEE 802.11 working group [98] is as follows:

$$\begin{aligned} \mathrm{PL}_n^m(d) = 40.05 + 20\times\log_{10}(f_c/2.4) + 20\times\log_{10}(\min(d,5)) \\ + (d>5)\times 35\times\log(d/5), \end{aligned} \tag{5.5}$$

where, f_c denotes 5 GHz, and d is the distance between the transmitter and the receiver in meters.

5.3 Proposed Algorithm

As described earlier, the main goal in this chapter is to allocate RUs with joint C-OFDMA and C-SR, where the C-OFDMA is attempted to solve RU assignment. The RU allocation problem defined in (5.4a)- (5.4e) is NP-hard, not feasible and the solution cannot be obtained with polynomial time complexity. In order to solve the problem, the problem is divided into two separated huristic algorithms first is with weighted-max-min algorithm and the second one is the graph coloring. WMM is used to solve the main objective (5.4a) and (5.4d). While graph coloring is used to solve the C-SR interference management with considered condition in (5.4e).

5.3.1 WMM

As earlier mentioned, it is expected that Sharing-AP allocates the overlapping STAs RUs based on channel fading consideration and avoiding overlapping-STAs with the

same edge having the same assigned RU. While each AP allocates based on a WMM algorithm allocated its own STAs's RUs.

The problem in (5.4a)- (5.4b) is reformulated via Lyapunov Optimization [97].

$$\underset{\gamma_n^m[t],\rho_{n,k}^m}{\text{maximize}} \quad \overline{\min_{n\in\mathcal{N}}\{\gamma_n^m[t]\}} \tag{5.6a}$$

$$\text{subject to} \quad \bar{\gamma}_n^m \le \frac{\bar{r}_n^m}{r_n^{\min}},\ \forall n \in \mathcal{N} \tag{5.6b}$$

To address the problem formulation in (5.6a) and its connection to the initial one (5.4), let's assume first $U(.)$ is denoted as the total utility function of total optimization problem in (5.4a), and $\gamma[t] = \big(\gamma_1[t],\ldots,\gamma_N[t]\big)$ is the optimization variable here. Therefore, based on Jensen's inequality if we could solve the problem $\overline{U(\gamma^\pi)} \ge U^\star - O(\varepsilon)$, then we are able to solve the $U\big(\bar{\tilde{r}}^\pi\big) \ge U^\star - O(\varepsilon)$. Considering $\bar{r}^\pi \ge \bar{\gamma}^\pi$, the Jensen's inequality is formulated as follows:

$$\bar{r}^\pi \ge \bar{\gamma}^\pi \ \Rightarrow\ U(\bar{r}^\pi) \ge U(\bar{\gamma}^\pi) \ge \overline{U(\gamma^\pi)}, \tag{5.7}$$

Therefore, to solve the main optimization problem in (5.4a)-(5.4e), it could be divided in two parts, first optimization in (5.6a)-(5.6b) needs to be solved, while needs to consider constraints on allocate RUs in (5.4c), (5.4d). In addition during the RU allocation the interference between overlapping-STAs with constraints (5.4e) needs to be addressed. So the total problem formulation is as following,

$$\underset{\gamma_n^m[t],\rho_{n,k}^m}{\text{maximize}} \quad \overline{\min_{n\in\mathcal{N}}\{\gamma_n^m[t]\}} \tag{5.8a}$$

$$\text{subject to} \quad \bar{\gamma}_n^m \le \frac{\bar{r}_n^m}{r_n^{\min}},\ \forall n \in \mathcal{N}, \tag{5.8b}$$

$$\sum_{n=1}^{N} \rho_{n,k}^m(t) \le 1, \forall k \in \mathcal{K}, \tag{5.8c}$$

$$\sum_{k=1}^{K} \rho_{n,k}^m(t) \le 1, \forall n \in \mathcal{N}, \tag{5.8d}$$

$$\sum_{n'=1}^{N'} \rho_{n',K}^m(t) \le 1, \forall k \in \mathcal{K}. \tag{5.8e}$$

The proposed reformulated problem can be solved with the WMM algorithm [99]. Initially the solution is to associate the (5.8b) to a virtual queue based on drift-plus-penalty (DPP), so we have,

$$Q_n^m[t+1] = \left(Q_k^m[t] - \frac{r_n^m[t]}{r_n^{\min}} + \gamma_n^m[t]\right)^+.$$

The DPP method suggests choosing $(\gamma_n^m[t], \rho_{n,k}^m)$ at each time slot to approximately minimize

$$\Delta(Q[t]) - V \cdot \min_n \gamma_n^m[t],$$

Following the standard analysis [99], the upper bound of the form is as following,

$$\Delta(Q[t]) - V \min_n \gamma_n^m[t] \;\le\; \text{const} - \left(V \min_n \gamma_k^m[t] - \sum_n Q_k^m[t]\gamma_n^m[t]\right) - \sum_k Q_n^m[t]\frac{r_n^m[t]}{r_n^{\min}}.$$

Therefore the problem can be solved with the Algorithm 3.

Algorithm 3: Weighted Max-Min Scheduling [97]

1 **for** $t = 0, 1, 2, \ldots$ **do**

2 Observe $\{Q_n^m[t]\}_{n\in\mathcal{N}}$ and the channel state $h_{n,k}^m$;

3 Determine $\gamma_n^m[t] \in \Gamma$ by solving the following optimization problem:

$$\max_{\gamma_n^m[t]\in\Gamma} V \min_{n\in\mathcal{N}}\{\gamma_n^m[t]\} - \sum_{n=1}^{N} Q_n^m[t]\gamma_n^m[t]$$

4 Determine the scheduling variable $\rho_{n,k}^m$ by solving the following maximization problem:

$$\max_{\rho_{n,k}^m} \sum_{n=1}^{K} Q_n^m[t]\frac{r_n^m[t]}{r_n^{\min}}$$

5 Update the virtual queues.

The WMM problem can be addressed by solving a set of deterministic subproblems at every scheduling period t. The optimal auxiliary variables can be achieved as,

$$\max_{\gamma_N^m[t]\in\Gamma} V \min_{n\in\mathcal{N}}\{\gamma_n^m[t]\} - \sum_{n=1}^{N} Q_n^m[t]\gamma_n^m[t] \quad \le \max_{\gamma_n^m[t]\in\Gamma} \gamma_{\min}[t]\left(V - \sum_{n=1}^{N} Q_n^m[t]\right), \tag{5.9}$$

where $\gamma_{\min}[t] = \min_{n\in\mathcal{N}}\{\gamma_n^m[t]\}$. Based on (5.9), it is straightforward to show that

$$\gamma_n^{m\star}[t] = \begin{cases} R^{\max}, & \text{if } V > \sum_{n=1}^{N} Q_n^m[t], \\ 0, & \text{otherwise.} \end{cases}$$

Then after calculating the γ_n^m, ρ_n^m for each AP's STAs RU allocation is based on solving the following equation,

$$\max_{\rho_{n,k}^m} \sum_{n=1}^{N}\sum_{n=1}^{N} \rho_{n,k}^m \phi_{n,k}[t], \tag{5.10}$$

where $\phi_{n,k}[t] = \frac{Q_n[t] r_{n,k}[t]}{r_n^{\min}}$. The problem can be solved with classical Hungarian method represented in [100].

5.3.2 Graph Coloring

After solving the problem of RU allocation for each AP, in this section, the solution for the main issue of interference between overlapping-STAs with graph coloring is explored. First an introduction of graph coloring is represented and after that the main proposed algorithm is presented.

A graph is a basic mathematical concept used to represent relationships or connections between objects or nodes, which is understood as a diagram comprising points and lines. In other words, a graph G is defined by a non-empty finite set $V(G)$ of elements called vertices (or nodes) and a finite family $E(G)$of unordered pairs of (not necessarily distinct) elements of $V(G)$ called edges. An edge joining two vertices, for example v and w, is typically abbreviated to vw. The set of vertices is known as the vertex set, and the collection of edges is referred to as the edge set. Graphs can include loops, which are edges connecting a vertex to itself, and multiple edges, indicating several connections between the same pair of vertices. If a graph contains no loops or multiple edges, it is specifically termed a simple. The representation of a graph emphasizes its connectivity rather than any metrical properties, such as distance or straightness of lines graph[101].

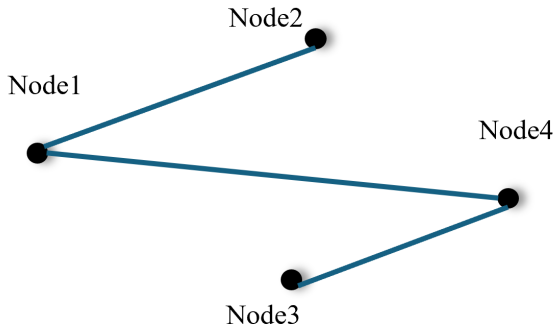


Fig. 5.2 Simple graph example.

Graph coloring can be applied to these defined graph structures. It means to assign colors to elements of a graph, commonly its vertices or edges, under specific rules. For instance, in vertex coloring, adjacent vertices must have different colors. [101].

5.3.3 Proposed Algorithm

The RU allocation problem defined in (5.4a)- (5.4e) is NP-hard, not feasible and the solution cannot be obtained with polynomial time complexity. To determine a feasible solution, we propose a heuristic algorithm based on the drift-plus-penalty method inspired from [97] and a graph coloring based on WMM algorithm. In this algorithm, the Sharing-AP initially gathers information on the channel amplitude response of overlapping STAs, and the BSSs in which these overlapping STAs are located. Then, the Sharing-AP is able to create the undirected graph $G = (\mathcal{N}, \mathcal{E})$ consisting of the set $\mathcal{N}$ of STAs' nodes and the set $\mathcal{E}$ of edges. The edges are between any two STAs' nodes located in at least two same BSSs whereas the nodes are not from the same parent AP. The Sharing-AP's role is to allocate RUs in a way that no two nodes with a common edge get the same color [102], while it chooses the best color (RU) for the nodes concerning their channel gain ($h^m_{k,n}$) that can give the highest data rate for the STAs which is based on WMM algorithm [97] and the Welsh-Powell algorithm[103]. After Sharing-AP colors the graph nodes, then, all APs attempt to assign the unallocated RUs to the rest of the STAs based on the WMM algorithm.

In Fig. 5.3 the proposed graph coloring model in the 4 OBSSs is depicted, while none of the STAs with the connected edges have the same color. Therefore, the assumptions that need to be considered for the graph coloring based on the WMM algorithm, are as follows:

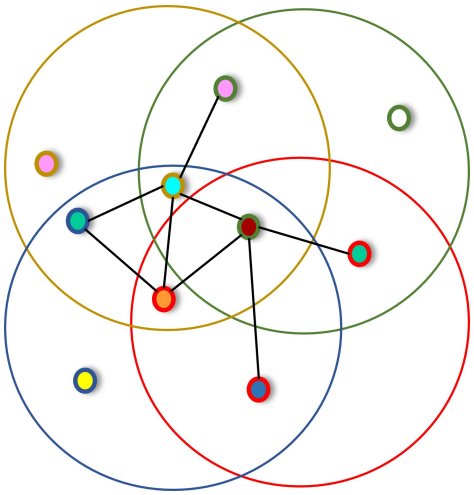


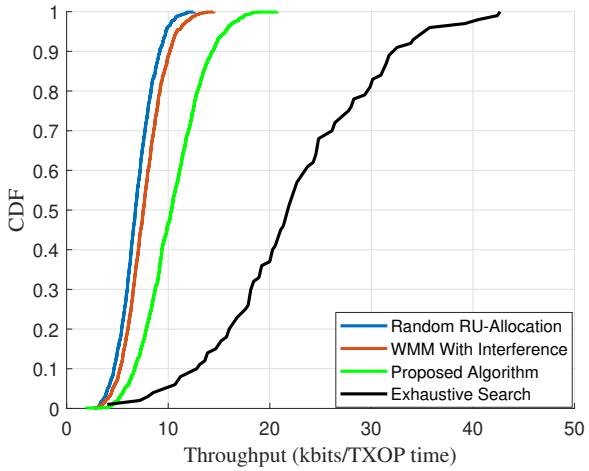
Fig. 5.3 Undirected Graph between overlapping STAs with colored nodes.

- All the APs have a fixed 20 MHz bandwidth, with 9-26-tones RUs.
- Each STA can get a maximum of one RU.
- In graph connection the edges are only between the STAs with different parent APs that are in at least two same BSSs.

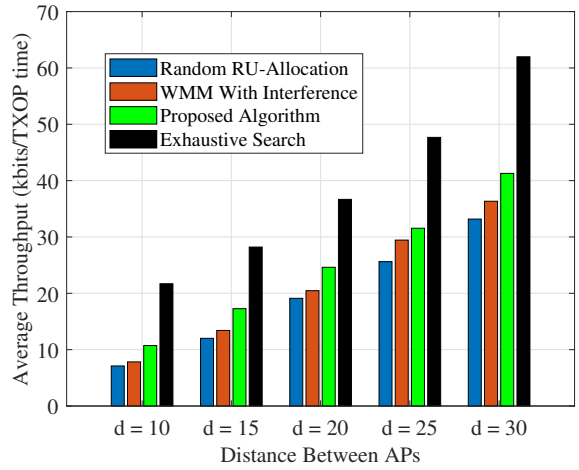
Algorithm 4 summarizes the proposed algorithm.

Algorithm 4: The proposed RU Allocation Technique

for $t = 1, 2, \ldots, T_{total}$ **do**
  Overlapping STAs RU allocation;
  Sharing-AP selection based on lowest backoff time;
  Sharing-AP and Shared-APs signalling to receive information about overlapping STAs' channel amplitude response and the BSSs in which these overlapping STAs are located;
  Sharing-AP creates a graph in which the graph nodes are overlapping STAs and the edges are between any two overlapping STAs located in more than one same BSSs from different assigned APs;
  Sharing-AP colors the nodes based on the Welch Powell and WMM algorithms in 5.3.1;
  Non-overlapping STAs RU allocation;
  for $m = 1, 2, \ldots, M$ **do**
    m-th AP for the rest of its STAs allocates RUs based on WMM algorithm in 5.3.1;
    for $n = 1, 2, \ldots, N$ **do**
      After RU allocation to all the STAs, m-th AP calculates $r^m_{n,k}$ for n-th STA and returns it;
    Update WMM algorithm;
  Repeat this algorithm until T_{total};
To evaluate the performance, compute $r^m_{n,k}$ considering SINR.



(a) Empirical CDF of mean aggregate throughput of total number of STAs



(b) Mean aggregate throughput of total number of STAs per APs' distance in meters

Fig. 5.4 Mean aggregate throughput of the total number of STAs for the scenario where $M = 4$, $N = 5$, $K = 3$.

5.4 Simulation Results

To evaluate the performance of the proposed algorithm, we consider a D-MAPC-based Wi-Fi network with a joint C-OFDMA and C-SR model, under residential settings. A residential Wi-Fi coverage area represented with a radius of $d_{\max}$ is considered. All APs are positioned at the center of BSSs. STAs are uniformly distributed within the BSS, following a uniform distribution of $\mathcal{U}(0,1)$, ensuring that all locations within the BSS are equally likely. We consider the 26-tone RU allocation

while all STAs transmit in every TXOP. The performance of the proposed algorithm is compared with that of a random RU allocation as a serving lower bound and an exhaustive search representing the upper bound. The simulations are performed on a compute node with 50 CPU cores and 80 GB of RAM. Additionally, another benchmark scheme is introduced, derived from [97], where APs do not coordinate with each other for RU allocation to their respective STAs. This means there is no interference management in OBSSs. In the simulation results, this scheme is referred to as WMM with Interference. Table 5.1 provides the simulation parameters used in this chapter. To determine the data rate, first, the RSSI for each STA is calculated. Using MCS in the look-up Table 5.2, the data rate is calculated with the 26-tone configuration. It is assumed that each AP has only its STAs' RSSI, in other words, each AP does not have any information about the interference that each STA from other APs can bring to it. The RSSI of n-th STAs in m-th AP with k-th allocated RUs is as follows:

$$\mathrm{RSSI}_{n,k}^{m} = 10\log_{10} p - \mathrm{PL}_{n}^{m}(d) + 10\log_{10}\left(g_{n,k}^{m}\right)(\mathrm{dBm}). \tag{5.11}$$

Then based on Table 5.2 STA's MCS level and its R_c is chosen and for each RU the data rate is calculated as follows:

$$r_{n}^{m} = N_R \times R_c \times \alpha \times \frac{T_{\mathrm{TXOP}}}{T_{\mathrm{OFDM}}} \quad (\mathrm{bits/\ TXOP\ time}). \tag{5.12}$$

Following the data rate calculation based on the proposed algorithm, the optimal RU combination is assigned to each AP. Finally, after RU allocation, the network performance is evaluated by analyzing the impact of shared RU interference between different APs (co-channel interference). The SINR is then computed, and the data rate, considering the SINR, is determined using (5.12).

Fig. 5.4 presents the mean aggregate network throughput per total number of STAs where the scenario is with 4 APs, 5 STAs per each AP and 3-26-tones RUs for each AP. Due to the high computational complexity of the exhaustive search algorithm each Monte Carlo iteration for the exhaustive search curve, involving 4 APs, 5 STAs per AP, and 3 RUs per AP, takes around 8 hours to complete with a compute node with 50 CPU cores and 80 GB of RAM, which shows its high complexity. Fig. 5.4a illustrates the empirical CDF of the mean aggregate throughput of the total STAs where the proposed algorithm achieves a higher mean aggregate network throughput near 25% compared to the WMM algorithm with no co-channel interference management and 30% compared to the random RU allocation algorithm. While the exhaustive search

algorithm with the unrealistic assumptions of the perfect channel state information of all the STAs with SINR calculation chooses the maximum throughput of the RU combination set for the proposed system model. However due to its high complexity and the unrealistic assumptions, it cannot be implemented in practice. Fig. 5.4b demonstrates as the distance between APs increases and the overlapping area decreases, the network's overlapping STAs generate less interference, which also demonstrates the interference level, as closest the APs are the interference level will be higher. As shown, the proposed algorithm achieves a better network throughput with approximately 30% improvement compared to the WMM algorithm due to its effective interference management. In this figure, exhaustive search depicts the upper bound.

Fig. 5.5 shows the mean aggregate network throughput of all STAs where 4 APs, 12 or more STAs per each AP, and 9-26-tones RUs for each AP is considered. Fig. 5.5a demonstrates the proposed algorithm mean aggregated throughput and shows about 30% increment compared to the WMM algorithm with no co-channe interference management and about 60% with random RU allocation as the lower bound. Additionally, Fig. 5.5b demonstrates that the aggregate throughput increases as the APs are positioned farther apart. Meanwhile, Fig. 5.5c reveals that the proposed algorithm continues to enhance the total network throughput even as the available STAs per AP increases. In contrast, the random RU allocation algorithm maintains a constant total network throughput of 7.10 Kbits/TXOP time, regardless of the number of STAs. For example, when $N = 15$, the proposed algorithm achieves a mean aggregate throughput that is 20% higher than that of the WMM algorithm without interference management.

The computational complexity of the considered problem formulation, assuming that STAs exceeds the number of RUs, is $O(N!^{(M)})$. However, the proposed algorithm reduces this complexity to $O\left((M \times N) - N'\right)^{(3)} + N' + \mathcal{E})$. Here, the computational complexity of the WMM algorithm combined with Hungarian RU assignment for each AP is $O\left((M \times N) - N'\right)^{(3)}$ [104], while the approximate computational complexity of graph coloring is $O(N' + \mathcal{E})$ [103], where $\mathcal{E}$ represents the number of edges between any two overlapping STAs. In this context, N, M, and N' denote the number of STAs in each BSS, the number of APs, and the number of STAs in the OBSSs, respectively.

Table 5.1 Simulation parameters

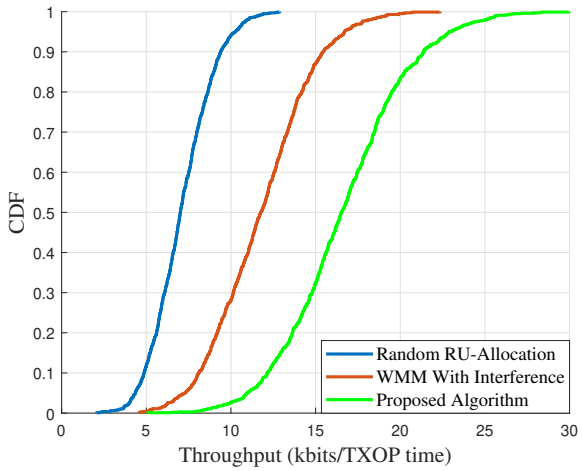
Parameter	**Notation Description**	**Value**
$d_{\max}$	Radius of each BSS	20 m
P	Power of each STA	20 dBm
T_{OFDM}	Duration of OFDM Symbol	16μs
T_{TXOP}	Duration of uplink OFDMA transmission	3.2 ms
T_{total}	Total number of TXOP for simulation in each Monte Carlo	400
V	drift-plus-penalty control parameter	100
d_m	Distance between two APs	20 m
N_R	Number of RU-tones	26

Table 5.2 MCS for the 20MHz channel and RUs of 26 subcarriers

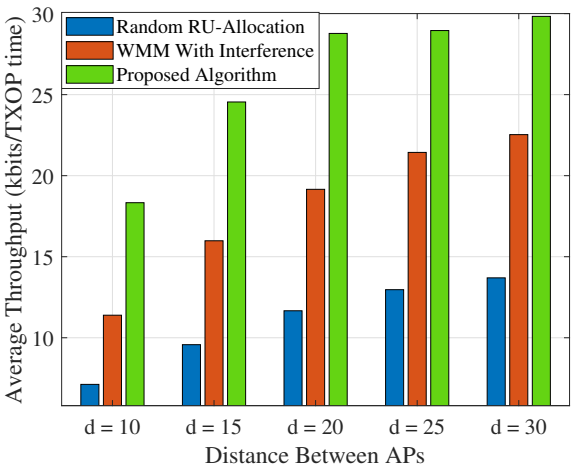
Index	**Modulation**	**Bit per symbol** α	**Coding rate** (R_c)	**Min. SNR**	**Min. RSSI** (dBm)
1	BPSK	1	1/2	2	−82
2	QPSK	2	1/2	5	−79
3	QPSK	2	3/4	9	−77
4	16-QAM	4	1/2	11	−74
5	16-QAM	4	3/4	15	−70
6	64-QAM	6	2/3	18	−66
7	64-QAM	6	3/4	20	−65
8	64-QAM	6	5/6	25	−64
9	256-QAM	8	3/4	29	−59
10	256-QAM	8	5/6	31	−57

While the proposed fairness-based scheduling mechanism demonstrates clear performance benefits under the considered simulation settings, practical deployment introduces additional challenges. In particular, the scheduling decisions assume stable channel estimates over the scheduling interval and neglect the impact of fast-varying channel dynamics and feedback imperfections. Moreover, the fairness optimisation becomes increasingly complex when the number of users and resource units grows, since the allocation space expands significantly. In practical implementations, this may require limiting the scheduling horizon or applying simplified allocation rules to meet latency constraints. Furthermore, interoperability with existing IEEE 802.11 scheduling procedures may require additional signalling and protocol adaptation. Future work could therefore investigate scalable heuristic

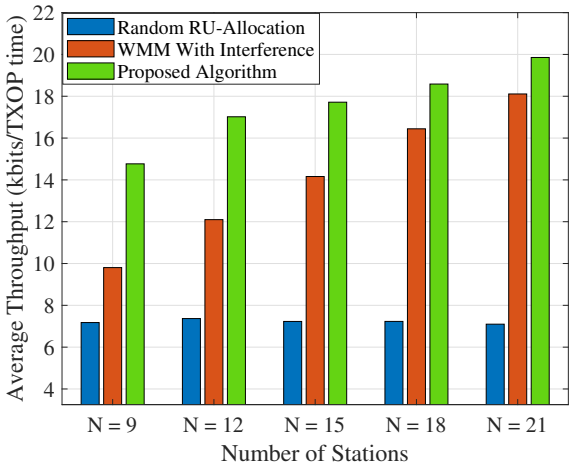
schedulers and real-time implementation strategies that maintain fairness while reducing operational complexity.



(a) Empirical CDF of Mean aggregate throughput of all STAs.



(b) Mean aggregate throughput of all STAs per APs' distance in meter.



(c) Mean aggregate throughput of all STAs per number of STAs in each BSS.

Fig. 5.5 Mean aggregate throughput of the total number of STAs for the scenario where $M = 4$, $N = 12$, $K = 9$.

5.5 Summary

In this chapter, we proposed a fixed RU scheduling algorithm for the multi-AP coordination system to maximize the total network throughput while satisfying the average QoS of each STA using a WMM-based graph coloring approach. The Sharing-AP received its requested information about the overlapping STAs channel amplitude response and the BSSs in which these overlapping STAs are located from the Shared-APs. Simulation results highlighted two main points first the importance of RU allocation in the MAPC system model for interference management. Second, the total mean network throughput increased by 30% with the proposed algorithm due to the interference management in the OBSSs with low complexity.

Chapter 6

Fair Variable RU Allocation in C-OFDMA

In this chapter, a variable-sized RU allocation scheme based on heterogeneous channel condition of STAs within a C-OFDMA system model is proposed. To examine the impact of heterogeneous channel conditions on RU size allocation, a toy scenario is analyzed and demonstrated. Subsequently, a max–min throughput fairness formulation is presented, aiming to ensure that all STAs achieve identical throughput. Since the objective is to design an RU allocation strategy that achieves max–min fairness while minimizing algorithmic overhead and enhancing practicality, the allocation algorithm is considered within a scheduling interval. The formulated problem is proven to be NP-hard, hence, a DRL algorithm is proposed to address the RU allocation challenge. Finally, proportional fairness throughput-based RU allocation is formulated within the same system model, and a closed-form solution is derived.

6.1 Introduction

RU allocation and scheduling algorithms in the IEEE 802.11be standards and beyond are required due to the increasing demand for network traffic and emerging new QoS requirements for emerging new use cases and applications. Most of the work in Wi-Fi networks for RU allocation has considered the fixed-RU allocation [50, 53, 97, 47]. For example, in [50], the authors investigated the different distribution of random and scheduled RU access for throughput maximization, with a fixed-RU allocation assumption. In [53], authors proposed variable-sized RU allocation with the goal

of sum rate maximization; however, the proposed problem formulation is based on IEEE 802.11ax, where each STA can only get one RU and each RU can only be assigned to one STA, which imposes more constraints on IEEE 802.11be and beyond. Authors in [105] considered a joint problem of RU allocation, beamforming design, channelization, and MCS selection. Due to the high complexity of the problem, heuristic algorithms are suggested. However, the work is based on IEEE 802.11ax with strict conditions, while the direct connection between channelization and MCS has not been investigated, and the main goal was the total system throughput maximization. In addition, in [35] authors proposed a variable-sized RU allocation with proportional gain with respect to a defined priority and fairness factor. In this work, the MCS is the same for all the STAs in the network.

It is clear that, despite different works and algorithms, RU allocation, especially with new applications and standardization, is an open problem. In addition, to the best of our knowledge, there has not been a clear investigation into the MCS index and the variable-sized RU allocation. Therefore in this chapter, first we investigate the variable-sized RU allocation and heterogeneous MCS index effect on this with practical assumptions in IEEE 802.11bn in the C-OFDMA system model. Channel gain and MCS index are used interchangeably in this chapter. A max-min throughput optimization is formulated. As the proposed problem formulation is NP-hard, a DRL algorithm is recommended to solve the high-complexity problem. Simulation results demonstrate the proposed DRL RU allocation algorithm in C-OFDMA with regard to maximization of the minimum STA's throughput and Jain's index fairness. In the next part of the chapter, an RU allocation with variable MCS index for proportional-throughput fairness is investigated in the considered system model. In addition, a closed-form solution for this problem formulation is derived.

6.2 System Model

The considered system model is based on a coordinated multi-AP framework. The set of APs is denoted by $\mathcal{M} = \{1, 2, \ldots, M\}$, and the set representing all available STAs in the network is defined as $\mathcal{N} = \{1, 2, \ldots, N\}$. For simplicity of analysis, it is assumed that the STAs are pre-assigned to the APs, and each AP is responsible for managing an equal number of STAs. Accordingly, the m-th AP serves $N_m = \frac{N}{M}$ STAs in uplink transmission. The transmission signaling is out of scope of this chapter.

The main goal of this chapter is to assign variable RU sizes based on STAs channel gain, considering the new IEEE 802.11bn standard. For practical consideration and

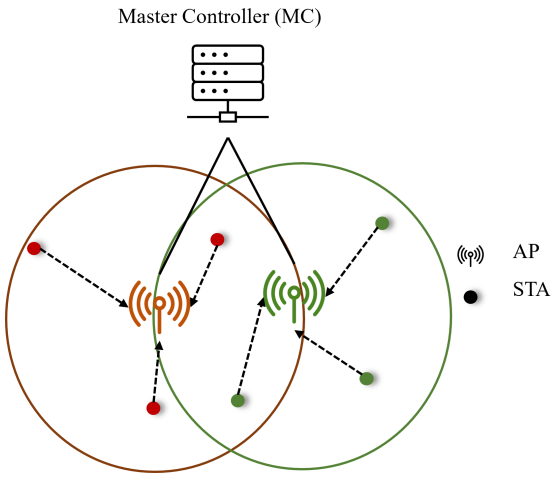


Fig. 6.1 An example of an OBSS with 2 APs and 3 STAs assigned in each AP's BSSs, in which C-OFDMA is applied using a master controller.

lowering overhead, it is assumed that the APs are agreed to have coordination for a scheduling interval. Within this scheduling interval, it is assumed L TXOPs are available and $L = N$, and in each TXOP, all STAs may perform transmission provided that they are allocated the appropriate RUs. It is considered that the total bandwidth available within each TXOP is B, and the minimum RU size is w. Consequently, the total bandwidth can be expressed as $B = \rho_{\max} \times w$, where $\rho_{\max}$ denotes the maximum number of RUs in B. In this chapter B as the total available bandwidth is 160 MHz, and the entire bandwidth can be dynamically shared among all STAs within the system model. The minimum RU size w is set to 20 MHz. As a result, the RU allocation vector is defined as $\mathcal{R} = [0, 20, 40, 60, 80, 100, 120, 140, 160]$ MHz, as illustrated in Fig. 6.2.

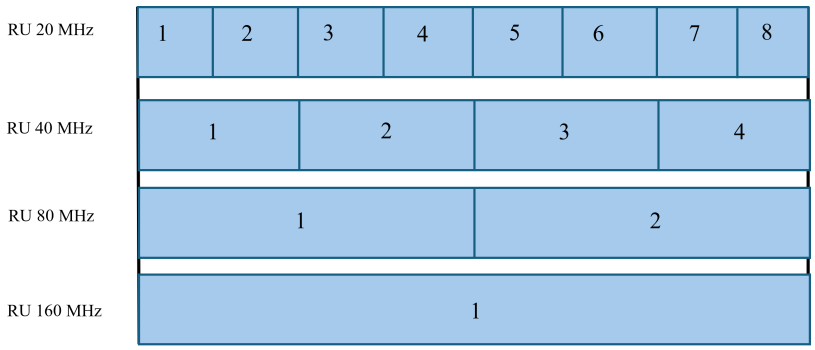


Fig. 6.2 Variable RU-sized allocation with consideration of 160 MHz bandwidth.

The channel gain of n-th STA with its m-th AP is $h_{n_m} = \mathrm{PL}_{n_m}(d) \cdot g_{n_m}$. In which g_{n_m} is the small-scale fading of the n-th STA and the m-th AP. $\mathrm{PL}_{n_m}(d)$, defined in

(6.1), represents the path loss of the k-th STA associated with the m-th AP. based on the Residential cases of the TGax model [98] and is as follows:

$$\begin{aligned}\mathrm{PL}_{n_m}(d) = 40.05 + 20\times\log_{10}(f_c/2.4) + 20\times\log_{10}(\min(d,5)) \\ + (d>5)\times 35\times\log(d/5),\end{aligned} \tag{6.1}$$

where f_c denotes 5 GHz, and d is the distance between the transmitter and the receiver in meters.

Based on each STA's channel gain and with assumption of equal maximum transmit power for each STAs, the RSSI is calculated as follows,

$$\mathrm{RSSI}_{n_m} = 10\log_{10}\mathrm{P_{n_m}} - \mathrm{PL}_{n_m}(d) - 10\log_{10}\left(g_{n_m}\right)(\mathrm{dBm}). \tag{6.2}$$

Where $10\log_{10}(P_{n_m})$ denotes the transmit power of the corresponding transmitter expressed in logarithmic scale (dB), obtained by converting the linear power value P_{n_m}. Then, using the RSSI threshold, the MCS index is chosen based on the Table in [106].

Then data rate of the n_m-th STA is calculated based on its assigned RUs and the selected MCS index derived from its channel gain. Then each TXOP throughput can be calculated as follows,

$$\eta^l_{n_m} = \frac{K^l_{n_m}}{T_l}, \tag{6.3}$$

where $K^l_{n_m}$ is the transmitted bits of n_m-th STA in the l-th TXOP in a way that each STA can be assigned RUs from different TXOPs of scheduling interval. T_l is the TXOP duration, which here is fixed to the maximum TXOP duration, and based on its signaling, it has the following overheads and data transmission.

$$\begin{aligned}T_l = T_{Backoff} + T_{MAP-RTS} + T_{SIFS} + T_{MAP-CTS} + T_{SIFS} + T_{MAP-TF} + T_{SIFS} \\ + T_{Data} + T_{SIFS} + T_{ACK} + T_{DIFS},\end{aligned} \tag{6.4}$$

where $T_{MAP-RTS}$ and $T_{MAP-CTS}$ are the multi-AP request to send and clear to send signals, respectively. T_{MAP-TF} is the multi-AP trigger frame to inform STAs about their allocated RUs, and other information for scheduling and transmission. The aggregated packets each STA can transmit during the T_l. Then, based on the TXOP duration, STAs' MCS indices and allocated RU, the total number of packets each STA

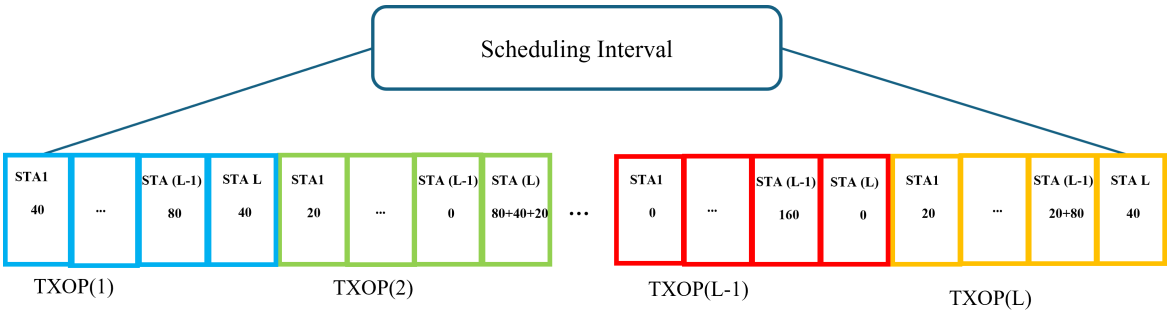


Fig. 6.3 Allocated RUs in MHz per scheduling interval.

transmits can be calculated. Let's assume the packet length of n-th STA is PK_{n_m} and the number of aggregated packets for n-th STA is $Z_{n_m}^{PK^l}$, therefore

$$K_{n_m}^l = PK_{n_m}^l \times Z_{n_m}^{PK^l}. \tag{6.5}$$

To calculate $Z_{n_m}^{PK}$, it is first necessary to determine the number of packets that can be transmitted within the l-th TXOP, denoted by T_l. First, assume that the physical data rate can be calculated as follows:

$$r_{n_m}^l = \frac{\rho_{n_m}^l \times w \times \mathrm{CR}_{n_m}^l \times \alpha_{n_m}^l}{T_{\mathrm{OFDM}} \times SC} \quad \text{(bits/sec)}, \tag{6.6}$$

where based on calculated $RSSI_{n_m}$ from Table in [106] STA's MCS index, $\alpha_{n_m}^l$, and its coding rate, $\mathrm{CR}_{n_m}^l$, are derived. In addition to the assigned RU, $\rho_{n_m}^l \times w$, the subcarriers spacing is denoted as SC. Finally $Z_{n_m}^{PK^l}$ is calculated as follows,

$$Z_{n_m}^{PK^l} = \frac{r_{n_m}^l \times T_{Data}^l - \mathrm{L}_{\mathrm{tail}}^l}{PK_{n_m}^l + H_{\mathrm{MAC}} + H_{\mathrm{PHY}}}, \tag{6.7}$$

where H_{MAC} and H_{PHY} are MAC and PHY headers. L_{tail}^l is the number of tail bits appended at the end of the coded transmitted data bits. In the next section, the proposed problem formulation is motivated by the heterogeneous channel conditions experienced by STAs in the context of RU allocation under the max–min fairness criterion which in the considered system model is elaborated with a simple example.

6.2.1 Motivation of the Proposed Problem Formulation

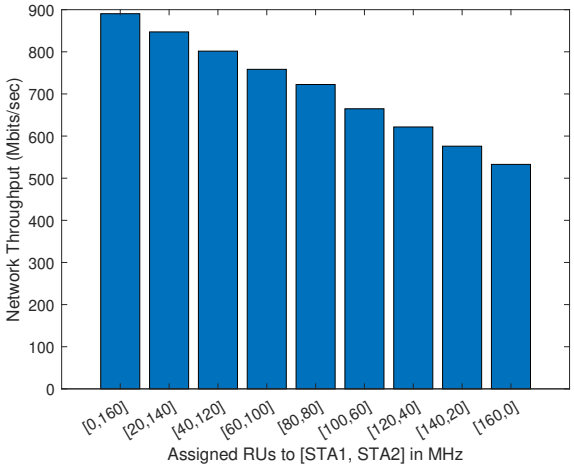
In this section, we dive into a very simple scenario to demonstrate the effect of channel gain of STAs for variable RU size allocation. For simplicity, consider a scenario with two APs, each serving a single STA, resulting in a total of two STAs in the network. It is assumed one TXOP where the STAs aggregated packet based on their data rate is calculated as (6.7). The MCS index is variable and determined based on the calculated SNR in Table 5.2.

Fig. 6.4 illustrates the network's performance where the MCS index of STA1 is 7, STA2 is 11 and there are several options for RU assignment in this model.

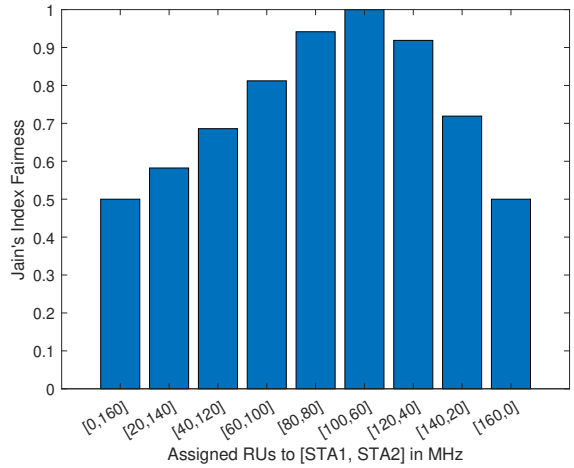
Initially, Fig. 6.4a shows while STA2 has is with favorable channel conditions compared to STA1, the maximum throughput is achieved in the network by allocating the 160 MHz to STA 2. This is due to the fact that the fastest STA or in other words the STA with better channel gain, accessing the whole bandwidth, gets the highest network throughput. However, this is not fair at all, as the STA1 has not given any resources for transmission. Fig. 6.4b demonstrates the mentioned point about fairness. So allocating all the resources to only STA2 will provide the worst Jain's index fairness respectively in the network. While also allocating all the resources for STA1 makes the fairness worse.

In addition Fig. 6.4b demonstrates that to have the highest Jain's index fairness in this toy scenario, the best option is to allocate 100 MHz bandwidth to STA1 and the rest to STA2. Jain's index fairness is a measurement of fairness that provides max-min fairness in the system. It means that the highest Jain's index fairness can provide quite equal throughput for all the STAs in the network. This could be required for scenarios where there are bottleneck STAs that cannot get resources, especially while in a dense environment, potentially leading to STAs with weaker channel conditions achieving the minimal throughput [107]. Therefore in this section, our main goal is to allocate RUs in a way that the STAs with the lowest channel gain can also increase their throughput.

Another important point is that to provide max-min fairness and as the resources are limited, it is assumed to provide fairness per each scheduling interval that has L TXOPs. That is, while the AP may not be able to provide the required throughput to each STA in a given TXOP, throughput fairness across all STAs is expected in the following TXOP within the scheduling interval, as shown in Fig. 6.3. In addition, it is not practical for an uplink system model with high overhead due to coordination signaling and trigger frame, to schedule RUs only for one TXOP. It should be noted



(a) Network Throughput versus all the possible assigned RU sets.



(b) Jain's index fairness versus all the possible assigned RU sets.

Fig. 6.4 Network's performance based on possible assigned RU sets.

that although the uplink signaling is assumed but the proposed model and problem formulation is applicable for downlink as well.

6.3 Max-Min Fairness

Max-min fairness main goal is to maximize the minimum throughput or other KPIs in a network. In addition with max-min fairness, it is possible to avoid scenarios where at least one user's QoS has not been satisfied [107]. In addition, authors in [97], proposed an algorithm to solve the max-min problem with RU allocation based on

channel fading. Therefore as observed in the previous section, the max-min fairness with variable-sized RU allocation is an open problem. Especially, it is different compared to throughput optimization and needs attention, as it is essential in new applications that the slow STAs required QoS could also be satisfied. In this section, the proposed problem formulation for max-min fairness for STAs' throughput is proposed. As the problem is NP-hard and has a huge search space, a DRL algorithm is proposed to solve the proposed problem.

6.3.1 Problem Formulation

The considered max-min fairness objective is as follows,

$$\underset{\rho_n^l}{\text{maximize}} \quad \min_{n \in \mathcal{N}} \eta_n, \tag{6.8a}$$

$$\text{subject to} \quad 0 \leq \rho_n^l \leq \rho_{\max}, \forall n \in \mathcal{N}, \tag{6.8b}$$

$$\sum_{n=1}^{N} \rho_n^l = \rho_{max}, \forall l \in \mathcal{L}, \tag{6.8c}$$

where ρ_n^l is the complex coefficient characterizing the 20MHz transmission for n-th STA in l-th TXOP. It defines the chosen RU size from $\mathcal{R}$ set for n-th STA in each TXOP. L is the TXOP per each scheduling interval. η_n is the total calculated throughput for the n-th STA in the scheduling interval. That is as follows,

$$\eta_n = \sum_{l=1}^{L} \eta_n^l, \forall n \in N. \tag{6.9}$$

Note that the total allocated RUs per TXOP is 160 MHz and it is also assumed that STAs are full-buffer and always have packets to transmit. Objective function (6.8a) is the max-min fairness of throughput of all STAs. Constraint (6.8b) ensures the maximum allocated RU-size for each STA is ρ_{max}. While constraint (6.8c) confirms the total amount of assigned RU are equal ρ_{max}.

The proposed optimization is an NP-hard problem as the search area is a mixed-integer programming problem [105]. In the next section we solve the considered problem with DRL approach.

6.3.2 Proposed DRL Algorithm

The formulated RU allocation problem is NP-hard. Deep learning (DL) algorithm has been promising for RU allocation however DL algorithms are data based algorithms which data collection is time consuming. Therefore in this chapter the DRL scheduling algorithm is considered for RU scheduling algorithm, moreover, as the wireless communications environment is fluctuate in time, using DRL is an advantage for these environment as it can adjust itself to the uncertainty. The proposed DRL algorithm is centralized, therefore it is assumed that the centralized controller allocates RUs to the STAs with the proposed DRL algorithm. It should be noted that the signaling of C-OFDMA for this system model is outside of the scope of this work.

To define the DRL-based RU allocation algorithm, we model the problem as a MDP to solve the complexity of the proposed problem formulation. In which the state S represents the current network's state space, A defines the action space. The reward R reflects the feedback based on the action the agent has taken. The transition T describes the probability of moving from one state to another after taking an action, and the policy π defines the strategy for selecting actions based on the current state to maximize the cumulative reward over time [76]. Therefore the elements of our proposed DRL-based RU allocation algorithm are as follows.

State Space

The goal is that the agent based on its previous selected RUs in a scheduling interval learns to assign RUs and the MCS index effect will be applied in the objective function of max-min fairness. In other words the agent based on its previous assigned RUs and channel gain effect on the STA with minimum throughput learns the best RU allocation assignment for the proposed problem formulation. Therefore the state space is defined as a matrix in which its elements are the assigned RUs in the previous scheduling interval. The state space is the assigned RUs in all the TXOPs of scheduling interval, and it is a $L \times N$ matrix (6.10). Each row represents a TXOP and all the N columns are the total available STAs in the network. Each element of the state matrix is the number of RUs assigned to the n-th STA in the l-th TXOP, which is ρ_n^l. Therefore the total state space is as follows,

$$S_{L\times N} = \begin{bmatrix} \rho_1^1, \rho_2^1, \ldots, \rho_N^1 \\ \rho_1^2, \rho_2^2, \ldots, \rho_N^2 \\ \vdots \\ \rho_1^L, \rho_2^L, \ldots, \rho_N^L \end{bmatrix}, \tag{6.10}$$

where constraints (6.8b) and (6.8c) should be satisfied for each element of the matrix and for each row of the matrix respectively.

Action Space

The agent in each step can see the current state which is each RUs that previously assigned for each STA in all the TXOPs. Therefore the action in each step is a vector with first-norm zero, in which one of the element is 1, one other element is -1 and with the an index element of the l row index. Which means the first-norm vector should be applied to the l-th TXOP. This means the action is added to the l-th TXOP of the scheduling interval, where the agent has taken the action based on the ϵ-greedy algorithm. Therefore the action space is all the combination of the 1, -1, and $N-2$ zeros and the the vector combinations with the TXOP set $\mathcal{L} = \{1, 2, \ldots, L\}$. Which shows the selected first-norm vector should be added to which of the coordinated-slots. The action space for a N STAs is as follows,

$$\begin{aligned} A = \Big\{ & [1,-1,0,\ldots,0,0,l], \ldots, [1,-1,0,\ldots,0,0,L], \\ & \ldots, [0,0,\ldots,1,-1,l], \ldots, [0,0,\ldots,1,-1,L], \\ & \ldots, [0,0,\ldots,0,0,l], \ldots, [0,0,\ldots,0,0,L] \Big\} \end{aligned} \tag{6.11}$$

where $\Big[1,-1,0,\ldots,0,0,l\Big]$ means the vector $\Big[1,-1,0,\ldots,0,0,l\Big]$ is added to l-th row of state matrix. Also $\Big[0,0,\ldots,0,0,l\Big]$ means that not only all the other rows RU assignment has not changed but also the l-th TXOP keeps the previous RUs as well, therefore the total action space size is the $(\binom{N}{2}+1)\times L$ of $(1\times N)$ vectors.

Reward

As our main goal is to maximize the minimum STAs' throughput based on their MCS index, we formulate the reward as the same as the defined objective function, and it is as follows,

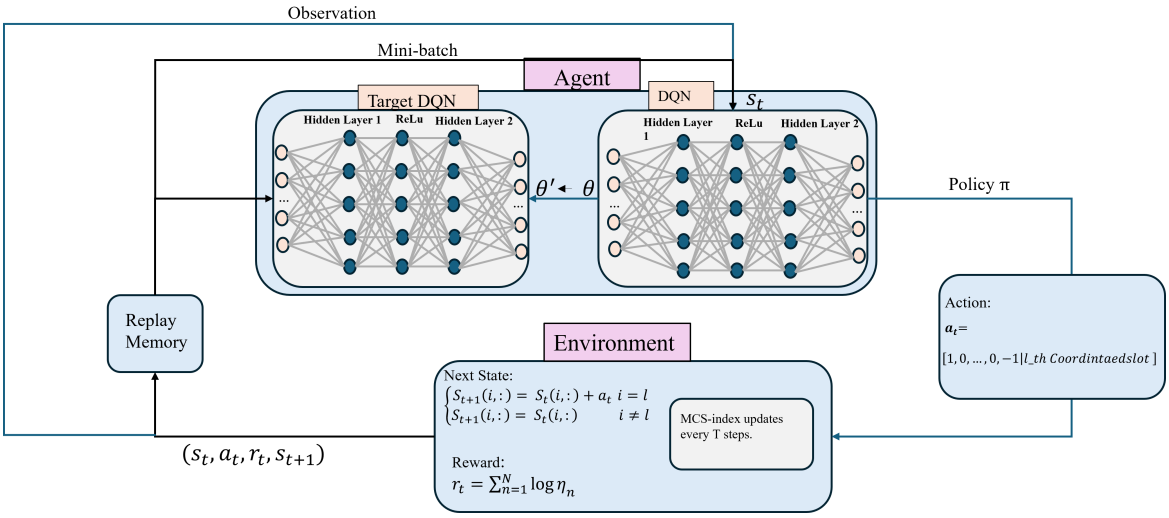


Fig. 6.5 Proposed DQN Algorithm-based RU allocation.

$$R = \min_{n \in \mathcal{N}} \eta_n \tag{6.12}$$

The DRL-based RU allocation algorithm is elaborated in Algorithm 5. The

Algorithm 5: The proposed DRL-based RU allocation algorithm

Initialize Model: Initialize replay memory and Q-network
for $Episode = 1, 2, \ldots$ **do**
  Initialized the environment state S^0
  for $t = 1, 2, \ldots, \tau$ **do**
    DQN agent chooses action a^t from A set using ϵ-greedy policy.
    Execute action a^t
    Receive reward r^t and next state $S^{(t+1)}$
    if $t \mod T = 0$ **then**
      Change MCS index based on STAs channel gains
    Store experience (S^t, a^t, r^t, S^{t+1})
    Sample mini-batch from memory and update Q-network

proposed algorithm in each episode, after initializing the environment state S^0, once the agent takes an action, it executes the selected action on the selected l-th TXOP. Then the reward is calculated, and the agent moves to the next action. In order to allow the agent to learn, it is assumed that the same MCS index for the STAs is used for T steps. Then the proposed reward is calculated, and the next state is updated. In other words, it could help the agent to learn based on different MCS indices to choose the best variable RU allocation.

6.3.3 Simulation Results

To assess the effectiveness of the proposed algorithm, multi-AP coordination with a centralized controller framework is employed within a residential Wi-Fi coverage scenario. All APs are assumed to be centrally located within the BSSs. Within each BSS, the STAs are uniformly distributed over the coverage area, with their locations relative to the serving APs modeled as a uniform random variable $\mathcal{U}(0,1)$. The APs are situated within mutual transmission range and operate on the same frequency channel. However, under C-OFDMA, RUs are allocated such that reusing RUs cannot occur within each TXOP. The simulation results are generated using MATLAB. The DQN network architecture consists of an input layer followed by a fully connected layer, and a ReLU activation layer, ReLU, which introduces non-linearity. This is followed by another fully connected layer, which further processes the features. The output from this layer is passed to the q-values, and the hidden layers' nodes are 256.

The proposed DRL algorithm is compared with Equal-sized and Random RU allocation benchmarks. Also an Exhaustive Search is compared as the upper bound for the total available 3 STAs. The Equal-sized RU allocation is a baseline algorithm where in each TXOP per scheduling interval the STAs get the same RU size, for example with 3 STAs in the network, each of the STAs are assigned 80 MHz with random allocation in each TXOP, where the total bandwidth is 160 MHz. So either the total bandwidth is divided with equal-sized RU if it is applicable. Or in case the number of STAs available for scheduling exceeds the number of available RUs, each STA in each TXOP per scheduling interval assigned randomly the minimum sized RUs for the Equal-sized RU allocation benchmark. The random RU allocation assigned random variable-sized RUs to random STAs in each TXOP per scheduling interval. Moreover, the Exhaustive Search is suggested the upper bound to allocate the best variable RUs in the scheduling interval for the STAs in order to maximize the proposed max-min fairness. However, because of its high complexity it could only be simulated for 3 STAs, which is in practical scenarios it is impossible to use for each scheduling interval in a system.

In more details, the proposed problem formulation computational complexity approximately is $O\left(\frac{\rho_{max}^{N-1}}{(N-2)!}\right)^{(L)}$. As its computational complexity is very high, that is not possible to solve it in a real-time scenario. While the DQN computational complexity per each time step is related to the cost of a single forward pass through the network $O\left(f(d)\right)$ [81]. Therefore, for example, the proposed DRL algorithm can solve the problem after training with only 40 steps, for 4 STAs.

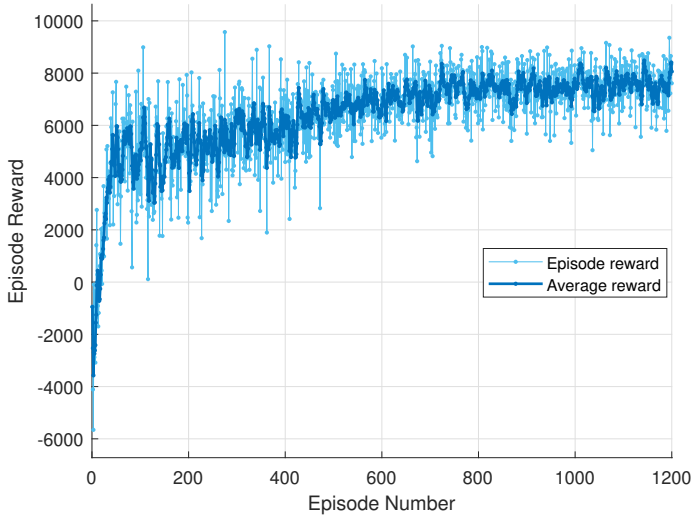


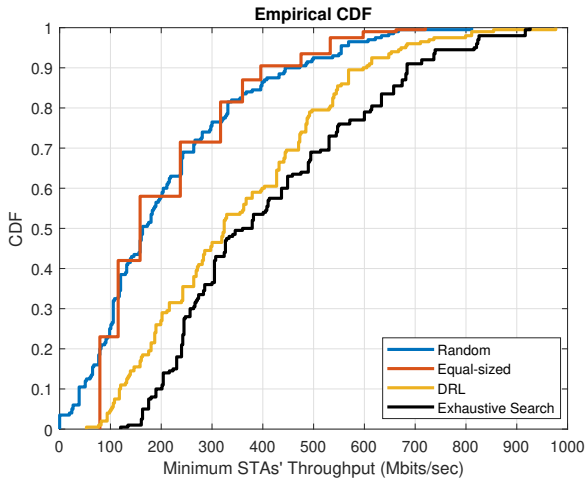
Fig. 6.6 Learning convergence of the proposed DQN agent for 4 STAs in the network. Where, an episode refers to a complete sequence of interactions between the DQN agent and the environment, starting from an initial state and ending when a terminal condition is reached.

Fig. 6.7 shows the CDF of guaranteed minimum rate of STAs and the Jain's index fairness in a scenario where the network has 3 APs where each of AP has 1 STAs. The main goal is to demonstrate with variable channel gains effect on guaranteed minimum throughput of STAs and the Jain's index fairness for all the benchmarks and the proposed algorithm, respectively. With the variable MCS index in 100 Monte-Carlos, in Fig. 6.7a it is observed the STAs' minimum throughput of all the benchmark schemes and the proposed DRL algorithm. The DRL algorithm in comparison with the Exhaustive Search shows it is following this upper bound. As it is observed the proposed DRL algorithm is about 10% lower than the Exhaustive Search. This shows the proposed DRL performance with lower complexity and only with 40 episode can achieve this, while Exhaustive Search space is 91125. The DRL algorithm in comparison with the random and equal-sized RU allocation, achieves about 50% higher minimum STAs' throughput. This shows its out performance to allocate RUs based on variable MCS indexes for a max-min fairness.

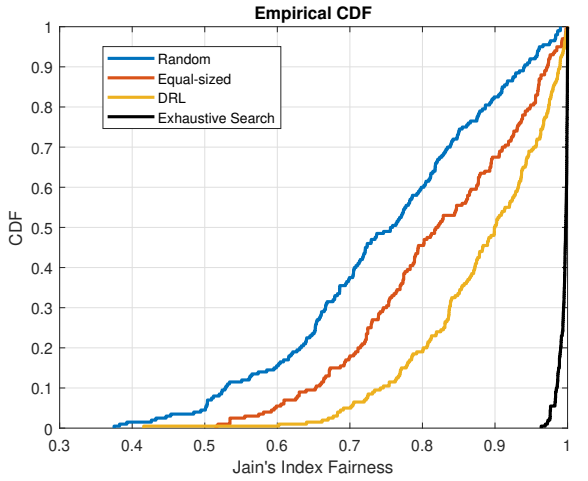
Fig. 6.7b demonstrates the outperformance of the proposed DRL algorithm in comparison with Equal-sized and Random RU allocation. Jain's index fairness measures how much all the STAs available in the network, in each different iteration simulation where each iteration has a different MCS index, compared to another iteration, can achieve the same throughput despite the variety in MCS index or

channel gain. As observed, the proposed DRL algorithm can achieve, for 50% of the total iteration simulations, more than 0.9 Jain's index fairness, while the corresponding values for Equal-sized and Random RU allocation algorithms are 0.82 and 0.75, respectively. Therefore, the proposed DRL algorithm can achieve a higher Jain's index fairness compared to the Random and Equal-sized RU allocation algorithms. In comparison with Exhaustive Search, the proposed DRL algorithm results in Jain's index fairness lower than 0.8 in less than 20% of the iteration simulations, while this number for Exhaustive Search is 0.98, which is expected as the proposed DRL does not know the exact RU allocation for max-min fairness and instead attempts to approximate it.

Following that, in Fig. 6.8, the the comparison of the 30%, 50%, and 90% CDF values of the minimum STAs' throughput under variable available STAs in the network is presented. The proposed DRL algorithm exceed better performance in comparison with other baselines, including Random and Equal-sized RU allocation algorithms. Fig. 6.8a demonstrates the algorithms performance in conditions where multiple STAs could have lower MCS indexes and it is the lower tail MCS indexes distribution. Therefore, the DRL algorithm consistently even with increasing available STAs in the network can achieves higher minimum throughput compared to the baselines. Fig. 6.8b illustrates the median performance of the algorithms. The DLR algorithm for example for 6 STAs can achieve about 28% and 65% higher minimum throughput compared to Equal-sized and Random RU allocation, respectively. As it is demonstrated the minimum throughput is decreased as the available STAs in the network decreased. This is due to the fact that with growing STAs, the variance of MCS indexes is lower between STAs because of the Popoviciu's inequality on variances [108]. Therefore, there are two main factors impacting this, one is that although the number of TXOPs per scheduling interval is as equal as the number of STAs and increasing with the STAs, with the goal of increasing the STAs' minimum throughput STAs required mostly shorter RU sizes unless there are very high variance between MCS indexes. With shorter RU sizes, this means having lower subcarriers in each RU for data transmission as more subcarriers are allocated to DC [8, 1]. In other words, while it is trying to allocate variable RU-sizes with higher STAs their MCS index variance is decreasing therefore the only factor that could have impact is the number of STAs for RU allocation, unless the MCS index is in first quartile of MCS indexes distribution. Fig. 6.8c demonstrates the 90% of CDF and it also shows that with increasing the STAs the DRL outperforms Random and Equal-sizes algorithms.

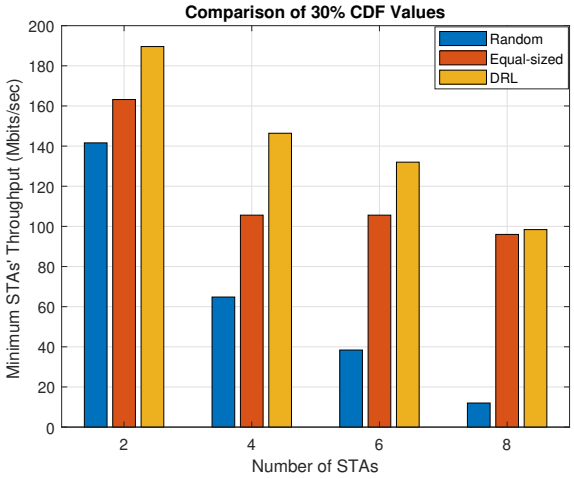


(a) CDF of STAs' minimum throughput for 100 iterations.

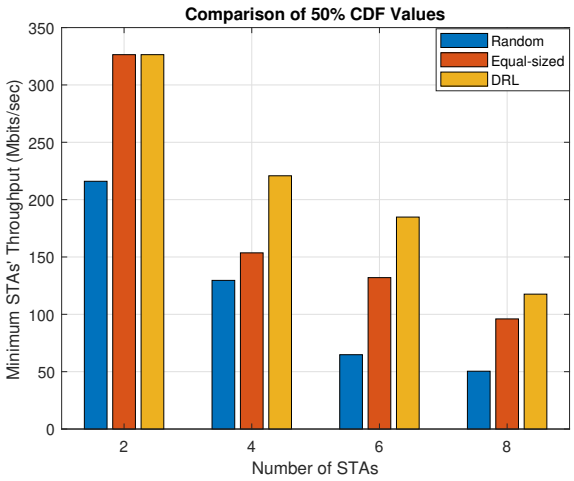


(b) CDF of Jain's index fairness for 100 iterations.

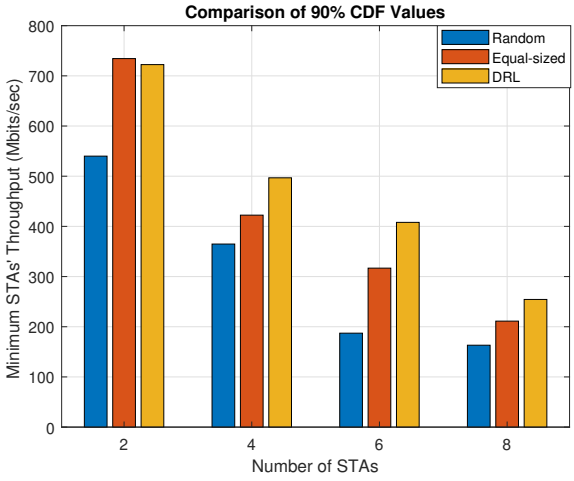
Fig. 6.7 Comparison of DRL algorithm with other benchmarks with 3 APs each with 1 STA.



(a) 30% CDF of STAs' minimum throughput versus variable number of STA.

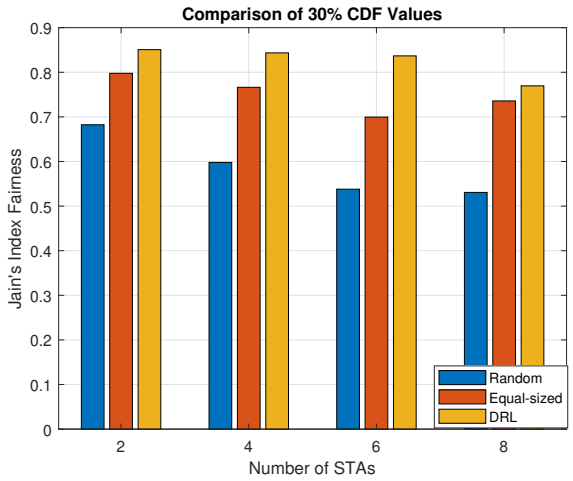


(b) 50% CDF of STAs' minimum throughput versus variable number of STA.

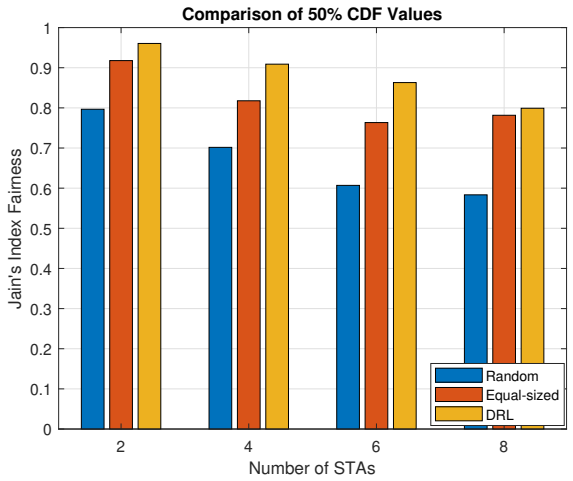


(c) 90% CDF of STAs' minimum throughput versus variable number of STA.

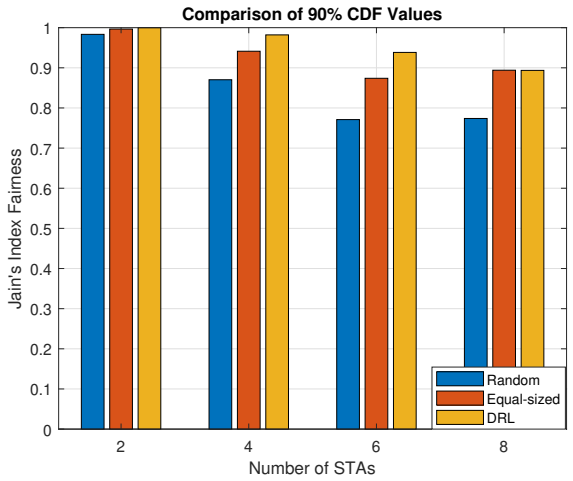
Fig. 6.8 STAs' minimum throughput versus variable number of STA.



(a) 30% CDF of Jain's index fairness versus variable number of STA.

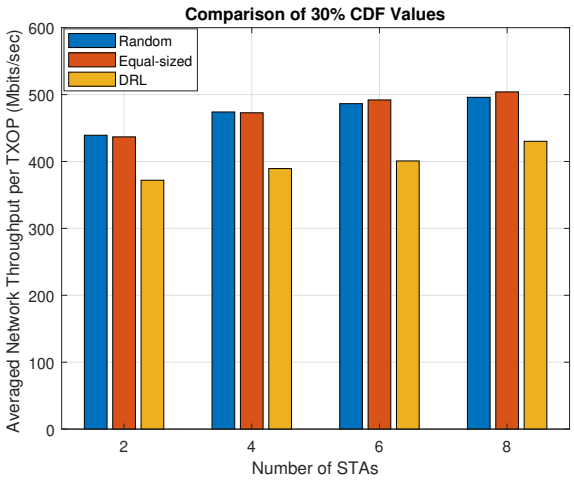


(b) 50% CDF of Jain's index fairness versus variable number of STA.

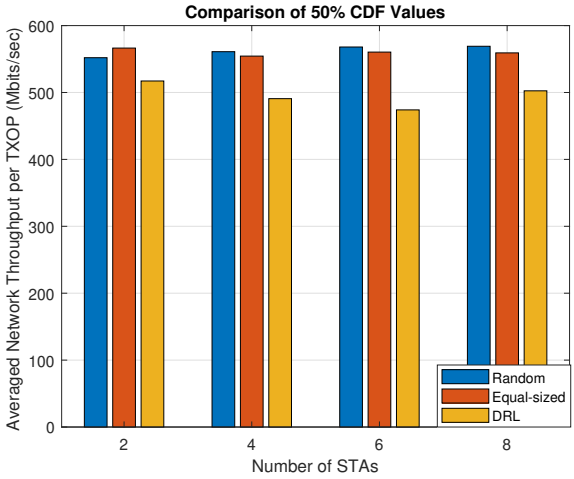


(c) 90% CDF of Jain's index fairness versus variable number of STA.

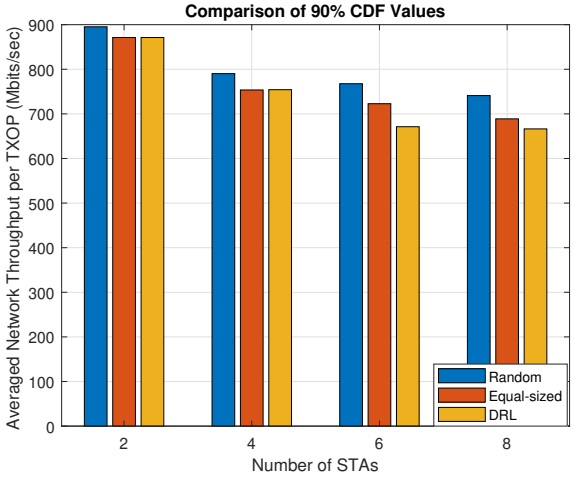
Fig. 6.9 Jain's index fairness versus variable number of STA.



(a) 30% CDF of Averaged network throughput versus variable number of STA



(b) 50% CDF of Averaged network throughput versus variable number of STA



(c) 90% CDF of Averaged network throughput versus variable number of STA

Fig. 6.10 Averaged network throughput versus variable number of STA.

Fig. 6.9 compares the 30%, 50%, and 90% CDFs of Jain's index fairness for $2-8$ STAs under Random, Equal-sized, and the proposed DRL RU allocations. This is another measurement to observe the proposed algorithm performance with respect to max-min fairness optimization. Across all available STAs and their MCS indexes, DRL algorithm achieves the highest Jain's index fairness, Equal-sized is consistently second, and Random performs worst. Fig. 6.9a represents the 30% CDF comparison where lower MCS indexes' STAs can achieve fairness higher than 0.8 once the STAs are between $0-6$ and for 8 available STAs it is about 0.75 Jain's index fairness. While the Equal-sized is between $0.7-0.8$ for different number of STAs. And Random RU allocation Jain's fairness is between $(0-0.7)$. For the median MCS indexes' distributions in Fig. 6.9b, it is illustrated that proposed DRL algorithm is steadily above all the other baselines with a Jain's index fairness range between $0.8-0.95$. In addition Fig. 6.9c shows the quartile of highest MCS indexes. All algorithms improve and the gaps narrow, while DRL fairness is higher than other baselines. That reflects its ability to exploit good channel gain without losing fairness. As the number of STAs increases, fairness generally declines for the baselines due to tighter contention, whereas DRL degrades more gracefully and remains near the top of the range.

In order to observe the network throughput performance, Fig. 6.10 demonstrates the averaged network throughput per each TXOP for variable number of STAs. It compares all the algorithms with the 30%, 50%, and 90% CDFs of averaged network throughput. As it is shown, in general the DRL algorithm with the main goal of max-min fairness has the lower network throughput compared to the Random and Equal-sized RU allocation algorithms. This is due to the fact that max-min fairness represent lower network throughput per each TXOP as it is trying to give the same throughput to all the STAs. Fig 6.10a shows the 30% CDF of averaged network throughput. Averaged network is driven because of variable TXOPs per each scheduling interval once we have variable number of STAs, therefore this is the averaged throughput of the total network. As the number of STAs increasing for 30% CDF, for all the algorithms, it can be seen that their throughput is slightly increasing. This is as a reason of in lower available STAs very low MCS indexes can cause a very low throughput. While the STAs expand, the chance that at least some STAs have decent MCS indexes enhancing, so the average network throughput in the worst 30% of cases improves. Also, with more STAs the DRL RU allocation algorithm wastes fewer resources, which lifts the lower tail. While Fig. 6.10b shows constant median for all the available STAs in the network whether with baseline

or DRL algorithm. As around the median, the benefit of extra candidates like we have in 30% CDF is roughly canceled by the shorter RU. Fig. 6.10c for 90% CDF of averaged network throughput demonstrates the slight decrease while STAs are increasing in the network. This is due to in higher MCS indexes cases with few STAs, each STA can get a larger RU. As STAs grows, RUs are shorter so the best-case average throughput falls for all algorithms.

Fig. 6.11 shows box-plots of the mean STAs' throughput achieved by the proposed DRL algorithm over 600 iterations for $2, 4, 6$, and 8 STAs. For clear demonstration the y-axis for different STAs number is enlarged. The median clearly decreases as the number of STAs grows, as explained earlier and expected. While the interquartile range (IQR) remains moderate and shrinks slightly relative to the median. Whisker asymmetry where (max > min) suggests, in a few iterations runs the network conditions are unusually good, so the achieved throughput in those runs is above the median, which can be due to either more STAs have higher MCS indexes/fewer STAs with lower MCS indexes or RU sizes that match STAs very well. These patterns are consistent with a max–min policy, where raising the slowest STA data rate reduces the overall mean and compresses the upper tail as the number of STAs increases. This shows the max-min fairness achieves its goal of giving the same throughput for all the STAs in the network.

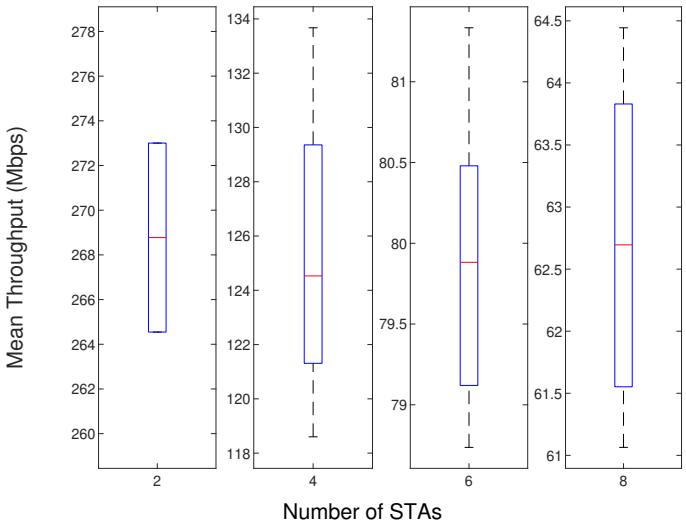
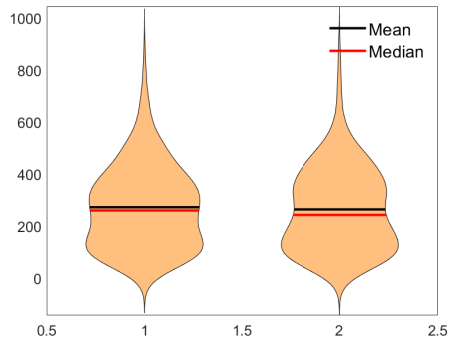
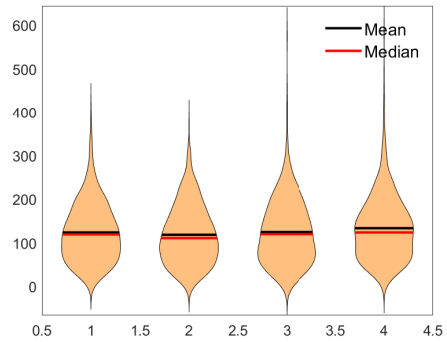


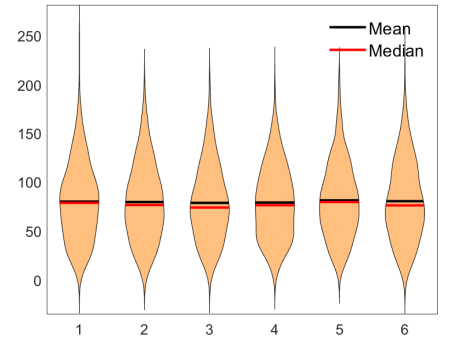
Fig. 6.11 Box plot of mean throughput for different STAs.



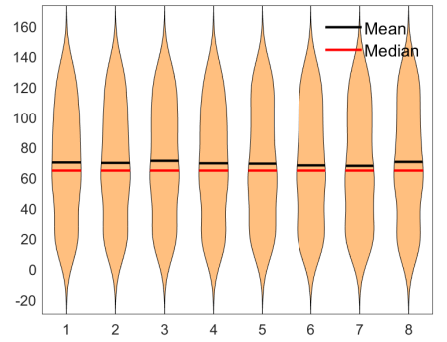
(a) Probability density of 2 STAs throughput.



(b) Probability density of 4 STAs throughput.



(c) Probability density of 6 STAs throughput.



(d) Probability density of 8 STAs throughput.

Fig. 6.12 Probability density of STAs throughput for different numbers of STAs. The x-axis in each subplot represents the STA index, while the y-axis represents the achieved throughput (Mbps). Subplots correspond to (a) 2 STAs, (b) 4 STAs, (c) 6 STAs, and (d) 8 STAs. Mean and median values are indicated by black and red lines, respectively.

Fig. 6.12 also demonstrates the STAs' throughput distribution achieved by the proposed DRL over 600 iterations for $2, 4, 6$, and 8 STAs which are shown respectively in Figs. 6.12a, 6.12b, 6.12c and 6.12d. For each STA on the x-axis, the orange violin graph shows the probability density estimation of its throughput, the red line is the median and the black line is the mean. Within each subplot, the violins for different STAs are aligned with the same means and medians. This shows that the DRL algorithm equalizes STAs throughput well despite MCS variability. Another point is that with increasing the available STAs in the network, the means and medians are degrading. This represents RUs split between more STAs. In other words the DRL RU allocation algorithm prioritizes increased the low MCS index STAs, which causes the STAs' throughput diminishes. The violins' shapes shows that mean line

is slightly higher than median one. Which means some iterations have higher MCS indexes compared to the most frequent iterations that stays near the median.

6.4 Proportional Fairness

In this section, instead of max-min fairness, proportional fairness is proposed as the main objective. Proportional fairness due to its practicality is an important goal to be considered in the system model. This is as a result that all the available STAs in the network based on their defined QoS and/or their feature could be assigned a proper resources. So not only should their requirements be quite satisfied, but also the network performance and resource usage should not be degraded. Throughout the literature review, several proportional fairness resource allocation are considered. Proportional fairness can be defined with variety of definitions, for example authors in [35] proposed a proportional fairness RU allocation based on QoS priority and buffered data of each STA. In addition to have the proportional fairness, a weighting controller based on the STAs' QoS priority and buffered data is calculated, and STA with higher weight gets the higher opportunity. However, within this scenario, it is complicated to ensure that all the STAs can achieve the minimum data rate that they might have required. In addition the effect of channel gains is not considered for RU allocation. Therefore, there is another way for defining the proportional fairness and it is based on the main definition provided in section 3.1.2. For example, authors in [94] defined the problem formulation based on the log-throughput optimization for airtime allocation. It is proved in the article that the problem formulation is a convex-optimization and a closed-form airtime allocation is driven. Concerning the related work, and with the basic observation, in this section, a closed-form proof is driven for RU allocation in a variable MCS index for the proposed system model in this chapter.

6.4.1 Problem Formulation

Therefore, considering the same system model described in 6.2, the objective is to allocate RUs based on the proportional fairness criterion defined in 3.1.2. The proposed problem formulation is as follows,

$$\underset{\rho_n^l}{\text{maximize}} \quad \sum_{n=1}^{N} \log(\sum_{l=1}^{L} \eta_n^l), \tag{6.13a}$$

$$\text{subject to} \quad 0 \leq \rho_n^l \leq \rho_{max}, \forall n \in \mathcal{N}, \tag{6.13b}$$

$$\sum_{n=1}^{N} \rho_n^l = \rho_{max}, \forall l \in \mathcal{L}, \tag{6.13c}$$

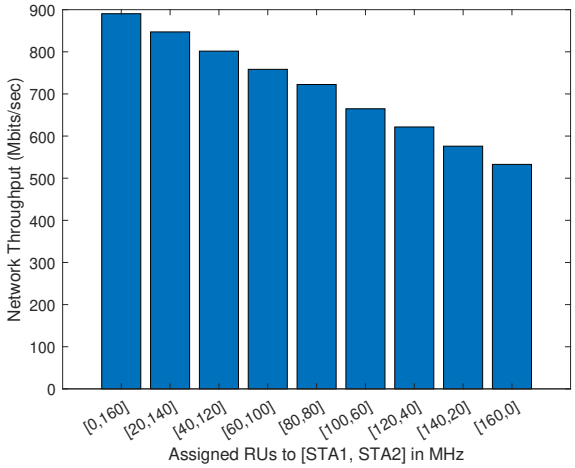
where ρ_n^l denotes the coefficient corresponding to the 20 MHz and defines the chosen RU size from $\mathcal{R}$ set for n-th STA in each TXOP per scheduling interval. L is the TXOPs per each scheduling interval. η_n^l is the calculated data rate for the n-th STA in l-th TXOP per scheduling interval. Please note that the total allocated RUs per TXOP is 160 MHz and it is also assumed that STAs are full-buffer and always have packages to transmit. Objective function (6.13a) is the proportional fairness of aggregated throughput of all STAs. Constraint (6.13b) ensures that each STA the maximum RU size that is able to achieve is ρ_{max}. While constraint (6.13c) confirm the total amount of assigned RU cannot be greater than ρ_{max}.

6.4.2 Motivation of the Proposed Problem Formulation

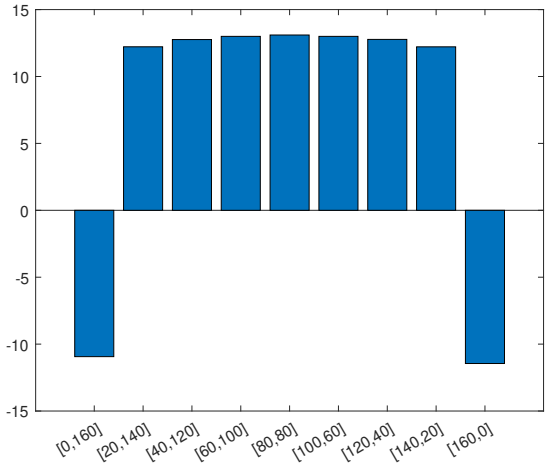
In this section, we dive into a very simple scenario to demonstrate the effect of channel gain of STAs for variable RU size allocation with proportional fairness objective optimization (6.13a). As discussed in Section 6.2.1, consider a scenario with two APs, each associated with a single STA (i.e., two STAs in the network). It is assumed that both STAs have identical packet lengths and transmit for the entire TXOP duration in each transmission opportunity. This means the STAs have airtime fairness as all of the STAs are allocated the same transmission duration. Therefore in this RU allocation it is expected that two main factors of MCS index and the available STAs in the network have effect the RU size.

However, based on the initial observation, while the STAs' MCS index is changing between $0-13$, it is observed that no matter what is the MCS index for proportional fairness RU allocation, the maximum value to have the highest proportional fairness in the network for throughput is the $[80, 80]$ RUs. Although the network throughput will be changed with different MCS indexes for each realization, still the pattern is the same, Fig. 6.13.

Therefore in the next session, we explore and proof of the reason behind this.



(a) Network Throughput versus all the possible assigned RU sets.



(b) Sum of logarithmic throughputs (proportional fairness metric) versus all possible assigned RU sets. Negative values occur when the throughput is smaller than 1 in the chosen unit, since $\log(x) < 0$ for $0 < x < 1$.

Fig. 6.13 Network performance for different RU allocation sets. (a) Total network throughput. (b) Proportional fairness metric computed as the sum of logarithmic throughputs of the STAs.

6.4.3 Closed-form Solution

In this section, initially it is proved that the η_n is linear with respect to ρ_n in the total scheduling interval, L TXOPs. All the notifications are the same as the defined ones in section 6.2. In order to derive the closed-form solution, the $\eta^l_{n_m}$ is formulated as follows,

$$\eta_{n_m}^l = \frac{PK_{n_m}^l \times \rho_{n_m}^l \times w \times \mathrm{CR}_{n_m} \times \alpha_{n_m} \times T_{Data}^l}{T_l \times SC \times T_{\mathrm{OFDM}} \times (PK_{n_m}^l + H_{\mathrm{MAC}} + H_{\mathrm{PHY}})} - \frac{PK_{n_m}^l \times L_{tail}^l}{T_l \times (PK_{n_m}^l + H_{\mathrm{MAC}} + H_{\mathrm{PHY}})}. \tag{6.14}$$

To express the linearity of $\eta_{n_m}^l$ under the assumptions of the considered system model. T_l, T_{OFDM}, $PK_{n_m}^l$, H_{MAC}, H_{PHY}, L_{tail}^l and T_{Data}^l are constant for all the STAs the network. Therefore, we can denote all constant parameters by C. The above expression reduces to a linear form in ρ_{nm}^l, CR_{nm}^l, and α_{nm}^l:

$$\nu_{n_m} = CR_{n_m} \times \alpha_{n_m}. \tag{6.15}$$

$$\eta_{n_m}^l = C \times \left(\rho_{n_m}^l \times \nu_{n_m}\right) - D. \tag{6.16}$$

It is important to note that all C and D parameters are constant and the same for any l-th TXOP and any n-th STA. In addition ν_{n_m} is related only to n-th STA and is fixed for all the l-th TXOPs in each scheduling interval. So, objective function in (6.13a) is rewritten as follows,

$$f \;=\; \sum_{n=1}^{N} \log\Big(\sum_{l=1}^{L} C \times \left(\rho_{n_m}^l \times \nu_{n_m}\right) - D\Big). \tag{6.17}$$

In the following, first it is proved that the objective function and the constraints (6.13a), (6.13b), and (6.13c) are concave, convex, convex respectively which means the problem formulation has a closed-form answer. Then it is proved the objective function is symmetric and its closed-form answer is driven.

Concavity of the objective

For fixed n, $s_n(\rho)$ is affine in $\{\rho_n^l\}_{l=1}^L$, and $\log(\cdot)$ is concave and nondecreasing on $(0,\infty)$. Hence $\rho \mapsto \log s_n(\rho)$ is concave on $\{s_n > 0\}$. Summing over n preserves concavity, so f is concave on its domain [3]. To confirm the concavity of it in another way, a direct Hessian calculation is derived. For n-th STA we have, $\nu_n = (\nu_n^1, \ldots, \nu_n^L)^\top$. Then, for each n,

$$\nabla_{\rho_n} f = \frac{\nu_n}{s_n}, \qquad \nabla_{\rho_n}^2 f = -\frac{1}{s_n^2}\,\nu_n \nu_n^\top \;\preceq\; 0, \tag{6.18}$$

a rank-1 negative semidefinite matrix. Cross-STA second derivatives vanish: $\frac{\partial^2 f}{\partial \rho_n^l \partial \rho_m^k} = 0$ when $m \neq n$. Therefore the full Hessian is block-diagonal with negative semidefinite blocks, confirming concavity.

Convexity of the feasible set

All constraints (6.13b), and (6.13c) are linear inequalities, so the feasible set is a polyhedron and thus convex. Therefore the problem is a convex optimization (maximization of a concave function over a convex set) [3].

Closed-form solution

Since $v_n^l = v_n$ (independent of l), only the *total* allocation to each STA matters. Let's define,

$$x_n \triangleq \sum_{l=1}^{L} \rho_n^l. \tag{6.19}$$

So, the objective function in (6.13a) can be simplified as follows,

$$\sum_{l=1}^{L} \eta_n^l = C v_n x_n - LD. \tag{6.20}$$

So we have,

$$\max_{x_n \geq 0} \sum_{n=1}^{N} \log\big(C v_n x_n - LD\big). \tag{6.21}$$

From the per-TXOP constraints, $\sum_{n=1}^{N} \rho_n^l = \rho_{\max}, \forall l \in L$, we obtain,

$$\sum_{n=1}^{N} x_n = L\rho_{\max}. \tag{6.22}$$

Therefore, the reduced optimization problem is:

$$\underset{x_n}{\text{maximize}} \quad \sum_{n=1}^{N} \log\left(C\nu_n x_n - LD\right) \tag{6.23a}$$

$$\text{subject to} \quad x_n \geq 0, \ \ \forall n \in \{1, \ldots, N\}, \tag{6.23b}$$

$$\sum_{n=1}^{N} x_n = L\rho_{\max}. \tag{6.23c}$$

With consideration of the the reduced problem formulation, the KKT derivation is formulated. λ is denoted as the multiplier for the equality constraint and multipliers $\mu_n \geq 0$ for the inequalities $x_n \geq 0$:

$$\mathcal{L}(x, \lambda, \mu) = \sum_{n=1}^{N} \log(C\nu_n x_n - LD) + \lambda\left(L\rho_{\max} - \sum_{n=1}^{N} x_n\right) + \sum_{n=1}^{N} \mu_n x_n. \tag{6.24}$$

As the problem is convex, KKT conditions are necessary and sufficient for optimality, so we have,

$$\begin{aligned}
\text{(Stationarity):} \quad & \frac{\partial \mathcal{L}}{\partial x_n} = \frac{C\nu_n}{C\nu_n x_n - LD} - \lambda + \mu_n = 0, \\
\text{(Primal feasibility):} \quad & \sum_{n=1}^{N} x_n = L\rho_{\max}, \qquad x_n \geq 0, \\
\text{(Dual feasibility):} \quad & \mu_n \geq 0, \\
\text{(Complementary slackness):} \quad & \mu_n x_n = 0.
\end{aligned} \tag{6.25}$$

If $x_n > 0$, then by complementary slackness $\mu_n = 0$. Thus

$$\frac{C\nu_n}{C\nu_n x_n - LD} = \lambda. \tag{6.26}$$

Therefore, with reformulating λ as follows, x_n is derived,

$$C\nu_n x_n - LD = \frac{C\nu_n}{\lambda}, \tag{6.27}$$

$$x_n = \frac{1}{\lambda} + \frac{LD}{C\nu_n}. \tag{$\star$}$$

So, with application of the equality constraint, we have

$$\sum_{n=1}^{N} x_n = L\rho_{\max}. \tag{6.28}$$

Plugging in ($\star$):

$$\sum_{n=1}^{N} \left(\frac{1}{\lambda} + \frac{LD}{C\nu_n} \right) = L\rho_{\max}. \tag{6.29}$$

This yields

$$\frac{N}{\lambda} + \sum_{n=1}^{N} \frac{LD}{C\nu_n} = L\rho_{\max}, \tag{6.30}$$

so

$$\frac{1}{\lambda} = \frac{L\rho_{\max} - \sum_{n=1}^{N} \frac{LD}{C\nu_n}}{N}. \tag{$\dagger$}$$

Therefore, the unique optimal solution is

$$x_n^\star = \frac{1}{\lambda} + \frac{LD}{C\nu_n}, \qquad \frac{1}{\lambda} = \frac{L\rho_{\max} - \sum_{n=1}^{N} \frac{LD}{C\nu_n}}{N}. \tag{6.31}$$

It should be noted that D is very small and is able to be considered approximately 0. In other words, it can be observed that once we have

$$x_n^\star = \frac{L\rho_{\max} - \sum_{n=1}^{N} \frac{LD}{C\nu_n} + N\frac{LD}{C\nu_n}}{N}, \tag{6.32}$$

Therefore

$$x_n^\star \simeq \frac{L\rho_{\max}}{N}, \tag{6.33}$$

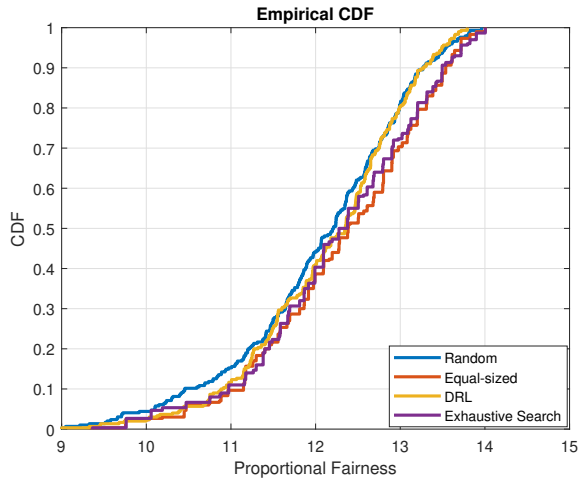
It should note D is very small that L_{tail} it can set 0. Since only the totals x_n matter, we can distribute them evenly across TXOPs:

$$\rho_n^{l\star} = \frac{x_n^\star}{L} = \frac{\rho_{\max}}{N}. \tag{6.34}$$

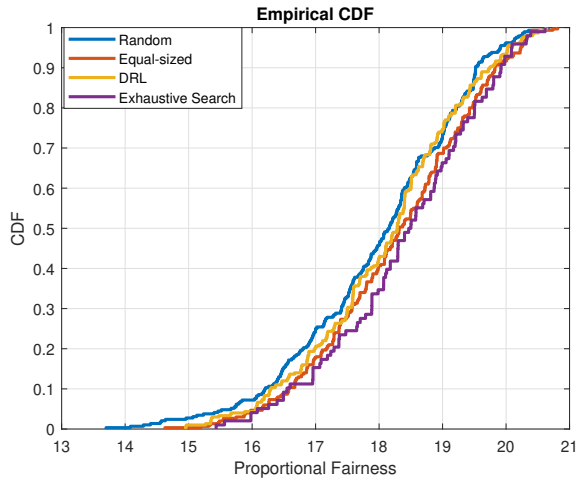
6.4.4 Simulation Results

The assumptions are the same as simulation results provided for the previous section about max-min fairness optimization. So a multi-AP coordination with a centralized controller framework is employed within a residential Wi-Fi coverage scenario. All APs are assumed to be centrally located within the BSSs. The STAs are uniformly distributed within the coverage area of each BSS, following a spatial uniform distribution around their associated APs. The APs are situated within mutual transmission range and operate on the same frequency channel. However, under C-OFDMA, RUs are allocated such that reusing RUs cannot occur within each TXOP.

First with the same assumption, for 3 APs with 1 STAs it is attempted to see the Exhaustive Search and other benchmark algorithms, Equal-sized and Random RU allocation. In addition the proposed DRL algorithm is also test for the proportional fairness where the reward function is the proportional throughput optimization proposed in (6.13a). As proved in the previous section, it is expected that with the proposed system model, all the STAs where in each TXOP all the STAs should get the equal-sized RUs, where in total scheduling interval, some of the RUs assigned to each STA should be the same. In scenarios, where the number of STAs is greater than the available RUs in the network, it is expected to allocate the shortest RU size where here is 20 MHz, to the STAs, and then assigned for the STAs not allocated any RUs in the next TXOP. As the number of TXOPs per scheduling interval, it is expected in total scheduling interval all the STAs could have the same RU-sizes allocated.



(a) CDF plot of proportional fairness of 2 STAs in the network.



(b) CDF plot of proportional fairness of 3 STAs in the network.

Fig. 6.14 Comparison of the different algorithms in the Proportional Fairness RU allocation

Fig. 6.14 demonstrates the CDF of proportional fairness, for 2 and 3 STAs. Exhaustive Search due to its high feasible search is not applicable with the current available compute nodes. As it is expected, with variable RU allocation in 300 realization, it is demonstrated that the considered Equal-sized RU allocation is able to achieve quite the same proportional fairness as the Exhaustive Search. In addition, the proposed DRL algorithm is also able to reach quite the same as the Equal-sized RU allocation. This proves the RU allocation especially as variance increased with more STAs in the network, the Equal-sized RU allocation with its low-complexity

allocation with complexity $O(1)$ can achieve the best RU allocation for proportional fairness.

Despite the performance gains achieved by the proposed coordinated multi-AP framework, several practical limitations remain. The proposed coordination mechanism assumes reliable exchange of network information among APs and centralised coordination decisions, which may introduce additional signalling overhead and latency in practical deployments. As the number of coordinated APs increases, the complexity of joint scheduling and resource assignment also grows substantially, potentially limiting scalability in dense enterprise environments. In addition, the current framework assumes synchronised operation among participating APs, whereas practical deployments may experience imperfect synchronisation and heterogeneous backhaul delays. Future improvements may therefore focus on decentralised coordination strategies, reduced signalling mechanisms, and robust scheduling methods that remain effective under imperfect coordination conditions.

6.5 Summary

This chapter investigated variable RU allocation in C-OFDMA systems with the objective of improving fairness among STAs while maintaining practical computational complexity. Two fairness objectives were studied: max-min throughput fairness and proportional throughput fairness, both formulated based on the heterogeneous channel conditions of STAs reflected by their channel gains or MCS indices.

For the max-min fairness objective, the RU allocation problem was shown to be NP-hard. To address this challenge, a DRL-based solution was proposed to learn efficient allocation policies. The obtained results demonstrated that the proposed DRL approach significantly improves fairness compared with baseline strategies such as equal-sized and random RU allocation. In particular, the learning-based approach is capable of adapting the allocation to heterogeneous channel conditions, thereby improving throughput balance among STAs without requiring exhaustive search.

For proportional fairness, the problem was analytically studied and proven to be convex under the considered system model. This enabled the derivation of a closed-form solution for optimal RU allocation when all STAs are saturated. The proposed proportional fairness solution provides an efficient allocation mechanism with constant computational complexity $O(1)$, making it highly suitable for practical implementations in real-time scheduling.

Overall, the results show that variable RU allocation based on channel conditions can significantly enhance fairness and system performance compared with fixed or random allocation schemes. Moreover, the proposed approaches maintain practical feasibility: the DRL-based solution avoids the exponential complexity of exhaustive search while still achieving near-optimal performance, and the proportional fairness formulation yields a low-complexity closed-form solution that can be directly implemented in real systems. These characteristics demonstrate that the proposed techniques not only improve fairness and throughput distribution but also maintain low algorithmic overhead, supporting their practicality for deployment in next-generation Wi-Fi systems employing C-OFDMA.

Chapter 7

Queue-Aware DRL-Based Variable RU Allocation in C-OFDMA for Latency and Reliability Optimisation

In the previous chapter, we discussed RU allocation in a full-buffer network, where it is assumed that all the STAs are saturated and always have packets to transmit. However, in a realistic scenario, this is not the case. In this chapter, we take one more step to consider the system model with practical and realistic assumptions. First, it is assumed that all STAs operate under non-saturated conditions, that is, the buffers of the APs associated with STAs are not fully occupied, and the packet arrival rate at the STAs remains variable. This means there are times when some STAs do not request any packets. In addition, the channel gain of STAs for RU allocation is considered. Throughout this chapter, the proposed system model is first introduced and motivated. Then the problem formulation is stated. As the system model is highly stochastic and RL has been promised as one of the solutions for this randomness, we propose a simple DRL algorithm to solve the considered problem. The main goal of the proposed DRL algorithm is to minimize buffer occupancy whilst aiming at minimizing packet loss and transmission latency in a stochastic environment. The proposed DRL algorithm is compared with other baseline algorithms, such as Random and Equal-sized RU allocation, and its performance advantage is illustrated and discussed with appropriate simulation results.

7.1 Introduction

Buffer queue state can help with managing the QoS of the network, especially for delay-sensitive services. Within IEEE 802.11bn, where the primary objective is to achieve ultra-high reliability with minimal latency, incorporating RU scheduling through a queue-aware scheme becomes an essential consideration [1]. The idea of queue-aware scheduling was initially introduced for network-level control of the multiple access mechanism, flow control, and routing with buffer occupancy scheduling policies [109]. For instance, authors in [110] proposed a queue-aware RU allocation algorithm within IEEE 802.11ax. The proposed algorithm, dynamic programming (DP) with the Timsort algorithm, attempted to solve the network spectral efficiency utilization and fairness with respect to packet waiting times. The authors claimed that their proposed algorithm outperforms the first in, first out (FIFO) method, especially as the number of connected stations increases. However, the authors considered a scenario with fixed MCS 11, and it should also be mentioned that the authors only considered a full-buffer scenario with fixed MCS. In [111], the authors proposed a genetic algorithm for data transmission mode selection, including OFDMA, SU-MIMO, MU-MIMO, and a combination of OFDMA and MU-MIMO. For RU allocation, the authors proposed a priority weight for each user based on their buffer length, spectral efficiency, and traffic priority factor. The allocation of RUs was performed by aligning user weights with normalized RU sizes to achieve the smallest Euclidean distance. The authors do not explicitly consider packet drops or buffer occupancy minimization, although queuing delays are included in their delay constraints. In addition, they do not address real-time decision-making mechanisms for unpredictable traffic scenarios. In [112], the authors proposed a transmission frame method selection to choose between CSMA/CA and OFDMA based on QoS of STAs in the network. While the authors did not discuss the RU scheduling based on QoS.

In this chapter, inspired by other work [111, 110, 109], we propose for the first time in IEEE 802.11bn variable real-time RU allocation based on queue-awareness. The proposed RU allocation is for decreasing the queue waiting time in the buffer, decreasing packet drop for reliability and decreasing the TXOP duration as much as possible for reducing latency. This also brings proportional fairness for the available STAs in the network. The trade-off here is how to allocate bandwidth based on the STAs' QoS. In other words, in this chapter, we propose a RU allocation algorithm with three main goals. First, it is considered to keep the buffer as empty as possible. This

helps the system with latency reduction especially queuing delay which technically it means packets spend less time waiting. Second, while the goal is to empty each buffer as much as possible then it is essential not to allow any single buffer to overflow causing a packet drop. Finally, it is also important to have the TXOP duration shorter once there are no needs for the maximum TXOP duration to transmit. The key thing is that TXOP duration defines the latency that the unscheduled stations must wait before being able to transmit and so to minimize latency, we need to keep TXOP durations as short as possible.

Therefore, here we propose an algorithm to allocate RUs based on variable channel gain and variable packet arrival rate. The primary objective is to maintain empty buffers while minimizing packet drops and latency. In addition, the proposed algorithm is a dynamic RU allocation which responds to any changes in packet arrival rate and MCS index. The simulation results demonstrate that our proposed DRL algorithm outperforms random RU allocation and the Equal-RU allocation, especially in scenarios where the total number of available RUs is less than the number of STAs in the network. In addition, the trade-off between each part of the main objective in the simulation results is discussed.

7.2 System Model and Motivation

The proposed system model considers downlink transmission with online RU allocation based on queue-awareness of STAs. The considered system model is based on a coordinated multi-AP framework. The set of APs is denoted by $\mathcal{M} = \{1, 2, \ldots, M\}$, and the set representing all available STAs in the network is defined as $\mathcal{N} = \{1, 2, \ldots, N\}$. For simplicity of analysis, it is assumed that the STAs are pre-assigned to the APs, and each AP is responsible for managing an equal number of STAs. Accordingly, the m-th AP serves $N_m = \frac{N}{M}$ STAs in the downlink transmission. The design of the transmission signaling is out of scope of this chapter. In addition it assumed that the RU scheduling and also some part of MAPC selection are happening in the master controller (MC) which has a wired connection with the APs, as illustrated in Fig. 7.1.

The main goal of this chapter is to assign variable RU sizes based on STAs buffer state and their channel gain in C-OFDMA. It is considered that the total bandwidth available within a TXOP is B, and the minimum RU size is w. Consequently, the total bandwidth can be expressed as $B = \rho_{\max} \times w$, where $\rho_{\max}$ denotes the maximum number of RUs in B. Therefore, in this chapter the total available bandwidth B is

160 MHz, and the entire bandwidth can be dynamically shared among all STAs within the same system. The minimum RU size w is set to 20 MHz. As a result, the RU allocation vector is defined as $\mathcal{R} = [0, 20, 40, 60, 80, 100, 120, 140, 160]$ MHz, as illustrated in Fig. 6.2.

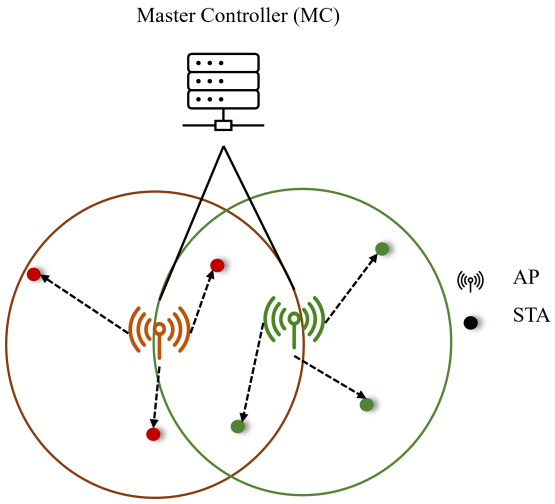


Fig. 7.1 The considered system model.

The channel gain of n-th STA with its m-th AP is $h_{n_m} = \mathrm{PL}_{n_m}(d) \cdot g_{n_m}$. The g_{n_m} is the small scale fading of n-th STA and the m-th AP. $\mathrm{PL}_{n_m}(d)$ is denoted as equation (7.1) which is the pathloss of n-th STA to the m-th AP based on the Residential cases of the TGax model [98]. We assume it as there are no internal walls, break point distance is 5 meters and with log normal fading, so it is as follows:

$$\begin{aligned} \mathrm{PL}_{n_m}(d) = {} & 40.05 + 20 \times \log_{10}(f_c/2.4) + 20 \times \log_{10}(\min(d, 5)) \\ & + (d > 5) \times 35 \times \log(d/5), \end{aligned} \tag{7.1}$$

where f_c denotes 5 GHz, and d is the distance between the transmitter and the receiver in meters. Based on each STA's channel gain and with assumption of equal maximum transmit power for each STAs, the SNR is calculated as follows,

$$\mathrm{SNR}_{n_m} = 10 \log_{10} \mathrm{P}_{\mathrm{n_m}} - \mathrm{PL}_{n_m}(d) - 10 \log_{10}\left(g_{n_m}\right) (\mathrm{dBm}). \tag{7.2}$$

Based on the minimum SNR threshold, the MCS index is chosen based on MCS Table 5.2.

Then data rate of the n_m-th STA is calculated based on its assigned RUs and its selected MCS index calculated using its channel gain. Then each TXOP throughput

can be calculated as follows,

$$\eta_{n_m} = \frac{K_{n_m}}{T_{\text{TXOP}}}, \tag{7.3}$$

where K_{n_m} is the transmitted bits of n_m-th STA in a TXOP. T_{TXOP} is the TXOP duration which once we want to calculate the aggregated packets APs can transmit is assumed fixed to the $T_{\max}$. However, after RU assignment and with considering the transmitted packets it will be updated as T_{TXOP}, which shows in the actual TXOP duration. So based on its signaling it has the following overheads and data transmission can be formulated as follows,

$$T_{\text{TXOP}} = T_{Backoff} + T_{MAP-RTS} + T_{SIFS} + T_{MAP-CTS} + T_{SIFS} + T_{MAP-TF} + T_{SIFS} + \max_{n}(T_{Data}) + T_{SIFS} + T_{ACK} + T_{DIFS}, \tag{7.4}$$

where $T_{MAP-RTS}$ and $T_{MAP-CTS}$ are the time required for multi-AP request to send and clear to send signals respectively [113]. T_{MAP-TF} is the multi-AP trigger frame to inform STAs about their allocated RUs, and other information for scheduling and transmission. To calculate the available aggregated packets the APs can transmit, initially it is assumed that $T_{\text{TXOP}} = T_{\max}$. Then based on the MCS index and allocated RUs, the total packet APs can transmit to STAs is calculated. The packet length of n-th STA denotes ζ_{n_m} and the number of aggregated packets for n-th STA is $Z^{PK}_{n_m}$, therefore we have,

$$K_{n_m} = \zeta_{n_m} \times Z^{PK}_{n_m}, \tag{7.5}$$

where K_{n_m} is the total bits the m-th AP can transmit to n-th STA. In order to calculate the $Z^{PK}_{n_m}$, the maximum number of packets that can be transmitted in a given T_{TXOP} for assigned RUs and the physical channel is calculated. For this purpose, first, the physical channel rate should be calculated as following,

$$r_{n_m} = \frac{\rho_{n_m} \times w \times \text{CR}_{n_m} \times \alpha_{n_m}}{T_{\text{OFDM}} \times SC} \quad \text{(bits/sec)}. \tag{7.6}$$

STA's MCS index, α_{n_m}, and coding rate, CR_{n_m}, are chosen based on calculated SNR_{n_m} from Table 5.2. In addition for assigned RU, $\rho_{n_m} \times w$, the subcarriers spacing is denoted as SC. Finally $Z^{PK}_{n_m}$ is calculated as follows,

$$Z^{PK}_{n_m} = \frac{r_{n_m} \times T_{Data} - \text{L}_{\text{tail}}}{PK_{n_m} + H_{\text{MAC}} + H_{\text{PHY}}}, \tag{7.7}$$

where H_{MAC} and H_{PHY} are MAC and PHY header. And L_{tail} is the number of tail bits appended at the end of the coded transmitted data bits. In the next section the main motivation of the proposed problem in the considered system model is elaborated with a simple example.

After calculating the possible aggregated packets STAs physical links are able to transmit, $Z_{n_m}^{PK}$, the APs based on the available information on the buffer occupancy, either transmit for the full TXOP or transmit whatever is in its buffer.

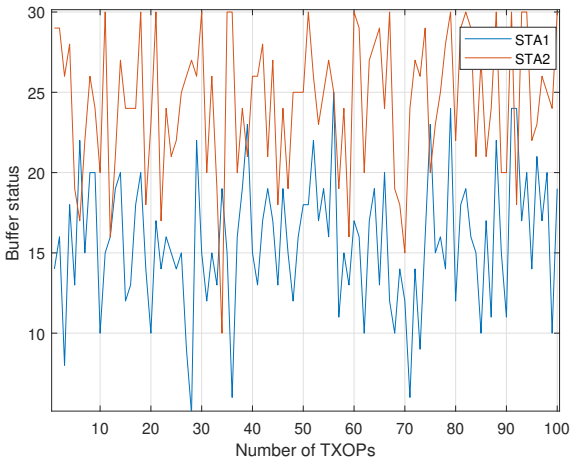
Fig. 7.2 illustrates a network scenario with a single AP and 2 STAs in the network, STA1's packet rate is described by Poisson distribution with mean $\lambda = 16$, and STA2's packet arrival rate is a Poisson's distribution with mean $\lambda = 24$. Each AP maintains a per STA queue capacity size 30 packets and with MCS index 8 for STA1 and 4 for STA2, respectively. The AP is configured for 80 MHz bandwidth with the allocated equal-sized RUs the simulation. Fig. 7.2a demonstrates the queue buffer occupancy of each STA during 100 TXOPs. As STA2 has a higher λ it is expected to fill its buffer at some points in time, resulting in dropped packets. Since STA2's physical channel is slower than STA1, it can transmit fewer packets per TXOP and is not able to empty its buffer quickly. Here the tradeoff between RU allocation, TXOP duration, buffer emptying and packet drop in the next section the considered problem is formulated and its solution described.

7.3 Problem Formulation

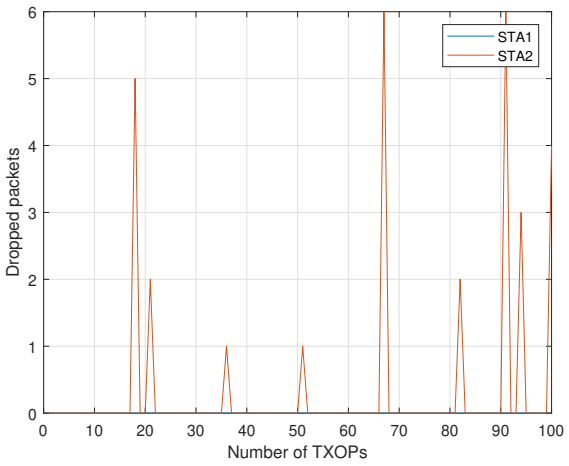
In this chapter we solve a variable RU allocation based on buffer emptying, packet drop control and TXOP duration management. c_{n_m} denotes the buffer capacity for n-th STA associated with m-th AP. Here we define the buffer occupancy as b_{n_m}. The m here is to demonstrate the n-th STAs is assigned to m-th AP. The buffer occupancy is related to n-th STA's physical data rate, r_{n_m}, its assigned RU $\rho_{n,k}^{m}$, the previous time-slot packet arrival rate, λ_{n_m}, and its buffer capacity c_{n_m}. Therefore, $b_{n_m}(t-1)$ before new arrival packets can be calculated as follows,

$$b_{n_m}(t-1) = \min(Z_{n_m}^{PK}, q_{n_m}(t-1)). \tag{7.8}$$

$b_{n_m}(t-1)$ is the minimum between the available packets in the buffer, denoted as $q_{n_m}(t-1)$, and the availability of total aggregated packets the physical channel based on allocated RUs can transmit and during transmission time can transmit. However, with new packet arrivals at t, $a_{n_m}(t)$, the $b_{n_m}(t)$ is as follows,



(a) Available packets in each STA's buffer per TXOP.



(b) Dropped packets in each STA's buffer per TXOP.

Fig. 7.2 Available buffer state and dropped packets of each STA's buffer per TXOP with assumption that STAs channel gain is fixed and the λ for STA1 and STA2 is 16, 24 respectively.

$$b_{n_m}(t) = \min(b_{n_m}(t-1) + a_{n_m}(t), c_{n_m}). \tag{7.9}$$

While the number of packets in the buffer exceeds the buffer size, packet will be lost and dropped in which is denoted as $d_{n_m}(t)$. This also related to the buffer occupancy, mean of current packet arrival rate, λ_{n_m}, physical channel and the assigned RUs.

Therefore, the variable RU allocation problem is formulated as follows:

$$\underset{\rho^{m}_{n,k}}{\text{minimize}} \quad \sum_{n=1}^{N} \Big(b_{n_m}(t-1) + d_{n_m}(t) \Big) + T_{\text{TXOP}} \tag{7.10a}$$

$$\text{subject to} \quad 0 \leq \rho_{n_m} \leq \rho_{\max}, \forall n \in \mathcal{N}, \tag{7.10b}$$

$$0 \leq T_{\text{TXOP}} \leq T_{\max}, \tag{7.10c}$$

$$\sum_{n=1}^{N} \rho^{l}_{n} = \rho_{max}, \tag{7.10d}$$

$$0 \leq b_{n_m}(t-1) \leq c_{n_m}, \forall n \in \mathcal{N}. \tag{7.10e}$$

Where the main objective in (7.10a) denotes the minimization of queue waiting time, dropped packet and TXOP duration. Constraints (7.10c) and (7.10e) makes sure the TXOP duration be within the maximum TXOP duration and the buffer queue of each STA not exceed the maximum buffer capacity of each STA, respectively. While (7.10d) ensures all RUs in the network to be used for STAs. And (7.10b) confirms that all the assigned RU should be between 0 and $\rho_{\max}$.

The formulated problem in (7.10) is NP-hard. Therefore in order to solve the problem, it is inevitable to use near optimal algorithms. DRL algorithms has been used in literature to solve NP-hard optimization problems. However their usage gives advantage to be used in wireless communication environment due to their ability for high stochastic and unpredictable environments. Therefore we propose DRL algorithm to solve our NP-hard problem where it is able to allocate RUs not only to solve the optimization problem but also to assign RUs dynamically with changes in packet arrival rates and MCS indices. In the next section the proposed DRL algorithm is elaborated.

7.4 Proposed DRL Algorithm

The variable RU allocation problem can be modeled as an MDP model. Here, S denotes the state space of the network, while A represents the corresponding action space. The reward R reflects the feedback based on the action the agent has taken. The transition function T characterizes the probability of moving from one state to another given a particular action, whereas the policy π specifies the strategy for selecting actions based on the current state so as to maximize the cumulative reward over time [76].

State Space

The goal is that the agent based on the current buffer status, current mean of the Poisson distribution λ_{n_m} and the MCS index, choose an action. The reason first to choose the current buffer status is once the agent transmit packets for each STA, the STAs buffer state will be updated and new packets will be added, therefore the agent is able to see its action effect. In addition the MCS index and λ_{n_m} are the context variables, are added to the state space to enrich the observation so the agent can make better decisions. Therefore the total state space is as follows,

$$S = \{b_1(t), \ldots, b_{n_m}(t), \lambda_1, \ldots, \lambda_{n_m}, MCS_1, \ldots, MCS_{n_m}\}. \tag{7.11}$$

Action space

The action space is the combination of all RU allocations for the available STAs in the network. In other words, an action is a vector of assigned RUs for each STAs. The action space is all of the combination of RU allocation and is as following,

$$A = \left\{\rho_1, \ldots, \rho_{n_m}\right\}. \tag{7.12}$$

Where ρ_{n_m} needs to satisfy the constraints in (7.10b) and (7.10d). Therefore the total action space size is $\binom{\rho_{\max}+N-1}{N-1}$ of vector $(1 \times N)$.

Reward

To define the reward function, it is essential to propose a function that the agent can see its action impact and optimizing the objective function. The objective function in (7.10a) proposes to minimize the packet queue in the buffer, packet drop and the TXOP duration. To convert this in a reward function and as the main goal is to keep the buffer as empty as possible with lower packet drop, $\frac{1}{(b_{n_m}(t-1)+d_{n_m}(t))+1}$ is proposed to make sure the reward decreases as the backlogs and packet drop grows. In addition the normalized delay component allows the agent to balance queue minimization with maintaining low latency, preventing it from focusing on only one aspect. Therefore the proposed reward function is as following,

$$R = \frac{1}{(b_{n_m}(t-1) + d_{n_m}(t)) + 1} - \frac{T}{T_{\max}}. \tag{7.13}$$

Finally the proposed DRL RU allocation algorithm is elaborated in Algorithm 6. In the environment, in each episode, a random mean Poisson distribution, λ_{n_m}, and

Algorithm 6: The proposed DRL-based RU allocation algorithm

Initialize Model: Initialize replay memory and Q-network
Packet arrival rate λ_{n_m} and MCS indices are chosen randomly
for $Episode = 1, 2, \ldots$ **do**
    Initialized the environment state S^0
    for $t = 1, 2, \ldots, \tau$ **do**
        DQN agent chooses action a^t from A set using ϵ-greedy policy.
        Execute action a^t
        Receive reward r^t and next state $S^{(t+1)}$
        Store experience (S^t, a^t, r^t, S^{t+1})
        Sample mini-batch from memory and update Q-network

MCS indices for STAs are generated. Once the agent takes an action, the selected action is applied in the environment and the buffer is updated based on transmitted packets and TXOP duration. The reward is calculated and then based on the new arrival packets, λ_{n_m}, the next state is updated.

7.5 Simulation Results

The simulation results are presented for a C-OFDMA based Wi-Fi network where a centralized controller framework is employed within a residential Wi-Fi coverage scenario. APs are assumed to have isotropic coverage. The STAs are positioned in their BSSs with uniform distribution, $\mathcal{U}(0,1)$. The APs are situated within mutual transmission range and operate on the same frequency channel. However, we assumed that reusing RUs cannot occur within each TXOP. The simulation results are generated using Matlab. The DQN network architecture consists of an input layer followed by a fully connected layer, and a ReLU activation layer, ReLU, which introduces non-linearity. This is followed by another fully connected layer, which further processes the features. The output from this layer is passed to the q-values and the hidden layers contain 256 nodes.

The performance of the proposed DRL algorithm is compared with Equal-sized and Random RU allocation benchmarks. The Equal-sized RU allocation is a baseline algorithm where in each TXOP the STAs get the same RU size, however with the 3 STAs it is different. For example with 3 STAs in the network, two of the STAs are assigned $40 + 20$ MHz and one of the STA's assigned RU is 40 MHz bandwidth, which are allocated randomly where the total bandwidth is 160 MHz. So either the total bandwidth is divided with equal-sized RU if it is applicable. Or in case the

number of STAs for scheduling is greater than the maximum number of RUs (i.e. bandwidth/minimum RU size), the minimum sized RUs are assigned randomly to STAs for the Equal-sized RU allocation benchmark. The random RU allocation assigned random variable-sized RUs to random STAs in each TXOP.

The computational complexity of the proposed solution approach approximately is $O\left(\binom{\rho_{\max}+N-1}{N-1}\right)$. This computational complexity is increasing with the number of STAs will be challenging in a real-time scenario. Conversely the DQN computational complexity per time step is related to the cost of a single forward pass through the network $O\left(f(d)\right)$ [81], and especially once it is trained it can give assigned RU simultaneously. However, the proposed DQN has its own complexity once STAs growing which is an open problem to be solved for our future work.

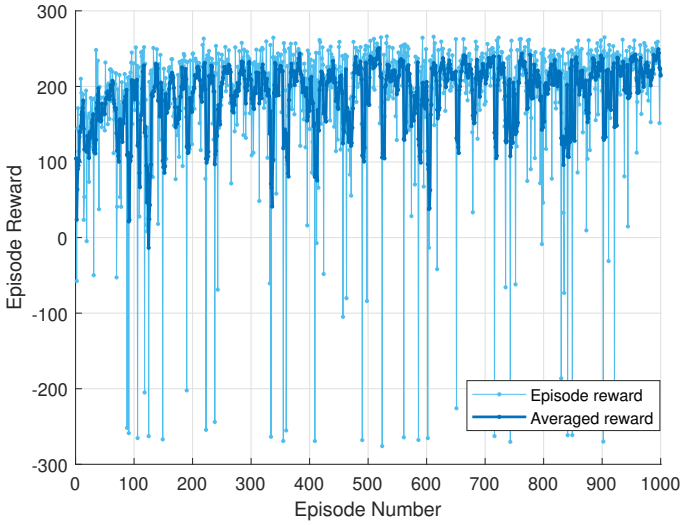


Fig. 7.3 Learning Curve for 2 STAs in the network.

Fig. 7.3 demonstrates the learning curve of scenario where there are 2 APs and each of them has one STA assigned, which means the total STAs in the network are 2. As it is illustrated in the initial episodes $(0-50)$, the episode rewards vary widely, including sharp drops into negative values below -200. This indicates the agent is still exploring and making suboptimal RU assignment decisions, leading to poor network performance. As training progresses, the reward stabilizes around a higher positive value, around $150-200$. The average reward curve shows a steady upward trend, confirming that the RL model is converging toward an optimal policy. The drops are also due to sudden change of mean packet arrival rate and MCS index in each step.

To observe the agent performance, initially in Fig. 7.4 examples of the distribution of the summation of buffer occupancy within 200 TXOPs per a realization with 200 realizations are demonstrated. This is important to demonstrate the initial objective of keeping the buffer as empty as possible. As illustrated the proposed DRL algorithm is compared with the Random and Equal-sized RU allocation. Fig. 7.4a compares the proposed DRL algorithm with other baselines where there are 2 STAs. As can be seen, the proposed DRL algorithm can allocate RUs slightly better than Equal-sized RU allocation (which allocates the same 80 MHz to both STAs). In addition the improvement compared to Random RU allocation is higher, which is demonstrated that with Random RU allocation the buffer is successfully emptied in 75% of realizations while for the DRL algorithm this is increased to 98%.

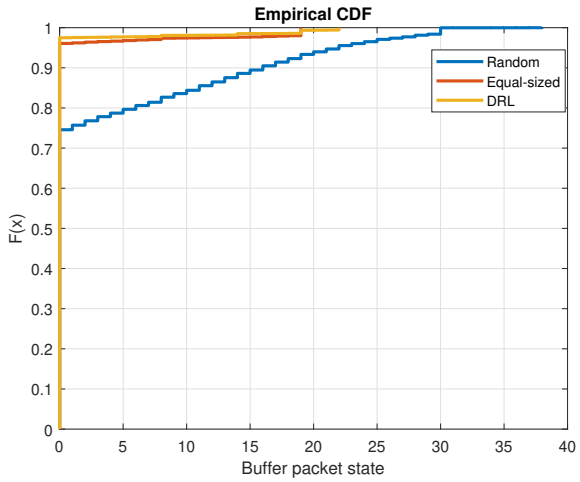
Fig. 7.4b compares the proposed DRL algorithm with the Random and Equal-sized RU allocation for total 3 STAs in the network. Where it is assumed that 3 APs are assigned only one STA. For Equal-sized RU allocation algorithm in order to make the comparison with other algorithms fair, the RU assignment is based on using the total 160 MHz where 2 of STAs randomly assigned $40+20$ MHz and one of them is assigned 40 MHz which give us the total bandwidth to use. However, another baseline named Equal-sized allocation is compared with the DRL algorithm. This is due to demonstrate the effect of scenarios where no RUs are available and without any scheduling allocating equal sized RUs for some of the STAs, how buffer state will be. Therefore, the DRL algorithm is compared with the RU assignment where 2 STAs randomly assigned 80 MHz and also one of them is not assigned any RUs. In addition the proposed DRL algorithm is slightly better than the Equal-sized RU allocation. While with comparison with Random RU allocation in about 93% of DRL algorithm realizations its buffer is empty while for Random RU allocation this number is 45%. Fig. 7.4c compared the DRL algorithm with the baseline algorithms in a scenario where there are 2 APs and each of them allocated 2 STAs. As it is demonstrated the proposed DRL algorithm is as good as the Equal-sized RU allocation for keeping the buffer empty. And in this case the DRL also outperform the Random RU allocation as before.

The second element of main objective optimization is to minimize the number of packet drop. Therefore, in Fig. 7.5 total packet drops between 200 realizations is demonstrated. The CDF plot of the proposed DRL algorithm with other baselines is compared. As demonstrated in Fig. 7.5a, the proposed DRL algorithm for two STAs manages not to have any packets drop in 95% of realizations. While with comparison with Random RU allocation this is true for only 87% of realizations. The number of

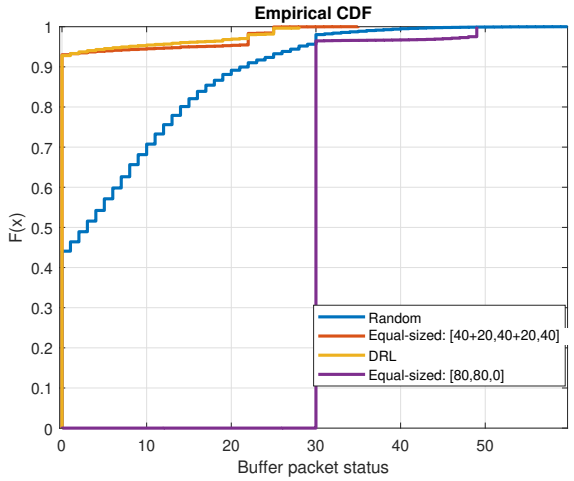
TXOPs for each realization is assumed 200 TXOPs. In addition the Equal-sized for two STAs is giving the packet drop rate the same as the DRL algorithm. This is the same for 4 STAs as illustrated in Fig.7.5c. However, in Fig. 7.5b the proposed DRL algorithm in comparison of the Equal-sized RU allocation algorithm for packet drop is slightly better. In addition in Fig.7.5b, it is shown that if the RUs are assigned with a random permutation of vector $[80, 80, 0]$ like in Equal-sized allocation algorithm the packet drop is drastically higher which demonstrates the importance of RU scheduling with changes especially in cases where the number of RUs is smaller than the number of STAs in the network.

In Fig. 7.6 the impact of the other parameters is presented by the proposed problem formulation and the DRL-based RU allocation. In Fig.7.6a the actual TXOP duration for all TXOPs of all realizations is demonstrated. The proposed DRL algorithm TXOP duration is quite the same as the Equal-sized RU allocation. This is reasonable due to the first initial goal is to keep buffer as empty as possible while maintain the reliability with packet drop decreases and then it is tried to manage the TXOP duration. While Random RU allocation for 40% of realizations is higher than DRL algorithm. On the other hand the Equal-sized allocation with 80 MHz bandwidth has shorter TXOP duration although its packet drop rate was not good at all. In Fig. 7.6b, the algorithms are compared with respect to total transmitted packet throughout each TXOP and as its demonstrated the DRL algorithm could transmit quite the same as the Random and Equal-sized RU allocation. This is due to the fact that the main objective is delay and reliability management and it is reasonable not much differences with respect to transmitted packets in the network compared to other baselines.

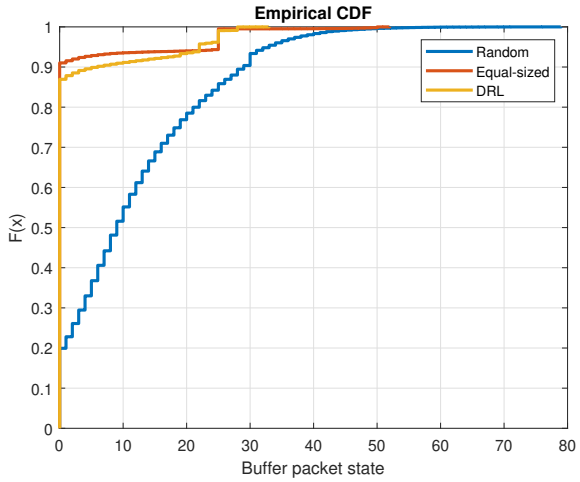
Finally Fig. 7.6c, illustrates the spectral efficiency. The main goal is to demonstrate the spectral efficiency we could achieve with the transmitted packet and the managed control while the main goal was optimizing. As it is shown , the proposed DRL algorithm could gain the same spectral efficiency as the Equal-sized allocation with 80 MHz bandwidth allocation. In addition for 50% CDF of spectral efficiency, it is about 10% better than Random RU allocation while 10% lower than the Equal-sized RU allocation.



(a) 2 STAs buffer status.

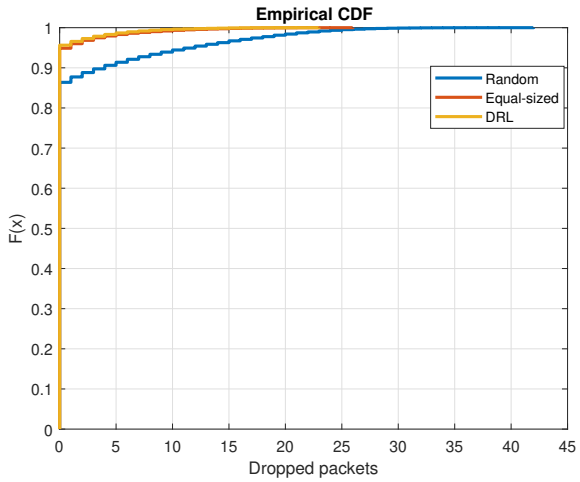


(b) 3 STAs buffer status.

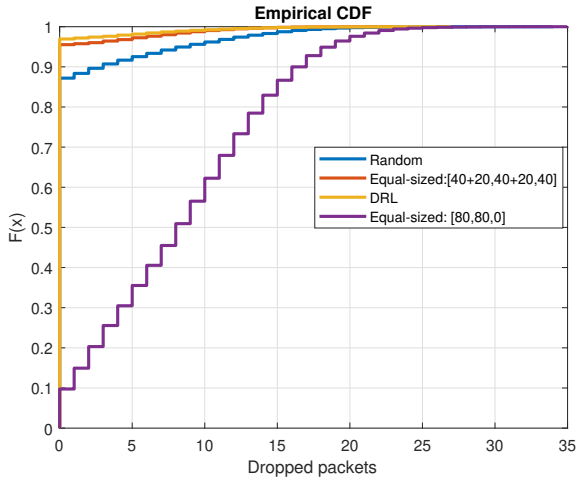


(c) 4 STAs buffer status.

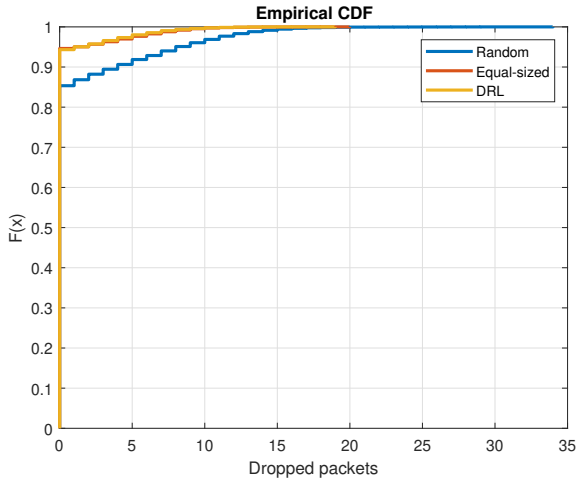
Fig. 7.4 Comparison of buffer status for different number of STAs.



(a) 2 STAs packets drop.

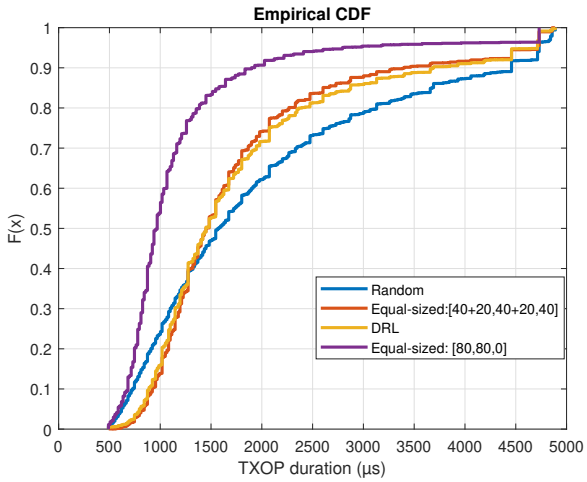


(b) 3 STAs packets drop.

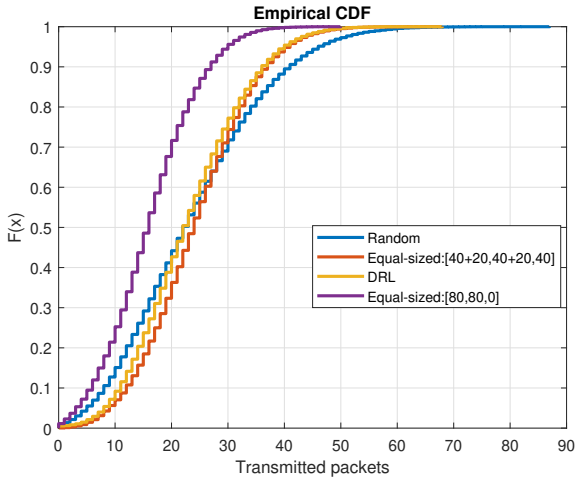


(c) 4 STAs packets drop.

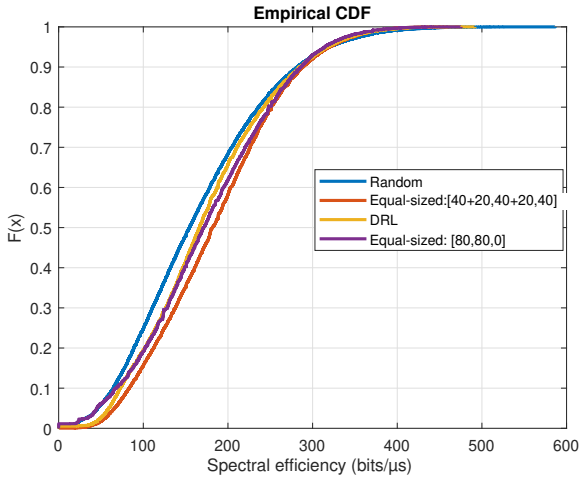
Fig. 7.5 Comparison of packets drop for different number of STAs.



(a) 3 STAs TXOP duration.



(b) 3 STAs transmitted packets.



(c) 3 STAs spectral efficiency.

Fig. 7.6 Comparison of TXOP duration and spectral efficiency for 3 STAs in the network.

Although the proposed learning-based framework offers adaptability under dynamic network conditions, several limitations should be recognised. Training complexity remains a key challenge, particularly when the state space expands with increasing numbers of users, APs, and scheduling parameters. In addition, learning-based decisions may require substantial training time before convergence, which can affect deployment in rapidly changing practical environments. The current implementation also assumes that sufficient system observations are available to support stable policy learning, while practical deployments may experience partial observability or delayed feedback. Furthermore, transferring trained policies across different deployment scenarios may require retraining or adaptation. Future research may therefore investigate lightweight online learning methods, transfer learning approaches, and hybrid schemes that combine optimisation-based scheduling with learning-assisted adaptation.

7.6 Summary

This chapter investigated real-time variable RU allocation in C-OFDMA systems under realistic non-saturated traffic conditions, where packet arrivals are stochastic, and STAs may not always have data to transmit. To address these conditions, the RU allocation problem was formulated with the objective of maximizing buffer emptying while simultaneously reducing queueing delay and packet drops. Due to the dynamic and stochastic nature of the system, a DRL based scheduling approach was proposed to learn efficient RU allocation policies based on both channel conditions and queue states.

The simulation results demonstrate that the proposed DRL-based approach significantly outperforms conventional baseline schemes such as random and equal-sized RU allocation. In particular, the proposed method achieves faster buffer emptying and more efficient utilization of available RUs by allocating resources according to both the instantaneous channel quality and the queue occupancy of STAs. As a result, the proposed approach substantially reduces packet drops and queueing delay compared with baseline allocation strategies that do not consider queue awareness or traffic dynamics.

Furthermore, the results show that incorporating queue information into the scheduling process allows the system to better adapt to fluctuating traffic conditions and heterogeneous channel environments. While baseline schemes tend to allocate resources in a static or uninformed manner, the learning-based approach dynamically

prioritizes STAs with urgent transmission needs, leading to improved overall network responsiveness and service reliability.

Overall, the proposed DRL-based RU allocation framework provides a practical and effective solution for real-time scheduling in next-generation Wi-Fi systems. By jointly considering queue states and channel conditions, the proposed approach improves buffer management, reduces latency and packet loss, and achieves better resource utilization compared with existing heuristic allocation methods. These results demonstrate the potential of learning-based scheduling mechanisms for supporting delay-sensitive and burst traffic in coordinated OFDMA Wi-Fi networks.

Chapter 8

Conclusions, Open Issues, and Future Works

8.1 Conclusions

The aim of this thesis was to propose practical resource allocation techniques for Wi-Fi systems operating under heterogeneous QoS requirements. The work began by analyzing the system parameters that influence resource allocation optimization, with the objective of improving network performance in terms of throughput, fairness, latency, and reliability. More specifically,

1. Chapter 1 introduced Wi-Fi network, its background, and the emerging technologies envisioned for next generations. The chapter highlighted the motivation for this study and outlined the specific contributions of each chapter.

2. Chapter 2 began with a comparison between Wi-Fi and cellular networks, emphasizing their general yet fundamental technical differences. The evolution of Wi-Fi standards was then reviewed up to the current and upcoming generations, namely Wi-Fi 7 and beyond. The revolutionary technologies and enhancements introduced at both the PHY and MAC layers were discussed, with multi-AP coordination identified as the central focus of this study. Multi-AP coordination was considered due to its potential to improve resource utilization efficiency, reduce latency, manage interference, and increase reliability especially in over crowded environments. Nevertheless, challenges related to its practical implementation, including transmission signaling, hardware limitations, overhead, and resource allocation, remain open research problems.

Therefore, the main focus of the thesis was on C-OFDMA, as a subcategory of multi-AP coordination, due to its benefits in improving spectrum utilization for different objectives despite its implementation complexity. In parallel, the channel access protocols of Wi-Fi were examined, particularly in light of IEEE 802.11bn's emphasis on reducing latency and improving reliability. The chapter also provided a technical overview of the conventional CSMA/CA channel access protocol and the signaling procedures of OFDMA/C-OFDMA, as these fundamental mechanisms are essential for understanding Wi-Fi operation. Finally, the related literature review was presented, highlighting the open problems associated with channel access protocol and C-OFDMA RU scheduling.

3. Chapter 3 presented the fundamental mathematical background required for this study, covering fairness metrics, convex optimization principles, and ML techniques. The chapter began by introducing different fairness metrics that were later employed in the problem formulations. Subsequently, the principles of convex optimization were discussed as a foundation for solving optimization problems. Finally, an overview of ML algorithms was provided, including the fundamentals of supervised learning, unsupervised learning, reinforcement learning, and federated learning.

4. Chapter 4 conducted the system implementation of the conventional CSMA/CA channel access protocol in Wi-Fi networks. Furthermore, CSMA/CA was compared with the Fair-MARL algorithm [85] under variable data rate scenarios. Finally, CSMA/CA was analyzed within a heterogeneous system model in which STAs were assigned different access categories. In this context, an open problem were proposed based on RL algorithms to achieve proportional fairness within the EDCA protocol.

5. Chapter 5 proposed an equal-sized RU allocation scheme within a joint C-OFDMA and C-SR system model. A semi-distributed multi-AP coordination framework was considered, in which the AP with the smallest TXOP utilized the full channel. The sharing AP then allocated RUs to overlapping STAs based on a WMM-based graph-coloring approach, while each AP independently assigned RUs to its non-overlapping STAs based on WMM. The objective of this design was to maximize overall network throughput while ensuring that the minimum data rate requirements of all STAs were satisfied. Numerical results demonstrated that the proposed RU allocation algorithm outperformed

baseline approaches in terms of total network throughput while maintaining lower computational complexity.

6. Chapter 6 proposed a centralized C-OFDMA multi-AP system in which variable RUs were allocated under variable channel gains, guided by two distinct throughput fairness objectives. In the first case, a DRL-based variable RU allocation algorithm was introduced to achieve max-min throughput fairness among STAs. To account for practical considerations, fairness was enforced over a scheduling interval. Since the formulated optimization problem was NP-hard, a novel DRL algorithm was proposed. Numerical results demonstrated the effectiveness of the approach, showing improved minimum throughput satisfaction and higher Jain's fairness index compared to baseline schemes. In the second part of the chapter, the same system model was extended to address proportional throughput fairness. It was shown that the proposed practical problem formulation was convex, allowing a closed-form solution to be derived. The analysis revealed that, in crowded environments, allocating equal-sized RUs is the optimal strategy to maintain proportional throughput fairness.

7. Chapter 7 introduced more practical assumptions for RU allocation. Motivated by the primary goal of UHR in IEEE 802.11bn, a system model was proposed in which RU scheduling was designed to be queue-aware. The objective was to minimize the waiting time of packets in the queue, packet drops, and TXOP duration jointly. In other words, based on the QoS requirements of STAs, including packet arrival rates and channel gains, the RU allocation aimed to keep low buffer occupancy while mitigating packet losses and reducing transmission overhead. The proposed problem formulation was identified as NP-hard. Furthermore, to support real-time RU allocation without requiring re-optimization after initial training, a DRL-based algorithm was introduced. This approach enabled practical and adaptive RU assignment while addressing the complexity of the problem.

8.2 Summary of Contributions

This thesis investigates resource allocation mechanisms for coordinated OFDMA-based Wi-Fi networks with the objective of improving fairness, resource utilization, and delay performance under both saturated and non-saturated traffic conditions. The main research contributions of this thesis are summarised as follows.

1. System modelling and problem formulation for coordinated OFDMA Wi-Fi networks (Chapter 3). A system-level framework for C-OFDMA Wi-Fi networks was developed, capturing the interactions between multiple BSSs, heterogeneous channel conditions, and RU allocation. Within this framework, fairness-aware resource allocation problems were formulated to address the challenges of heterogeneous STA channel quality and limited spectrum resources. The modelling provides the analytical foundation for the optimisation and learning-based solutions proposed in later chapters.

2. Max–min fairness-based RU allocation using deep reinforcement learning (Chapter 6). The RU allocation problem under the objective of max–min throughput fairness was formulated and shown to be NP-hard due to the combinatorial nature of variable RU assignments. To address this challenge, a DRL-based scheduler was proposed to learn efficient RU allocation policies from the system state. The proposed approach adapts RU assignments according to STA channel conditions and achieves significantly improved fairness compared with baseline allocation strategies, such as equal-sized or random RU allocation, while avoiding the exponential complexity of exhaustive search.

3. Analytical formulation of proportional fairness RU allocation (Chapter 6). In addition to the max–min fairness objective, the RU allocation problem was analysed under a proportional fairness objective. The problem was shown to be convex under the considered system model, enabling the derivation of a closed-form optimal RU allocation solution when all STAs are saturated. This contribution provides a computationally efficient scheduling mechanism with constant complexity, demonstrating the feasibility of fairness-aware RU allocation in practical Wi-Fi scheduling systems.

4. Queue-aware DRL-based RU allocation for non-saturated traffic (Chapter 7). To address realistic traffic conditions where packet arrivals are stochastic, a queue-aware DRL-based RU allocation framework was proposed. Unlike

conventional approaches that consider only channel conditions, the proposed method jointly incorporates queue states and channel quality information into the scheduling decision process. The learning-based scheduler enables adaptive prioritisation of STAs with urgent transmission demands, leading to improved buffer emptying, reduced queueing delay, and lower packet dropping probability.

5. Performance evaluation and comparison with baseline allocation schemes (Chapters 6 and 7). Extensive simulations were conducted to evaluate the performance of the proposed algorithms under heterogeneous channel conditions and stochastic traffic arrivals. The results demonstrate that the proposed approaches consistently outperform conventional RU allocation schemes such as equal-sized allocation and random allocation. In particular, the proposed methods improve fairness among STAs, increase resource utilisation efficiency, and reduce queueing delay and packet losses in dynamic network conditions.

Overall, the contributions of this thesis demonstrate that intelligent and fairness-aware RU allocation mechanisms can significantly enhance the performance of coordinated OFDMA Wi-Fi networks while maintaining practical computational complexity. The proposed analytical and learning-based solutions provide a foundation for efficient scheduling mechanisms in future high-efficiency Wi-Fi systems.

In conclusion, this thesis has addressed new practical challenges in RU allocation with the objectives of maximizing throughput, ensuring fairness, minimizing latency, and enhancing reliability, while maintaining the QoS of STAs across diverse practical scenarios. The findings further demonstrate the significant impact of RU scheduling on C-OFDMA, particularly in real-time scenarios, which warrant further investigation. The next section discusses potential future research directions related to the work presented in this thesis.

8.3 Future Works

8.3.1 Channel Access with Fair-MARL in HetNet Environment

As discussed in Chapter 4, the Fair-MARL algorithm was evaluated under a single-AP scenario with homogeneous and variable data rate conditions. However, three key limitations were identified. First, the algorithm was not tested in coexistence with

legacy STAs operating under conventional CSMA/CA, which is a necessary condition for practical deployment, since aggressive channel access by either RL agents or legacy STAs would compromise system fairness. Second, the weight-transfer mechanism in Fair-MARL introduces signalling overhead whose impact under realistic conditions remains unquantified; if excessive, alternative DRL formulations with revised reward functions should be explored. Third, and most importantly, the current Fair-MARL algorithm treats all STAs equally regardless of their access category, making it incompatible with the QoS differentiation framework of EDCA. A natural extension is therefore to design an RL-based channel access protocol that allocates channel access proportionally in accordance with EDCA access categories, enabling priority-aware scheduling within the multi-agent framework.

8.3.2 New approaches for joint C-OFDMA and C-SR RU allocation

The algorithm proposed in Chapter 5 always prioritises RU allocation to overlapping STAs first, after which remaining APs independently assign RUs to their non-overlapping STAs. As noted in the chapter discussion, this sequential allocation strategy does not guarantee global optimality, particularly in scenarios with unbalanced traffic loads or asymmetric OBSS topologies. A more principled extension would incorporate a priority factor into the allocation, for example, a proportional fairness weight per STA, to achieve a globally more equitable and efficient RU assignment. Furthermore, the feasibility of the sharing AP obtaining the required buffer state and channel information from neighbouring APs in real time should be validated, and a DRL-based solution for this joint problem warrants further investigation.

8.3.3 Multi-agent RL for max-min fairness RU allocation based on channel gain

In Chapter 6, the max-min fairness problem was formulated with transmit power assumed constant and equal across all STAs. As discussed, this is a practical simplification, but since transmit power directly influences channel gain and therefore RU efficiency, fixing it limits the optimality of the allocation. A meaningful extension is to jointly optimise RU allocation and transmit power, which would require revisiting the problem formulation, which likely renders it non-convex, and motivating more advanced multi-agent RL approaches capable of handling the enlarged joint action space. Additionally, the number of TXOPs per scheduling interval was treated as a

fixed constant in the system model; future work could treat this as an optimisation variable, tailored to the QoS requirements and channel conditions of individual STAs.

8.3.4 Online RU allocation with Actor-Critic RL algorithms

The DRL algorithm proposed in Chapter 7 suffers from the curse of dimensionality: the action space grows combinatorially with the number of STAs, which limits scalability to dense deployments. As identified in the chapter, replacing the standard DQN with more advanced actor-critic methods such as PPO, which can handle high-dimensional and continuous action spaces, is a principal direction for future work. Furthermore, the proposed solution targets downlink transmission only. In the uplink, per-TXOP RU scheduling incurs prohibitive signalling overhead, meaning that RU allocation must be performed over an entire scheduling interval, a problem structure that exceeds the practical capabilities of the current DQN formulation and requires the development of new algorithms capable of efficiently solving uplink RU allocation at scale.

8.3.5 Multi-Antenna Transmission and Enhanced Channel-Aware Scheduling

An important direction for future research is the extension of the proposed coordinated OFDMA framework to frequency-selective channel environments, where resource unit allocation may adapt to sub-band channel variations across users. Although heterogeneous and time-varying channel gains are already considered in this thesis, finer frequency-selective scheduling was not included because the additional channel feedback and signalling overhead required in practical Wi-Fi systems may offset its potential performance gains. Investigating this trade-off under realistic implementation constraints remains an open research problem. In addition, the present work assumes single spatial stream transmission between each AP and STA. Extending the optimisation framework to multi-antenna coordinated transmission, including beamforming-aware scheduling and joint spatial multiplexing, would provide an important next step toward more complete multi-AP coordination in future Wi-Fi systems.

8.4 Summary

In this chapter, the main contributions of the thesis were outlined. In addition, open problems that remain unresolved were highlighted, and possible directions for future work were discussed. This emphasizes both the significance of the work presented and the opportunities it creates for further advancement in the field.

References

[1] Giovanni Geraci, Francesca Meneghello, Francesc Wilhelmi, David Lopez-Perez, Iñaki Val, Lorenzo Galati Giordano, Carlos Cordeiro, Monisha Ghosh, Edward Knightly, and Boris Bellalta. Wi-Fi: Twenty-five years and counting. arXiv preprint arXiv:2507.09613, 2025.

[2] Shikhar Verma, Tiago Koketsu Rodrigues, Yuichi Kawamoto, Mostafa M. Fouda, and Nei Kato. A survey on Multi-AP coordination approaches over emerging WLANs: Future directions and open challenges. *IEEE Communications Surveys & Tutorials*, 26(2):858–889, 2024.

[3] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[4] Biljana Bojovic. *Cellular and Wi-Fi technologies evolution: From complementarity to competition*. PhD thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, 2022.

[5] Ramia Babiker Mohammed Abdelrahman, Amin Babiker A Mustafa, and Ashraf A Osman. A comparison between IEEE 802.11 a, b, g, n and ac standards. *IOSR Journal of Computer Engineering (IOSR-JEC)*, 17(5):26–29, 2015.

[6] Hassan Aboubakr Omar, Khadige Abboud, Nan Cheng, Kamal Rahimi Malekshan, Amila Tharaperiya Gamage, and Weihua Zhuang. A survey on high efficiency wireless local area networks: Next generation WiFi. *IEEE Communications Surveys & Tutorials*, 18(4):2315–2344, 2016.

[7] Anton Vikulov and Alexander Paramonov. Practical retrospective of 5-year evolution of the IEEE 802.11 client device capabilities. In *2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pages 296–300. IEEE, 2020.

[8] Cailian Deng, Xuming Fang, Xiao Han, Xianbin Wang, Li Yan, Rong He, Yan Long, and Yuchen Guo. IEEE 802.11 be Wi-Fi 7: New challenges and opportunities. *IEEE Communications Surveys & Tutorials*, 22(4):2136–2166, 2020.

[9] Dr V Vityanathan Rathnakar Acharya and Pethur Raj Chellaih. WLAN QoS issues and IEEE 802.11 e QoS enhancement. *International Journal of Computer Theory and Engineering*, 2(1):1793–8201, 2010.

[10] Anna Lina Ruscelli, Gabriele Cecchetti, Angelo Alifano, and Giuseppe Lipari. Enhancement of QoS support of HCCA schedulers using EDCA function in IEEE 802.11 e networks. *Ad Hoc Networks*, 10(2):147–161, 2012.

[11] Godwin Onyekachi Ugwu, Udora Nwabuoku Nwawelu, Mamilus Aginwa Ahaneku, and Cosmas Ikechukwu Ani. Effect of service differentiation on QoS in IEEE 802.11 e enhanced distributed channel access: a simulation approach. *Journal of Engineering and Applied Science*, 69(1):1, 2022.

[12] Wi-Fi Alliance. Global economic value of Wi-Fi® to reach $5 trillion in 2025. https://www.wi-fi.org/system/files/Economic_Value_of_Wi-Fi_Highlights_202407.pdf, 2024. [Online; accessed 16-Sep-2025].

[13] Evgeny Khorov, Anton Kiryanov, Andrey Lyakhov, and Giuseppe Bianchi. A tutorial on IEEE 802.11 ax high efficiency WLANs. *IEEE Communications Surveys & Tutorials*, 21(1):197–216, 2018.

[14] Lorenzo Galati-Giordano, Giovanni Geraci, Marc Carrascosa, and Boris Bellalta. What will Wi-Fi 8 be? a primer on IEEE 802.11 bn ultra high reliability. *IEEE Communications Magazine*, 62(8):126–132, 2024.

[15] Stefano Avallone, Pasquale Imputato, Getachew Redieteab, Chittabrata Ghosh, and Sumit Roy. Will OFDMA improve the performance of 802.11 WiFi networks? *IEEE Wireless Communications*, 28(3):100–107, 2021.

[16] Mingqi Han, Xinghua Sun, Wen Zhan, Yayu Gao, and Yuan Jiang. Multi-agent reinforcement learning based uplink OFDMA for IEEE 802.11ax networks. *IEEE Transactions on Wireless Communications*, 2024.

[17] Dheeraj Kotagiri, Koichi Nihei, and Tansheng Li. Distributed convolutional deep reinforcement learning based OFDMA MAC for 802.11 ax. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE, 2021.

[18] Evgeny Khorov, Ilya Levitsky, and Ian F. Akyildiz. Current status and directions of IEEE 802.11be, the future Wi-Fi 7. *IEEE Access*, 8:88664–88688, 2020.

[19] Mustafa Ergen. IEEE 802.11 tutorial. *University of California Berkeley*, 70:105, 2002.

[20] Shixun Wu, Xinrui Zeng, Miao Zhang, Kanapathippillai Cumanan, Abdulhamed Waraiet, Zheng Chu, and Kai Xu. LCVAE-CNN: Indoor Wi-Fi fingerprinting CNN positioning method based on LCVAE. *IEEE Internet of Things Journal*, 2025.

[21] Yi Chu, Kanapathippillai Cumanan, Sathish K Sankarpandi, Stephen Smith, and Octavia A Dobre. Deep learning-based fall detection using WiFi channel state information. *IEEE Access*, 11:83763–83780, 2023.

[22] Gaurang Naik, Jung-Min Park, Jonathan Ashdown, and William Lehr. Next generation Wi-Fi and 5G NR-U in the 6 GHz bands: Opportunities and challenges. *IEEE Access*, 8:153027–153056, 2020.

[23] Chenguang Rao, Zhiguo Ding, Kanapathippillai Cumanan, and Xuchu Dai. LDMA-NOMA beamforming in near-field communication systems. *IEEE Transactions on Vehicular Technology*, 2025. Early Access.

[24] IEEE Standards Association et al. P802. 11bn. *Institute of Electrical and Electronics Engineers (IEEE) Standards Association, Piscataway, NJ, USA*, 2023.

[25] Edward Oughton, Giovanni Geraci, Michele Polese, Vijay Shah, Dean Bubley, and Scott Blue. Reviewing wireless broadband technologies in the peak smartphone era: 6G versus Wi-Fi 7 and 8. *Telecommunications Policy*, 48(6):102766, 2024.

[26] Francesc Wilhelmi, Boris Bellalta, Szymon Szott, Katarzyna Kosek-Szott, and Sergio Barrachina-Muñoz. Coordinated multi-armed bandits for improved spatial reuse in Wi-Fi. In *2025 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*, pages 1–6. IEEE, 2025.

[27] Alvaro López-Raventós and Boris Bellalta. Multi-link operation in IEEE 802.11 be WLANs. *IEEE Wireless Communications*, 29(4):94–100, 2022.

[28] Lyutianyang Zhang, Hao Yin, Sumit Roy, and Liu Cao. Multiaccess point coordination for next-gen Wi-Fi networks aided by deep reinforcement learning. *IEEE Systems Journal*, 17(1):904–915, 2022.

[29] Boris Bellalta and Katarzyna Kosek-Szott. AP-initiated multi-user transmissions in IEEE 802.11 ax WLANs. *Ad Hoc Networks*, 85:145–159, 2019.

[30] J. Chun. Discussion on C-OFDMA operation. Tech. Rep. 11-23/0768r0, IEEE 802.11 Working Group, LG Electronics, May 2023.

[31] Pasquale Imputato and Stefano Avallone. Meeting latency constraints in Wi-Fi through coordinated OFDMA. In *2024 22nd Mediterranean Communication and Computer Networking Conference (MedComNet)*, pages 1–4. IEEE, 2024.

[32] Samyuktha Sena Indrasena and Rute C Sofia. Probabilistic Co-OFDMA: Performance in overlapping Wi-Fi 6 networks. In *2024 27th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pages 1–6. IEEE, 2024.

[33] Szymon Szott, Katarzyna Kosek-Szott, Piotr Gawłowicz, Jorge Torres Gómez, Boris Bellalta, Anatolij Zubow, and Falko Dressler. Wi-Fi meets ML: A survey on improving IEEE 802.11 performance with machine learning. *IEEE Communications Surveys & Tutorials*, 24(3):1843–1893, 2022.

[34] Douglas Dziedzorm Agbeve, Andrey Belogaev, Wim Sandra, Carl Lylon, and Jeroen Famaey. A2P: A scalable OFDMA polling algorithm for time-sensitive Wi-Fi networks. *arXiv preprint arXiv:2502.00430*, 2025.

[35] David González Filoso, Ryogo Kubo, Kazutaka Hara, Shinya Tamaki, Katsuya Minami, and Kohji Tsuji. Proportional-based resource allocation control with QoS adaptation for IEEE 802.11ax. In *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pages 1–6, 2020.

[36] Rong Yan, Mingjun Du, Xiao-Ping Zhang, and Yuhan Dong. Deep reinforcement learning based contention window optimization for IEEE 802.11 bn. In *2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)*, pages 1–5. IEEE, 2024.

[37] Jiaming Yu, Le Liang, Ziyang Guo, and Shi Jin. Heterogeneous multi-agent reinforcement learning for channel access in WLANs. In *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2025.

[38] Ziyang Guo, Zhenyu Chen, Peng Liu, Jianjun Luo, Xun Yang, and Xinghua Sun. Multi-agent reinforcement learning-based distributed channel access for next generation wireless networks. *IEEE Journal on Selected Areas in Communications*, 40(5):1587–1599, 2022.

[39] Hui Zhou and Yansha Deng. Federated reinforcement learning for uplink centric broadband communication optimization over unlicensed spectrum. *IEEE Transactions on Wireless Communications*, 2025.

[40] David Lopez-Perez et al. IEEE 802.11be extremely high throughput: The next generation of Wi-Fi technology beyond 802.11ax. *IEEE Commun. Mag.*, 57(9):113–119, Sep. 2019.

[41] Masashi Kobayashi et al. IEEE 802.11ad/WiGig based millimeter-wave small cell systems with adjacent channel interference suppression. In *IEEE Conference on Standards for Communications and Networking (CSCN)*, pages 1–5, 2016.

[42] Hongyang Du et al. Performance and optimization of reconfigurable intelligent surface aided THz communications. *IEEE Trans. Commun.*, 70(5):3575–3593, 2022.

[43] Francesc Wilhelmi et al. Throughput analysis of IEEE 802.11bn coordinated spatial reuse. In *IEEE Conference on Standards for Communications and Networking (CSCN)*, pages 401–407, 2023.

[44] Cailian Deng et al. IEEE 802.11be Wi-Fi 7: New challenges and opportunities. *IEEE Commun. Surveys Tuts*, 22(4):2136–2166, 2020.

[45] Yayu Gao et al. Throughput optimization of multi-BSS IEEE 802.11 networks with universal frequency reuse. *IEEE Trans. Commun*, 65(8):3399–3414, 2017.

[46] Pasquale Imputato and Stefano Avallone. Adaptive scheduling for downlink OFDMA in IEEE 802.11ax. In *European Wireless 2023; 28th European Wireless Conference*, pages 168–173, 2023.

[47] Mahboubeh Irannezhad Parizi, Mostafa Rahmani Ghourtani, Frank Scahill, and Kanapathippillai Cumaman. Resource unit allocation in coordinated OFDMA multi-user Wi-Fi systems. *IEEE Wireless Communications Letters*, 2025. Early Access.

[48] Thijs Havinga, Xianjun Jiao, Wei Liu, Baiheng Chen, Robbe Gaeremynck, and Ingrid Moerman. Fine-grained coordinated OFDMA with fiber backhaul enabled by open wifi and white rabbit. *arXiv preprint arXiv:2507.10210*, 2025.

[49] Guillermo Lacalle, Iñaki Val, Óscar Seijo, Mikel Mendicute, Dave Cavalcanti, and Javier Perez-Ramirez. Multi-AP coordination PHY/MAC management for industrial Wi-Fi. In *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–8. IEEE, 2022.

[50] Sudeep Bhattarai, Gaurang Naik, and Jung-Min Jerry Park. Uplink resource allocation in IEEE 802.11 ax. In *ICC 2019-2019 IEEE international conference on communications (ICC)*, pages 1–6. IEEE, 2019.

[51] Dmitry Bankov, Andre Didenko, Evgeny Khorov, and Andrey Lyakhov. OFDMA uplink scheduling in IEEE 802.11 ax networks. In *2018 IEEE international conference on communications (ICC)*, pages 1–6. IEEE, 2018.

[52] Naman Gupta, Syamantak Das, and Mukulika Maity. FairSplit - an efficient near-optimal bandwidth splitting strategy for OFDMA in IEEE 802.11ax. In *ICC 2022 - IEEE International Conference on Communications*, pages 285–290, 2022.

[53] Kaidong Wang and Konstantinos Psounis. Scheduling and resource allocation in 802.11ax. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, pages 279–287, 2018.

[54] Frank Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8(1):33–37, 1997.

[55] Frank P Kelly, Aman K Maulloo, and David Kim Hong Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research society*, 49(3):237–252, 1998.

[56] Kamil Szczech, Maksymilian Wojnar, Katarzyna Kosek-Szott, Krzysztof Rusek, Szymon Szott, Dileepa Marasinghe, Nandana Rajatheva, Richard Combes, Francesc Wilhelmi, Anders Jonsson, et al. Towards specialized wireless networks using an ML-driven radio interface. *arXiv preprint arXiv:2502.20996*, 2025.

[57] Francesc Wilhelmi, Boris Bellalta, Szymon Szott, Katarzyna Kosek-Szott, and Sergio Barrachina-Muñoz. Coordinated multi-armed bandits for improved spatial reuse in Wi-Fi. *arXiv preprint arXiv:2412.03076*, 2024.

[58] Maksymilian Wojnar, Wojciech Ciezobka, Katarzyna Kosek-Szott, Krzysztof Rusek, Szymon Szott, David Nunez, and Boris Bellalta. IEEE 802.11bn multi-AP coordinated spatial reuse with hierarchical multi-armed bandits. *IEEE Communications Letters*, 29(3):428–432, 2025.

[59] Raj Jain, Dah-Ming Chiu, and William Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems. Technical Report DEC-TR-301, Digital Equipment Corporation, September 1984.

[60] Violet Xinying Chen and John N Hooker. A guide to formulating fairness in an optimization model. *Annals of Operations Research*, 326(1):581–619, 2023.

[61] Haitham Al-Obiedollah, Haythem Bany Salameh, Ammar Gharaibeh, Kanapathippillai Cumanan, Zhiguo Ding, and Octavia A Dobre. Harvested power fairness-based multi-carrier NOMA IoT networks with SWIPT. *IEEE Wireless Communications Letters*, 12(2):381–385, 2022.

[62] Haitham Al-Obiedollah, Kanapathippillai Cumanan, Haythem Bany Salameh, Gaojie Chen, Zhiguo Ding, and Octavia A Dobre. Downlink multi-carrier NOMA with opportunistic bandwidth allocations. *IEEE wireless communications letters*, 10(11):2426–2429, 2021.

[63] Manijeh Bashar, Kanapathippillai Cumanan, Alister G Burr, Merouane Debbah, and Hien Quoc Ngo. On the uplink max–min SINR of cell-free massive MIMO systems. *IEEE Transactions on Wireless Communications*, 18(4):2021–2036, 2019.

[64] Kanapathippillai Cumanan, Leila Musavian, Sangarapillai Lambotharan, and Alex B Gershman. SINR balancing technique for downlink beamforming in cognitive radio networks. *IEEE Signal Processing Letters*, 17(2):133–136, 2009.

[65] Bozidar Radunovic and Jean-Yves Le Boudec. A unified framework for max-min and min-max fairness with applications. *IEEE/ACM Transactions on networking*, 15(5):1073–1083, 2007.

[66] Hamoud S Bin-Obaid and Theodore B Trafalis. Fairness in resource allocation: Foundation and applications. In *International Conference on Network Analysis*, pages 3–18. Springer, 2018.

[67] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, 8(5):556–567, 2000.

[68] Peng Xue, Peng Gong, Jae Hyun Park, Daeyoung Park, and Duk Kyung Kim. Max-min fairness based radio resource management in fourth generation heterogeneous networks. In *2009 9th International Symposium on Communications and Information Technology*, pages 208–213. IEEE, 2009.

[69] Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.

[70] Michael Grant. Cvx: Matlab software for disciplined convex programming, version 1.21. *http://cvxr. com/cvx*, 2011.

[71] Johan Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the 2004 IEEE International Symposium on Computer Aided Control Systems Design*, pages 284–289, Taipei, Taiwan, September 2004.

[72] Hayssam Dahrouj, Rawan Alghamdi, Hibatallah Alwazani, Sarah Bahanshal, Alaa Alameer Ahmad, Alice Faisal, Rahaf Shalabi, Reem Alhadrami, Abdulhamit Subasi, Malak T Al-Nory, et al. An overview of machine learning-based techniques for solving optimization problems in communications and signal processing. *IEEE Access*, 9:74908–74938, 2021.

[73] Phil Kim. Matlab deep learning. *With machine learning, neural networks and artificial intelligence*, 130(21):151, 2017.

[74] Yifan Chen, Shuang Wang, and Yunpeng Ge. A survey on the applications of image classification based on convolution neural network. In *2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, pages 381–384, 2022.

[75] Miao Zhang, Kanapathippillai Cumanan, Jeyarajan Thiyagalingam, Yanqun Tang, Wei Wang, Zhiguo Ding, and Octavia A Dobre. Exploiting deep learning for secure transmission in an underlay cognitive radio network. *IEEE Transactions on Vehicular Technology*, 70(1):726–741, 2021.

[76] Richard S Sutton, Andrew G Barto, et al. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

[77] Abdulhamed Waraiet, Kanapathippillai Cumanan, Zhiguo Ding, and Octavia A Dobre. Robust design for IRS-assisted MISO-NOMA systems: A DRL-based approach. *IEEE wireless communications letters*, 13(3):592–596, 2023.

[78] Abdulhamed Waraiet and Kanapathippillai Cumanan. Outage-constrained robust resource allocation framework for IRS-empowered NOMA systems: A DRL-based joint design. *IEEE Open Journal of the Communications Society*, 5:2748–2764, 2024.

[79] Abdulhamed Waraiet, Kanapathippillai Cumanan, Zhiguo Ding, and Octavia A Dobre. Deep reinforcement learning-based robust design for an IRS-assisted MISO-NOMA system. *IEEE Transactions on Machine Learning in Communications and Networking*, 2:424–441, 2024.

[80] Manijeh Bashar, Ali Akbari, Kanapathippillai Cumanan, Hien Quoc Ngo, Alister G Burr, Pei Xiao, Merouane Debbah, and Josef Kittler. Exploiting deep learning in limited-fronthaul cell-free massive MIMO uplink. *IEEE Journal on Selected Areas in Communications*, 38(8):1678–1697, 2020.

[81] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, February 2015.

[82] Lucian Busoniu, Robert Babuska, and Bart De Schutter. A comprehensive survey of multiagent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(2):156–172, 2008.

[83] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[84] Tian Li, Maziar Sanjabi, Ahmad Beirami, and Virginia Smith. Fair resource allocation in federated learning. *arXiv preprint arXiv:1905.10497*, 2019.

[85] Lyutianyang Zhang, Hao Yin, Zhanke Zhou, Sumit Roy, and Yaping Sun. Enhancing WiFi multiple access performance with federated deep reinforcement learning. In *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*, pages 1–6. IEEE, 2020.

[86] Christian Grasso, Raoul Raftopoulos, and Giovanni Schembra. Oscar: A contention window optimization approach using deep reinforcement learning. In *ICC 2023-IEEE International Conference on Communications*, pages 459–465. IEEE, 2023.

[87] Yiding Yu, Taotao Wang, and Soung Chang Liew. Deep-reinforcement learning multiple access for heterogeneous wireless networks. *IEEE Journal on Selected Areas in Communications*, 37(6):1277–1290, 2019.

[88] Lyutianyang Zhang, Hao Yin, Zhanke Zhou, Sumit Roy, and Yaping Sun. Enhancing WiFi Multiple Access Performance with Federated Deep Reinforcement Learning. volume 2020-November. Institute of Electrical and Electronics Engineers Inc., 11 2020.

[89] Lyutianyang Zhang, Hao Yin, Zhanke Zhou, Sumit Roy, and Yaping Sun. FLDRL-in-wireless-communication. https://github.com/Mauriyin/FLDRL-in-Wireless-Communication, 2020. [Accessed on 19 Aug 2025].

[90] Max Jaderberg, Valentin Dalibard, Simon Osindero, Wojciech M Czarnecki, Jeff Donahue, Ali Razavi, Oriol Vinyals, Tim Green, Iain Dunning, Karen Simonyan, et al. Population based training of neural networks. *arXiv preprint arXiv:1711.09846*, 2017.

[91] Madhavapeddi Shreedhar and George Varghese. Efficient fair queuing using deficit round-robin. *IEEE/ACM Transactions on networking*, 4(3):375–385, 1996.

[92] Yiding Yu, Taotao Wang, and Soung Chang Liew. Deep-Reinforcement Learning Multiple Access for Heterogeneous Wireless Networks. *IEEE Journal on Selected Areas in Communications*, 37(6):1277–1290, June 2019.

[93] Giuseppe Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on selected areas in communications*, 18(3):535–547, 2002.

[94] Paul Patras, Andrés Garcia-Saavedra, David Malone, and Douglas J Leith. Rigorous and practical proportional-fair allocation for multi-rate Wi-Fi. *Ad Hoc Networks*, 36:21–34, 2016.

[95] Ronald Y. Chang et al. Multicell OFDMA downlink resource allocation using a graphic framework. *IEEE Transactions on Vehicular Technology*, 58(7):3494–3507, 2009.

[96] Yu-Jung Chang et al. A graph-based approach to multi-cell OFDMA downlink resource allocation. In *2008 IEEE Global Telecommunications Conference*, pages 1–6, 2008.

[97] Konstantinos Dovelos and Boris Bellalta. Optimal resource allocation in IEEE 802.11 ax uplink OFDMA with scheduled access. *arXiv preprint arXiv:1811.00957*, 2018.

[98] S. Merlin et al. TGax simulation scenarios. *doc.: IEEE 802.11-14/0980r16*, Nov. 2015.

[99] Michael Neely. *Stochastic network optimization with application to communication and queueing systems*. Morgan & Claypool Publishers, 2010.

[100] Harold W Kuhn. The hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97, 1955.

[101] Robin J Wilson. *Introduction to graph theory*. Pearson Education India, 1979.

[102] Adrian Garcia-Rodriguez et al. IEEE 802.11 be: Wi-Fi 7 strikes back. *IEEE Commun. Mag.*, 59(4):102–108, 2021.

[103] Adrian Kosowski et al. Classical coloring of graphs. *Contemporary Mathematics*, 352:1–20, 2004.

[104] Hamza Chakraa et al. Optimization techniques for multi-robot task allocation problems: Review on the state-of-the-art. *Robotics and Autonomous Systems*, 168:104492, 2023.

[105] Vu Nguyen Ha, Georges Kaddoum, and Gwenael Poitau. Joint radio resource management and link adaptation for multicasting 802.11 ax-based wlan systems. *IEEE Transactions on Wireless Communications*, 20(9):6122–6138, 2021.

[106] Yousri Daldoul, Djamal-Eddine Meddour, and Adlen Ksentini. Ieee 802.11 n/ac data rates under power constraints. In *2018 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2018.

[107] Yeongjun Kim, Jonggyu Jang, and Hyun Jong Yang. Distributed resource allocation and user association for max-min fairness in HetNets. *IEEE Transactions on Vehicular Technology*, 73(2):2983–2988, 2023.

[108] Constantin P Niculescu. The integral version of Popoviciu's inequality. *J. Math. Inequal*, 3(3):323–328, 2009.

[109] Ning Wang and T Aaron Gulliver. Queue-aware transmission scheduling for cooperative wireless communications. *IEEE Transactions on Communications*, 63(4):1149–1161, 2015.

[110] Yang Yu and Mark Davis. Resource allocation for the IEEE 802.11 ax OFDMA mechanism based on dynamic programming combined with Timsort. In *2024 35th Irish Signals and Systems Conference (ISSC)*, pages 1–6. IEEE, 2024.

[111] Yang Lin, Rong He, Chuang Wang, and Xuming Fang. Adaptive transmission modes selection and resource allocation in Wi-Fi networks. In *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2025.

[112] Hayato Yamakata, Yosuke Tanigawa, and Hideki Tode. Selective OFDMA transmission method based on downlink communication quality requirements in IEEE 802.11 ax wireless LANs. In *2024 IEEE 21st Consumer Communications & Networking Conference (CCNC)*, pages 851–856. IEEE, 2024.

[113] David Nuñez, Pasquale Imputato, Stefano Avallone, Malcolm Smith, and Boris Bellalta. Enabling reliable latency in Wi-Fi 8 through multi-AP joint scheduling. *IEEE Open Journal of the Communications Society*, 6:2090–2101, 2025.