The Next Wi-Fi Generation: Towards Intelligent and Multi-Link Enabled Networks

Author: Álvaro López Raventós

TESI DOCTORAL UPF / Year of the thesis: 2022

THESIS SUPERVISOR
Boris Bellalta i Jiménez
Department of Information and Communication Technologies

Universitat Pompeu Fabra
Barcelona

To my parents, my sister,
and my beloved Alisha.

Acknowledgments

First and foremost, I want to express my deepest gratitude to my advisor Dr. Boris Bellalta. Pursuing a PhD would not have been possible without your trust, support and guidance through all these years. Also, thank you for encouraging me to take this opportunity, as it has been an amazing and life-changing journey.

Indeed, I want to give a special thanks to all the people, also the ones that have already departed, from the Wireless Networking Group (WNG). You embraced and took me as a member, giving me the best environment to grow as a researcher.

I would like to thank also to all the professors that have guided me to this day. From elementary school to university, all deserve the credit. Thank you for your patience and teaching.

The warmest gratitude is dedicated to my family. Without all of you, none of this would have been possible. Francisca and José Luis, my parents, thank you for all the values that you taught me, for teaching me the importance of hard work, and most important, for all the love and support that you give me. Marta, my sister, you have always been a role model to follow since I was a kid, your hard work inspires and encourages me to give always my best effort.

At last, to my beloved Alisha. I can not express how thankful I am for having you in my life. Thank you for helping me during the ups and downs of this thesis, and thank you for giving me the strength when I needed it the most. Thank you for being by my side, and loving me the way you do.

Abstract

The next Wi-Fi generation poses in front of a massive challenge as the main enabler for new services and applications. Traffic requirements are expected to keep rising year over year, challenging Wi-Fi networks to cope with vast amounts of data. To face such an issue, network densification seems to surge as a principal solution with the intention to create smaller service areas. However, such an approach is not exempt from drawbacks, as creating high-density deployments has its own costs. In this context, providing good network management is key to grant a good quality of experience. Yet, it is also a really challenging task.

In this thesis, we tackle such management and traffic constraints in two different parts. From the former, we address the management problem through the implementation of the Multi-Armed Bandits (MABs) framework, a simple machine learning mechanism that has been proven to be highly effective to combat the non-stationary nature of wireless systems. To that end, we employ MABs to perform a dynamic channel allocation (DCA) and a dynamic AP selection (DAPS) to overcome traditional and static approaches. Our results show that their adoption is able to provide a better performance than the traditional mechanisms, which worked under lower user density scenarios, but failed when tackling high-density (HD) ones. In particular, we observe that throughput losses are reduced by more than 15%, while being able to keep performance with a larger number of users.

On the other hand, new ways to grant more throughput are being developed under the next great evolution of the standard: the IEEE 802.11be Extremely High Throughput (EHT). We assess the proposed Multi-Link Operation (MLO), as the main candidate to improve network throughput, from an efficient traffic allocation perspective. We shed some light on how to perform traffic allocation through multiple interfaces, as such a question remains unresolved. In this context, we evaluate the application of different policy-based schemes, showing that MLO's performance is tied to a good allocation strategy. Our performance results show promising and relevant insights in this area while tackling coexistence issues with legacy devices.

Resum

La propera generació de Wi-Fi es postula davant d'un repte important com a principal promotor de nous serveis i aplicacions. Durant els pròxims anys, s'espera que els requisits d'aquests nous serveis continuïn augmentant, desafiant així les xarxes Wi-Fi amb quantitats massives de dades. Per fer front a aquest problema, densificar la xarxa amb més punts d'accés sembla ser la principal solució. Tanmateix, aquesta no està exempte d'inconvenients, ja que crear desplegaments d'alta densitat comporta els seus propis desavantatges. En aquest context, tenir una bona gestió de la xarxa és clau per permetre una bona experiència a l'usuari; tasca realment complicada donats els mecanismes actuals.

En aquesta tesi, estudiem tant la gestió de recursos com d'abasti-ment de tràfic en dues parts diferents. Primer, ens centrem en el problema de gestió mitjançant la implementació d'un dels mecanismes d'aprenentat-ge autònom, els Multi-Armed Bandits (MABs). Aquest mecanisme, tot i ser senzill, ha demostrat ser molt eficaç per combatre la naturalesa no estacionària dels sistemes sense fils. Així doncs, hem utilitzat els MABs per crear dues estratègies independents: una assignació dinàmica de canals i una selecció dinàmica dels punts d'accés. Totes dues tenen com a objectiu millorar el rendiment de les estratègies usades fins ara, les quals poden tenir una eficiència molt inferior en escenaris d'alta densitat. Els nostres resultats mostren que l'adopció d'ambdós mecanismes és capaç de proporcionar de manera eficient un millor rendiment de la xarxa. En particular, observem una reducció en les pèrdues del tràfic de més d'un 15%, alhora que podem mantenir un nombre més gran d'usuaris.

D'altra banda, noves maneres d'abastiment de tràfic s'estan desenvolupant sota la següent gran evolució de l'estàndard Wi-Fi: l'IEEE 802.11be Extremely High Throughput (EHT). En aquest context, nosaltres avaluem l'operació multi-enllaç com a principal candidat per millorar el rendiment de la xarxa, des de la perspectiva de l'assignació del tràfic, donant resposta a com realitzat aquesta assignació en sistemes multi-interfície. Per fer-ho possible, hem avaluat l'aplicació de diferents polítiques, demostrant que el rendiment de l'operació multi-enllaç està lligat a una bona estratègia d'assignació del tràfic. Els resultats presentats són prometedors, alhora que rellevants sobre aquesta àrea de recerca, els quals a més a més també consideren la coexistència amb sistemes més de primera generació.

Contents

List of Figures

List of abbreviations

AP	Access Point
BE	Best Effort
BO	Backoff
BSS	Basic Service Set
CA	Channel Allocation
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
DAPS	Dynamic AP Selection
DCA	Dynamic CA
DCF	Distributed Coordination Function
DIFS	Distributed Inter Frame Space
EHT	Extremely High Throughput
EDCA	Enhanced Distributed Channel Access
HD	High-Density
IDC	In-Device Coexistence Interference
ISM	Industrial, Scientific and Medical
KP	Knowledge Plane

MABs	Multi-Armed Bandits
MAC	Medium Access Control
MB-SL	Multi-Band Single Link
MCAA	Multi Link Congestion-Aware Load Balancing at Flow Arrivals
MCAB	Multi-Link Congestion-Aware Load Balancing
ML	Machine Learning
MLDs	Multi-Link Devices
MLO	Multi-Link Operation
MPDU	MAC Protocol Data Unit
MPTCP	Multipath TCP
MSLA	Multi-Link Same Load to All Interfaces
NG	Next-Generation
NFV	Network Function Virtualization
PHY	Physical
QoS	Quality of Service
RL	Reinforcement Learning
RSSI	Received Signal Strength Indicator
SDN	Software Defined Networking
SIFS	Short Inter Frame Space
SINR	Signal-to-Interference-plus-Noise Ratio
SL	Single Link
SLR	Supervised Learning

SLCI Single Link Less Congested Interface

SSF Strongest Signal First

STAs Stations

STR Simultaneous Transmission and Reception

TCP Transmission Control Protocol

TGbe Task Group be

TID Traffic Identifier

TPC Transmission Power Control

USL Unsupervised Learning

VDS Video and Data Separation

WLANs Wireless Local Area Networks

List of publications

The work presented in this thesis has been published in the following documents, which are sorted by publication date.

- **Journals**

  1. López-Raventós Álvaro, and Boris Bellalta. "Concurrent decentralized channel allocation and access point selection using multi-armed bandits in multi BSS WLANs". Computer Networks 180 (2020): 107381.
  2. López-Raventós Álvaro, and Boris Bellalta. "Multi-link Operation in IEEE 802.11be WLANs". IEEE Wireless Communications. IEEE, 2022.
  3. López-Raventós Álvaro, and Boris Bellalta. "Dynamic Traffic Allocation in IEEE 802.11be Multi-link WLANs". IEEE Wireless Communications Letters. IEEE, 2022.

- **Conferences**

  1. López-Raventós Álvaro, et al. "Combining software defined networks and machine learning to enable self organizing WLANs". 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). IEEE, 2019
  2. López-Raventós Álvaro, and Boris Bellalta. "IEEE 802.11 be Multi-Link Operation: When the Best Could Be to Use Only a Single Interface". 2021 19th Mediterranean Communication and Computer Networking Conference (MedComNet). IEEE, 2021.

Chapter 1

INTRODUCTION

1.1 Motivation

Based on the IEEE 802.11 standard family, Wi-Fi has been established as the main access technology for Wireless Local Area Networks (WLANs). This story of success is driven by Wi-Fi's simplistic design and operation, which allows to provide users with a cost-effective solution to satisfy their traffic demands. Operating under the unlicensed Industrial, Scientific and Medical (ISM) spectrum bands, Wi-Fi has gained interest from network operators, as a practical tool to ease congestion from cellular networks by means of traffic offloading. Such fact is corroborated in [1], which forecasts that 59% of the global mobile data traffic generated by 2022 (i.e., 111.4 exabytes/month) will be offloaded from cellular to both public and private Wi-Fi networks.

The trend is set, then, to the idea that Wi-Fi will have an even more pivotal role on future society, as it may become a feasible enabler for new services and applications such as, real-time 8K video, virtual and augmented reality, and cloud data services. In this context, ensuring sufficient network capacity for those services is fundamental to properly provide a good Quality of Service (QoS). To that end, network densification has been used by operators (i.e., more than 17 million of commercial[1] Wi-Fi

[1]Commercial Wi-Fi spots refer to those that are installed to offer public Wi-Fi at *cafés* and restaurants, retail chains, hotels, airports, planes, and trains for customers and guests.

spots are forecast by 2022 [1]) as an straightforward solution to provide more system capacity, since other factors like the theoretical limits of spectral efficiency, alongside to the spectrum scarcity, are the main bottlenecks in wireless communications. From corporate offices to mass events, outdoor hotspots, shopping malls, airports, exhibition halls, and stadiums, we find an increasingly number of Access Point (AP), which task is to create smaller service areas to provide connectivity to a reduced set of users.

On the downside, creating High-Density (HD) deployments has its own costs. Such fact relies on different factors, some of them being, likewise, great contributors to Wi-Fi's fame. On one hand, Wi-Fi operates in the unlicensed ISM bands, which are well-known for being globally available to any system. Essentially, this means that several technologies can coexist under the same spectrum of frequencies, and so, creating wild environments with multiple heterogeneous contenders and high inter-network interferences. This effect may be further exacerbated if considering the emergence of the internet of things.

On the other hand, there is adoption of the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) to perform channel access. The main advantage to use this protocol is that it does not involve high coordination techniques, unlike cellular networks, allowing to preserve high network scalability, while keeping a simple operation that favors to maintain devices at cheaper costs. However, this approach constrains Wi-Fi nodes to a half-duplex system, in which they need to remain silent if sensing the medium busy. Such an effect is critical in HD deployments as it implies a reduced chance to gain access to the medium. Additionally, the use of a randomized Backoff (BO) counter to resolve contention is not ideal in dense scenarios as the probability to have collisions[2] increases, which results into experience a significant throughput degradation.

In this context, providing good network management and throughput performance for future applications is going to be a really challenging task to achieve, if current mechanisms are not improved. To that end, we find that Machine Learning (ML) are gaining a lot of interest to address management issues, since they can exploit the high amounts of data generated within the network to understand interactions, and pro-

[2] A collision is said to happen when more than one node attempts to send a packet through the wireless medium at the same time.

vide intelligent responses to them. Moreover, new ways to improve the throughput performance are tackled under the IEEE 802.11be Extremely High Throughput (EHT) [2] amendment. There, the Multi-Link Operation (MLO)[3] stands out as a major feature, since it promotes the use of multiple wireless interfaces to allow concurrent data transmission and reception in APs and Stations (STAs) with dual- or tri-band capabilities. Hence, enabling higher throughput values, while achieving a better use of the spectrum resources.

This thesis, first, evaluates the application of ML mechanisms to perform an intelligent channel allocation and AP selection. Such approaches, which are taken under decentralized adversarial conditions (i.e., each action is taken independently from actions of the others), provide a flexible solution to adapt to different scenarios, while creating reactive systems to combat anomalies. Second, it evaluates the MLO as candidate to highly improve network performance, unraveling that such feature opens up currently unsuitable scenarios for nowadays network operations.

1.2 Contributions

In light of network management optimization for HD WLANs, in this thesis, we study the application of ML strategies to assess an intelligent channel allocation and AP selection. Additionally, we aim to shed some light on future MLO feature to provide high throughput improvements. Specifically, the main contributions of this thesis are summarized as follows:

1. We consider an state-of-the-art ML solution for improving network performance in dense WLANs. In particular, we assess the contributions of the Multi-Armed Bandits (MABs) framework to resolve different problems in HD and complex deployments.

2. We evaluate the implications of using decentralized MABs in adversarial WLAN environments, with the aim to observe the feasibility of

[3]Throughout this thesis, we will refer to the multi-band/multi-channel operation feature as the MLO, following the notation in [2].

their implementation. In this context, we consider the joint combination of decentralized channel allocation and AP selection problems in large WLAN scenarios.

3. We provide an overview on how the Next-Generation (NG) of WLANs are envisioned. Particularly, we delve into the MLO framework proposed by the Task Group be (TGbe), pointing out the different modifications in regards the nodes' architectural changes, transmission modes and management functions.

4. We discuss the benefits, and challenging open issues related to the integration of the MLO framework in IEEE 802.11be WLANs. Besides, we take advantage of MLO's capability to seamlessly manage the traffic to implement a traffic manager to avoid load balancing issues.

5. We extend the MLO functionalities by empowering the use of multiple policy-based strategies to tackle the traffic allocation problem with favorable results. Besides, we present different insights regarding coexistence issues between legacy and 802.11be WLANs from the the perspective of both types of devices.

6. We design and evaluate a flow-based implementation of the MLO in a custom-based network simulator that allows the simulation of HD scenarios. In this context, we validated against the ns-3 our simulation platform providing accurate results, while highly reducing the simulation time.

1.3 Document structure

This thesis is a compendium of articles resulting from the research activity on network management optimization for HD WLANs. We refer to the publications attached at the end of this document as **paper #1** through **paper #5**. Apart from the list of publications, a monograph is provided to introduce the research topic and give some background on the same. This document is structured as follows. Chapter 2 depicts different main performance issues caused by Wi-Fi's conceptual design under HD scenarios. Chapter 3 delves into the different actions taken to increase the performance of WLANs, as a combination between intelligent mechanisms and new proposals on the amendment. Chapter 4 presents the analytical and simulation tools used throughout this thesis. Chapter 5 provides the performance evaluation and results obtained, whereas Chapter 6 outlines the concluding remarks, and future work directions.

Chapter 2

HIGH-DENSITY 802.11 WLANS

In this chapter, we address the problem statement by means of characterizing different performance challenges for HD WLANs. Then, we present some background literature on those identified challenges to overview different research directions to solve them.

2.1 Introduction

Since its appearance back in the late '90s, Wi-Fi has been continuously innovating to adapt its performance to the new user demands. From the oldest amendments, which only introduced new modulation and coding schemes, to the newest IEEE 802.11ax, which presents newer techniques to enhance network efficiency in HD scenarios, all have captured, and shaped today's Wi-Fi. Yet, its basic operation has kept the same. Indeed, keeping the same principle throughout the years has enabled backwards compatibility from older versions, which also contributed to Wi-Fi's popularity and success.

The basic operation in 802.11 WLANs is performed through the Distributed Coordination Function (DCF), which employs the CSMA/CA protocol as underlying mechanism to perform a listen-before talk access strategy. Under the DCF, if a MAC Protocol Data Unit (MPDU), namely a packet, is ready to be transmitted, a BO counter is randomly picked from a uniform distribution, so the transmission process can start. Once the BO is fixed, the transmitting node starts sensing the medium for a Distributed Inter Frame Space (DIFS) period before starting the BO countdown. Then, if during the DIFS period the medium has been sensed as idle, the BO starts to be decremented until the packet is eventually transmitted. If during the BO countdown the channel is sensed as busy, the BO is paused and resumed once the channel is detected idle again. At last, if the transmission succeeds, the destination node acknowledges the received packet after waiting a Short Inter Frame Space (SIFS) interval.

The DCF works under the assumption that the traffic of the different contenders has the same latency requirements, and so, it only provides a Best Effort (BE) service. However, the explosion of multimedia applications lead to include some mechanism to provide service differentiation. In this context, traffic prioritization is tackled in the 802.11e amendment, which introduced the Enhanced Distributed Channel Access (EDCA). The EDCA is an extension of the DCF, which provides concurrent access to four independent queues of different priorities, each one with different parameters (i.e., CW and DIFS). Thus, high priority queues have lower values on their access parameters, allowing to reduce access delays for high priority traffic.

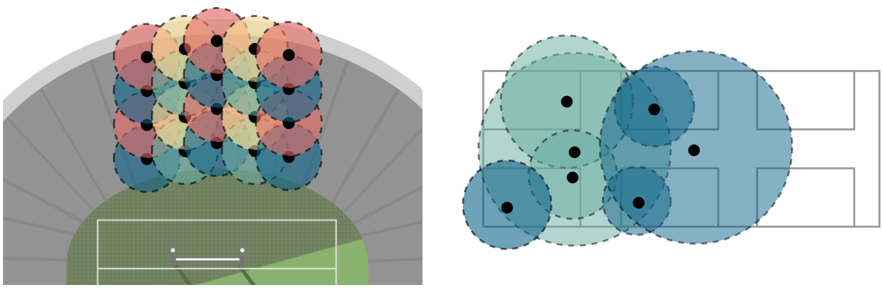
Although DCF and EDCA operations are extremely flexible, having a distributed Medium Access Control (MAC) has its own cost. The lack of coordination could make multiple nodes to end their BO at the same time, and therefore, accessing the channel all at once. Additionally, since CSMA/CA is based on sensing the channel, the heterogeneous environment of Wi-Fi networks may result into the well-known hidden and exposed node problems [3,4], which may be exacerbated under HD WLANs. Hence, in the next section, we characterize different performance challenges in HD WLANs, in order to get a bigger picture on different flaws of Wi-Fi, and what concerns need more attention.

2.2 Performance challenges in dense networks

Dense networks are known for having higher density numbers of users (e.g., 1 user/m^2) over the same location size [5], if compared to traditional non-dense scenarios. Consequently, the number of serving APs to provide the same quality of experience need also to be higher. In [5], some examples are described giving a very accurate representation on the requirements and issues for future dense WLANs. From public to private sector locations, it is expected, then, to find multiple networks with large number of APs, given the traffic and latency requirements of future services and applications.

Following the dense scenarios described in [5], we identify how different performance issues may appear over different HD WLANs, as a result of their conceptual design and application. On one hand, in public spaces such as stadiums, conference halls, and airports, among others, we find networks carefully planned by service providers. In such scenarios, APs are placed in optimal or semi-optimal locations, and they are configured to avoid frequency overlaps and excessive transmission power configurations, trying to reduce as much as possible inter-network interferences. Figure 2.1a depicts an schematic of a dense scenario in which service areas are kept small, while interferences are avoided by creating a frequency reuse pattern. However, we identify a potential issue that escapes the control of the operators: the mobility of users. Since wireless networks are about mobility, the distribution of users may not be uniformly spread over the area, creating situations where multiple APs may become highly saturated. Such issue is related to AP selection mechanism, as the standard scheme does not use any information about the load of the channel.

On the other hand, shown in Figure 2.1b, we find apartment buildings, where APs tend to be placed in non-optimal places. Also, they are



(a) Stadium (b) Apartment building floor plan

Figure 2.1: Characterization of dense network scenarios

normally configured with selfish and aggressive transmission power configurations (i.e., maximum power), and zero or limited awareness on the channel[1] being used. Hence, such harmful configurations lead to have an uncontrolled interference environment (i.e., several devices may use the same channel), with excessive power configurations, creating significant overlaps between multiple networks. As a result, different issues such as the hidden and expose node problems, the flow starvation and/or the well-known performance anomaly [6] may appear due to the heterogeneous nature of devices (i.e., short and high range) that coexist under the same area, reducing the overall performance in multiple independent WLANs.

2.3 Channel allocation and transmission power control

Channel Allocation (CA) poses as a main constraint in unplanned networks (i.e., home scenarios), as the uncoordinated nature of such scenarios may prevent from a proper interference mitigation. However, the appearance of dynamic approaches such as the works in [7, 8], overcomed traditionally fixed channel assignments over the APs. Besides, as a desperate measure to improve the Signal-to-Interference-plus-Noise Ratio (SINR), transmission power configurations tend to be excessively aggressive, which cause even more chaotic deployments. In this context, solutions regarding CA and Transmission Power Control (TPC) mechanisms are needed to try to mitigate the lack of network planning.

In literature, we find in [9] that the CA problem is treated as a map coloring problem with the objective to assign non-overlapping channels to adjacent APs. Also, [10] presents an scheme where the AP placement and CA are optimized under sequentially and jointly approaches. In such work, the objective relies on a maximization problem to enhance the total throughput on the service area while satisfying a specified number of APs. Additionally, [11] presents a mechanism in which each AP autonomously searches for the most emptiest channel, switching channels if during the scanning process another less congested channel is found. Authors in [12]

[1] In wireless communications, channels are identifiers that refer to a combination of frequency and bandwidth.

attempted to follow a min-max approach in order to assign channels in an adaptive manner, so the maximum channel utilization at the most overloaded AP is minimized. Also, in [13], authors propose an approach that exploits an smart channel selection strategy by classifying the traffic pattern on primary channels, and choosing the channel with the longest idle time. It is shown that using their approach the amount of collisions can be reduced drastically, increasing the system performance. In [14–16] opportunistic channel allocation schemes are intended to overcome frequency holes, promoting a higher frequency utilization by accessing them in a short-interval basis. More recently, authors in [17] proposed a channel allocation method based on graph analysis, linear programming and regression to minimize the overlap among APs. In [18], authors propose a dynamic-wise, light-weight and decentralized, online primary channel selection algorithm for performing dynamic channel bonding, which considers the activity on both target primary and secondary channels in order to maximize the expected throughput. Regarding the use of ML over the CA problem, we find different works that already tackle such approach. For instance, [19] propose an online CA by adopting the MABs framework. They implement a weighted algorithm in order to carry the action selection process, in which the probability of selecting a certain action is adjusted according to the regret observed. An study about the exploration-exploitation trade-off for different learning algorithms with the objective to achieve the best pair of channel and power allocation is presented in [20]. In addition, authors compare the performance of the different considered action selection strategies, while studying the implications of applying them under an adversarial setting.

Regarding TPC solutions, in [21] we find that authors propose a framework to determine the optimum setting with the objective to maximize the network throughput for elastic traffic. In [22], the authors propose a TPC algorithm that tunes the transmission power of the APs in order to be able to support all its clients at the highest transmission rate. Also, in [23] an adaptive TPC jointly with a rate selection scheme is proposed to maximize the energy efficiency of IEEE 802.11 stations. Other forms to maintain under control excessive transmission power configurations have come up in the form of controlling the carrier sense range. In this context, under the 802.11ax amendment the spatial reuse (SR) operation is proposed to mitigate the effects of excessive transmission powers, for instance, [24] extensively reviews such operation providing relevant insights on such operation. Additionally, its potential gains are shown under different density conditions in [25].

2.4 Access point selection

The standard association method for IEEE 802.11 WLANs is based on the Strongest Signal First (SSF) scheme, which determines the best AP to be associated based on the Received Signal Strength Indicator (RSSI). However, it seems unfeasible to apply the SSF strategy in HD scenarios. The main reason is that the SSF is based on the RSSI, purely a Physical (PHY) metric, and it does not take into consideration how much traffic is already being handled by the destination AP. Therefore, in networks with high number of APs, very close to each other, decisions taken through the SSF may create highly unbalanced situations, where some APs may be overload and others practically unused.

In literature, we find multiple works that propose different AP selection mechanisms, some of them jointly considered alongside load balancing techniques. For instance, an association control algorithm is evaluated in [26] where the users' bandwidth demands are considered to estimate the AP utilization. In [27], a decentralized AP selection scheme is considered to select the AP that offers the best SINR. The average workload of the network is used in [28] to redistribute the traffic when a new station joins the network. A similar scheme is presented in [29] where the stations are migrated to the least loaded AP in order to balance the traffic load upon a new association. In [30], an online Markov process based AP selection scheme for 802.11n users is evaluated under coexistence with other heterogeneous clients (802.11a/b/g/n). The AP selection problem is modelled using graph theory in [31]. Cell breathing is used in [32] to tune the cell size of the APs accordingly to their load and their neighbors' upon new associations. Similarly, the same technique is used in [33] over an Software Defined Networking (SDN) architecture. Also, in [34] it is proposed an association scheme over a SDN-based WLAN, which is capable to detect situations where traffic is not efficiently distributed, and reschedule to other APs the clients whose transmissions are causing performance issues. Mininet is used in [35] to test an algorithm over an SDN controller to decide whether to accept or reject new stations based on their load.

From centralized to decentralized, those works show different implementations to perform the AP selection under different conditions. However, most of them are expected fail under dynamic environments due to the rigid approach of their solution. In particular, some of them may show a ping-pong effect, without converging into a solution, due to their threshold based approach.

2.5 Summary

In this chapter, we overviewed the basic operation of 802.11 WLANs, as well as, characterizing different scenarios and their issues. Depending on their design, we have seen that the major concerns are related to channel and power allocation, and the AP selection. Also, we reviewed multiple relevant works on those topics. Next, we detail future Wi-Fi ecosystems in order to create a reactive networks, as the static approaches of the presented works may fail to resolve unexpected situations.

Chapter 3

TOWARDS A NEW WI-FI ECOSYSTEM

In this chapter, we present the start-of-the-art and research directions to enhance network resource management and performance. First, we point future network architectures as enablers to accommodate adaptive and flexible mechanisms. Then, we overview such mechanisms, as well as, presenting future directions of Wi-Fi technology.

3.1 Adaptable and flexible networks

Next generation WLANs, need next generation solutions. As well as protocols and technologies, the underlying systems need to evolve. Therefore, advancing towards more flexible and adaptable systems is key to provide the desired performance, as network deployments get denser. A paradigm gaining popularity for wireless networking is the SDN [36, 37], which relies on decoupling control and data planes to allow a centralized controller to orchestrate and manage the entire network according to a set of programmed applications. Also, SDN is normally related and/or combined with the Network Function Virtualization (NFV) [38]. Although those architectures provide some sort of network flexibility through their software-based approach [39], there are still multiple features that need to be carefully considered in terms of adaptability. In particular:

- The network needs to be able to automatically detect nodes' capabilities, and their requirements, as well as, their radio environment, in order to match the network resources to the users it is serving;

- since wireless networking is about mobility and dynamic environments, the network should identify mobility patterns to perform resource reservation, as well as, to provide seamless mechanisms to avoid or minimize service cuts due to the handover process;

- the network needs to match users' traffic requirements even in high dynamic conditions. Thus, adaptation to non-static conditions is crucial to provide a good quality of experience. Then, architectures need to be prepared to create reactive mechanisms to prevent anomalies, or unexpected situations.

To meet those features, the adoption of ML is gaining some interest, as potential enabler for all of them [40]. We find, for instance, that the first one can be fulfilled by classification strategies, which can optimize nodes' configurations based on their radio environment [41]. Also, mobility pattern recognition is possible through clustering techniques as presented in [42], enabling the second feature. Lastly, the third one can be met by the use of the MABs framework, which provides a reactive mechanism to adapt network parameters dynamically based on actions taken by the

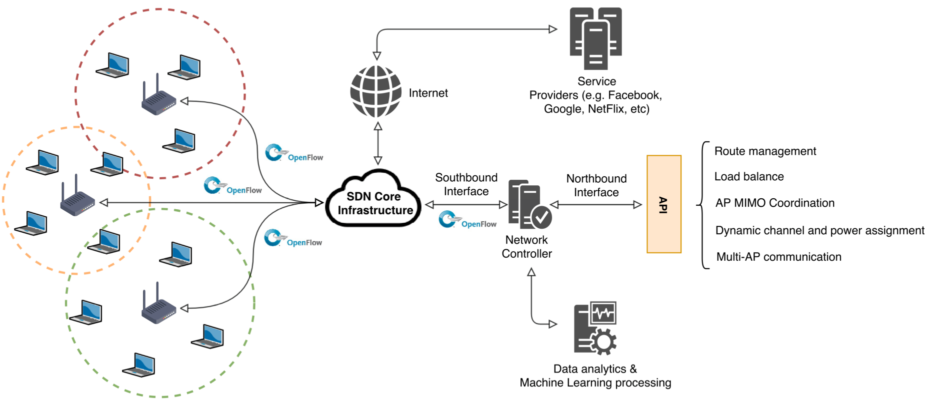


Figure 3.1: Joint use of SDN and ML over an enterprise WLAN

surrounding nodes [20].

In **paper #1**, we provide a conceptual design, as well as a feature evaluation of a fully programmable WLAN architecture based on the SDN paradigm. We consider a Knowledge Plane (KP) [43], which is responsible for learning the behavior of the network and the decision-making process. Therefore, information is collected into the KP, and transformed into knowledge via ML algorithms, which actively participate in the network orchestration. Figure 3.1 outlines an architecture that merges both KP and SDN, allowing an adaptable and flexible wireless environment.

3.2 Machine learning for 802.11 WLANs

ML stands up as a tool to improve the network performance by improving its resource management [44]. To do so, algorithms are fed with the data collected through the network, which allows them to learn from past experiences, unraveling multiple usage patterns that are applied to manage the entire system. Besides, under a network environment as complex as the wireless medium, ML provides reactive mechanisms to handle unexpected situations, or anomalies, without being specifically designed for it.

3.2.1 Taxonomy of ML Techniques

ML tasks depend on the nature of their training data. In this context, there exists, in general, three main classes of learning approaches:

- **Supervised Learning (SLR):** these algorithms use a learning technique in which a labeled dataset (i.e., inputs, and known outputs) is used to build the model that represents the relations between the inputs and the outputs. To fine tune these models, the error between predictions and measurements is evaluated in order to minimize it [45, 46].

- **Unsupervised Learning (USL):** under USL algorithms are given a set of inputs without their outputs. In this context, these techniques aim to find patterns, structures, or knowledge in unlabeled data by clustering samples into different groups according to the similarity between them [47, 48].

- **Reinforcement Learning (RL):** RL algorithms normally involve an agent, an action and an environment. There, agents are learning entities that interact with its environment to learn the best action from its action space[1] to maximize its long-term reward [49, 50].

The adoption of any of the aforementioned ML approaches, and their subsequently implementation, is followed by a common dilemma: where to put the ML mechanism in order to maximize its potential. Indeed, depending on the taxonomy of the problem, and the underlying network capabilities, we may prefer to go with a centralized [40] or decentralized [51] approach, with its own advantages and drawbacks. On one hand, centralized ML-driven decisions can gather information of more nodes, giving a great perspective on the global state of the network. Thus, the system has lots of data to take more accurate decisions about future network states, but at a cost of large response delays. We find SLR and USL to be the most suitable techniques to be applied under a centralized environment, as they are typically used for classification, regression, and clustering.

On the other hand, moving ML processes to the edge allows to have more particular solutions with quicker responses to a dynamic environment. However, they tend to lose resolution about the state of the network, since their limitation to collect data. For such cases, RL mechanisms are the most suitable ones, as their learn-by-interaction approach allows them to learn the most convenient action to take based on the observed reward. Hence, RL algorithms are useful to overcome the non-stationary conditions of the wireless medium by providing reactive mechanisms to address possible system malfunctions, such as scenarios where the load is unbalanced, or with high interference issues. To overcome such problems, the MABs, as well as, Q-learning algorithm are some of the most common RL methods used.

To embrace the fact that wireless environments can change from second to second, we explored in **paper #1** and **paper #2** a multi-agent MAB setting for different WLAN optimization purposes. Next, we present the MABs framework, and some preliminaries on our work.

[1]We refer to the action space to the set of actions that a given agent can take in order to obtain a reward from its environment.

3.2.2 Multi-Armed Bandits for decentralized WLAN optimization

In literature, we find that MABs have been extensively used to enhance the performance for wireless networks [44]. In this context, we find that the traditional MABs problem models the interaction between a learning agent, often called a player, and an unknown environment. To do so, the learner sequentially chooses one of its k available actions receiving a reward from the environment, which is used to evaluate the performance of the action, as well as, to select the subsequent actions. Then, the goal of the learning agent is to maximize the long-term reward to reach an optimal result. In such quest, agents need to deal with a natural trade-off between exploration (actions that have already provided rewards) and exploitation (actions that have to be chosen to learn their rewards), since a lack of exploration may lead to a suboptimal solution when the rewards from the arms are still unexplored. A further and extensive introduction to the MABs can be found in [52].

The MABs problems are typically classified depending on the nature of their rewards. In particular, stochastic [53, 54], adversarial [55], contextual [56] and bayesian [57] are the four main groups of MABs. For instance, in stochastic bandits actions have and independent and identical reward distribution, whereas in bayesian bandits, an arm is selected following a probability distribution, which is proportional to the historic of the rewards experienced by that arm. However, independently of the type, the main objective still is to find the arm maximizing the payoff. To that end, a common way of measuring the performance of MABs algorithms is by means of the regret function. The regret for a player i at time t, after T rounds is

$$R_{i,t} = \mu^*_{i,t} \cdot T - \sum_{t=1}^{T} \mu_{i,t}, \tag{3.1}$$

where $\mu^*_{i,t}$ is defined as the reward given by the optimal action at time t, and $\mu_{i,t}$ is the reward obtained by the current action selected. From the regret definition, learn is said to happen if the cumulative regret function grows sublinearly, and therefore, the algorithm is able to identify the action with the highest reward. In this case, the expected regret, $\mathbb{E}[R_{i,t}]$, will decrease over time, converging to zero.

Algorithm 1: TS MAB with a Gaussian prior

Input: set of possible actions, $\mathcal{A} = \{1, ..., K\}$

for each arm $k \in \mathcal{A}$, set $\hat{\mu}_k = 0$ and $n_k = 0$

while *active* **do**

  For each arm $k \in \mathcal{A}$, sample $\theta_k(t)$ from $\mathcal{N}(\hat{\mu}_k(t), \sigma_k^2(t))$

  Select arm $k = \underset{i=1,...,K}{\operatorname{argmax}}\ \theta_i(t)$

  Observe and compute the reward experienced $r_k(t)$

  $\hat{\mu}_k(t) \leftarrow \frac{\hat{\mu}_k(t) \cdot n_k(t) + r_k(t)}{n_k(t)+2}$

  $n_k(t) \leftarrow n_k(t) + 1$

end

To address the action-selection strategy, we considered throughout this thesis, the use of the Thompson Sampling (TS) [58] algorithm. The TS algorithm is a Bayesian algorithm that selects a given action based on its past noticed performance. To do so, during the learning stage, TS observes the reward, and updates its prior belief in a way that the probability of a particular arm being optimal matches with the probability of each arm being selected. In practice, this is done by sampling each arm k from its posterior distribution $\theta_k(t)$, and selecting the one that returns the maximum expected reward. This property results very useful, allowing us to tackle the intrinsic non-stationarity of our environment. Hence, arms that were chosen initially because of their good rewards, can be discarded over time if they start to perform badly. Algorithm 1 shows the implementation of the TS using a Gaussian prior [57] for a given agent.

Following the MAB approach, in **paper#1**, we proposed a collaborative behavior to maximize a shared reward defined by a max-min throughput policy. Although keeping decentralized learning procedure for each agent, we achieve collaboration between the different players by addressing an online learning [59] setting through a software defined wireless network. Thus, agents' rewards are distributed at each step across the different participants, so they can update their payoff for the action taken. Figure 3.2a depicts one of the multiple scenarios evaluated in **paper#1**, where $\mathrm{AP_A}$ and $\mathrm{AP_C}$ are considered to be active from the start, whereas AP_B is triggered at the half of the simulation. Such an scenario gives a clear representation on how MABs work, as it tackles not only non-stationary conditions, but the flow starvation problem.

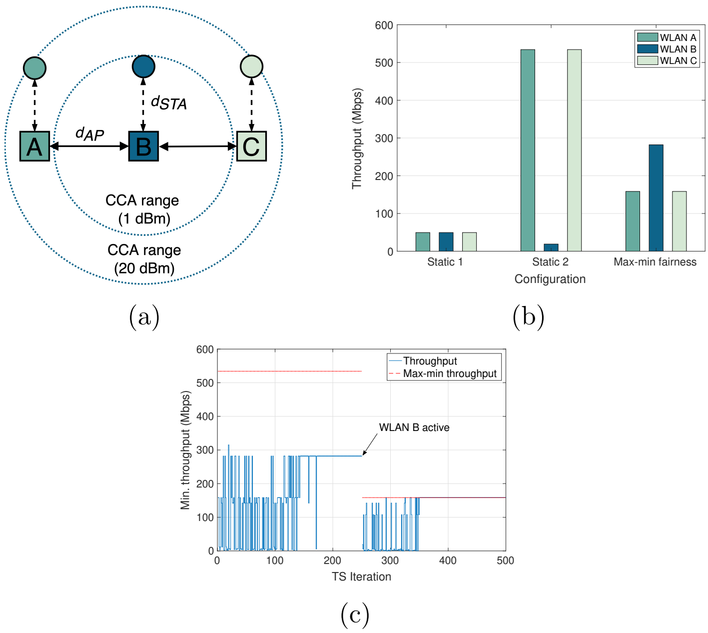


Figure 3.2: Optimization through decentralized MABs

In Figure 3.2b there are represented the throughput values obtained for different configurations under this scenario. On one hand, selecting the action named *static 1* (i.e., the combination of the lowest power value and channel bandwidth) allows to avoid harmful situations, as all APs get the same throughput, but at expenses of not leveraging all the available resources. Additionally, selecting the action named *static 2* (i.e., the combination of the highest power and channel bandwidth) is neither an option, as it only allows AP_A and AP_C to achieve their maximum throughput but at the cost of causing high interferences to AP_B, which is unable to get practically any.

On the other hand, the MAB implementation based on the max-min throughput policy allows to achieve a network state in which all three APs are able to get significant values of throughput. Indeed, the collaboration between APs is found on the fact that AP_A and AP_C give up on throughput in favor of AP_B, which clearly is the most benefited. Besides, Figure 3.2c shows the TS evolution through each iteration for AP_A. Here, it is observed an interesting behavior of the AP_A agent, since it is not able to reach the optimal value for the first half of the iterations. Such an effect happens due to the low probability of AP_A and AP_C selecting the best action at the same time. Consequently, a suboptimal solution is equally achieved at both APs. On the contrary, when AP_B is activated, the AP_A agent is able to achieve the optimal value, as a result of the reduced actions that lead to have a favorable reward. It is found, then, that such type of collaborative environment between agents empowers a fairer system solution.

Later, in Section 5.1, we delve into MABs with the findings obtained from **paper#2**, in which an adversarial multi-agent MAB setting was tested for HD WLAN deployments.

3.3 The Multi-link operation

Wireless data services will continue to grow since upcoming applications such as virtual/augmented reality, video/game streaming and cloud based services, are expected to request more and more data with the most demanding throughput, latency, and reliability requirements [60]. Although ML mechanisms can help to better manage the network, squeezing its performance to the limits, the 802.11 standard needs to continue its evolution. In this context, the 802.11 TGbe was created to address the development of new specifications to fuel the upcoming Wi-Fi 7 [2]. Through the multiple 802.11be features [2], we find specially interesting the MLO. This feature is intended to promote the use of multiple wireless interfaces to allow concurrent data transmission and reception, while standing out for being extremely flexible and enabling a seamless use of the different resources to provide, for instance, an opportunistic spectrum access [61].

To support such implementation, we find the redefinition of classical APs or STAs into the so-called Multi-Link Devices (MLDs)[2], whose MAC architectural implementation is characterized by presenting a unique MAC

[2]Either AP MLDs or STA MLDs, refer to single devices with multiple wireless interfaces.

instance to the upper layers, without losing the independent parameters of each interface [62]. That is, common and link specific functionalities are split into the so-called upper and lower MAC layers, respectively. Figure 3.3 depicts the MLD architecture, representing both MAC sub-layer levels.

The motivation behind this two-tier architecture is to permit MLO-capable devices to make a concurrent use of all available resources without modifying the actual implementation of the upper layers. In this context, we find that the MAC layer follows a natural evolution, which has been already present in the transport layer. We refer to the Multipath TCP (MPTCP) [63], which implementation set the trend to towards the simultaneous use of multiple interfaces. MPTCP stands as a natural extension[3] of Transmission Control Protocol (TCP) [64] allowing a pair of hosts to use several paths to exchange the segments that carry the data from a single connection[4]. To do so, MPTCP creates multiple underlying TCP connections, also known as subflows, and presents a unique instance of the transport layer to the upper layers, allowing complete interoperability to any application without the need of TCP being modified [63]. Indeed, MLO and MPTCP have a similar approach in the sense that multiple independent MAC or TCP instances coexist under a unique one. However, MPTCP presents multiple limitations since its implementation

[3]Although being a natural extension, MPTCP is not mandatory, and therefore, multiple systems keep the basic TCP implementation.

[4]A connection is defined to be a logical link, in which traffic of a certain application is exchanged between two end hosts.

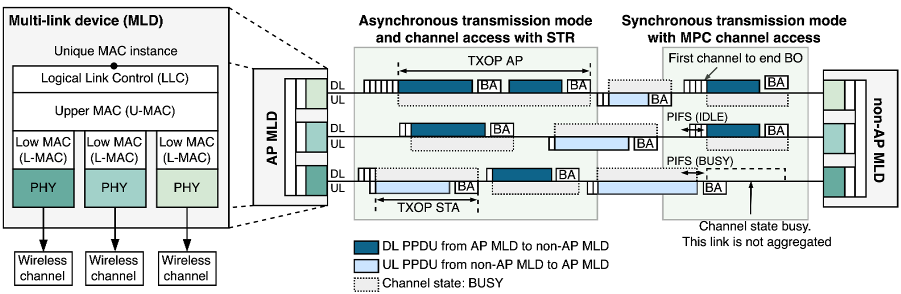


Figure 3.3: Multi-link architecture and transmission modes representation.

is tied to the rigidity of the transport layer. For example, MPTCP depends on having different IP addresses assigned to the available interfaces, as well as, it needs to establish a complete TCP three-way handshake for each subflow [63]. Besides, once a segment is assigned to a given path, it can not be relocated to another unless a retransmission of that segment is needed. In this context, MLO removes those limitations, as it is implemented at a lower layer (i.e., MAC layer), avoiding the need of multiple IP addresses, and empowering a much more efficient traffic management, which allows to move traffic from one link to another in a seamless manner.

Under the MLO umbrella, we identify the asynchronous transmission mode to enable the Simultaneous Transmission and Reception (STR) capability, which allows frames to be sent or received through multiple interfaces, as depicted Figure 3.3. Although such scheme may provide an enormous throughput improvement, it is not exempt of different constraints. Certainly, such an aggressive implementation (i.e., having multiple asynchronous interfaces operating at the same time) may be followed by energy saving mechanism, specially for handheld devices, which consumption may be significantly affected. However, it is not in the energy consumption, but in the In-Device Coexistence Interference (IDC) where we find its major issue. That is, the power leakage between interfaces may prevent a frame reception on one link, during an ongoing transmission on the another one. Such issue appears as a result of not having enough separation between operating bands/channels (e.g., two channels in the 5 GHz band), and therefore, causing excessive levels of interference on the receiving interface that prevent the asynchronous operation.

To avoid the IDC issue, it is also defined the synchronous mode, which relies on synchronized frame transmissions across the available links. Devices operating under a synchronous mode are referred to as constrained MLDs, or non-STR MLDs, since they are not allowed to transmit through an idle interface at the same time they are receiving through another. To perform synchronization, the end-time alignment or the defer transmission mechanisms [2] may be implemented. While the former relies on ending transmissions on different channels at the same time, the latter defers the transmission of a link that has finished its BO, until the end of the same counter in other links. With that, APs or STAs are prevented to perform the STR operation, avoiding IDC problems, but at the cost of a lower throughput, if compared to the asynchronous scheme. At last, MLD-capable APs or STAs may change its transmission modes (e.g., asynchronous to synchronous, and vice versa) at any time, as depicted in Figure 3.3.

We encourage the reader to **paper #4** for further details on the MLO implementation, as we review more in-depth the nodes' architecture, and transmission schemes, as well as, other relevant aspects on the operation and management.

3.3.1 Open issues and challenges

Although the MLO represents a promising functionality to be implemented in next generation WLANs, the concurrent use of multiple interfaces brings multiple challenges to face off. In this context, we point out some open issues that require further research:

- *non-STR and legacy blindness.* This issue relates to the fact that non-STR and legacy STAs may cause different collision scenarios, as a consequence of their constrained operation. First, non-STR STAs may be unable to detect an intra-BSS transmission (either downlink or uplink) in one of its available links, because of performing a transmission on another. Therefore, a collision may occur if non-STR STAs attempt to transmit over that link already in use. This issue has been already tackled in [65] by allowing AP MLDs to inform non-STR stations about the channel state in other links in use by the AP MLD to prevent such a situation. On the other hand, similarly, legacy devices may not know if a transmission is taking place in others links, since they only operate in a single one. Hence, some indication, as the proposed in the non-STR case, is needed to inform legacy nodes of the activities happening in the other links.

- *Spectrum inefficiency.* Conservative approaches to avoid the IDC interference or collisions can lead to an inefficient use of the spectrum, because of suspending the BO procedure in one link, if medium access is granted in another one. In this regard, an opportunistic BO mechanism to maximize the spectrum utilization of non-STR nodes is proposed in [65], so transmission attempts can be resumed only when the channel state guarantees a collision with not happen.

- *Channel access fairness.* Since MLO allows to perform transmission opportunity aggregation over different links, nodes with single link availability may experience starvation due to their higher difficulties to access the channel. Therefore, in presence of legacy stations the usage of link aggregation techniques should be limited or restricted, in order to minimize unfair situations.

- *Load balancing.* Further research is required to fully understand which is the best strategy to balance the traffic in MLO WLANs. For instance, it is important to consider also how MLO can be used for uplink traffic, as it may require a completely different approach than its downlink counterpart. In this aspect, load balancing strategies can benefit from the use of machine learning solutions to predict future traffic and network dynamics.

3.3.2 Traffic policing

The multi-interface availability in MLDs allows thinking about the implementation of a traffic manager, in order to resolve load balancing issues. Following the proposals of the TGbe, such a logical entity should be placed at the upper MAC level, so the interface assignation is performed as traffic goes through it [66]. Then, once a connection is established between an AP-STA pair, and traffic streams start to flow, the manager is in charge to allocate traffic based on the set of rules defined by the selected policy. Such an approach allows to achieve not only an efficient use of the network resources, but a better control the capabilities of MLDs supporting, for instance, advanced traffic differentiation, beyond the default MLO's Traffic Identifier (TID) to link mapping functionality. In the following, we introduce the different allocation policies presented in **paper #3**, **paper #4** and **paper #5**, which can be classified into non-dynamic and dynamic, depending on their behavior.

Non-dynamic policies

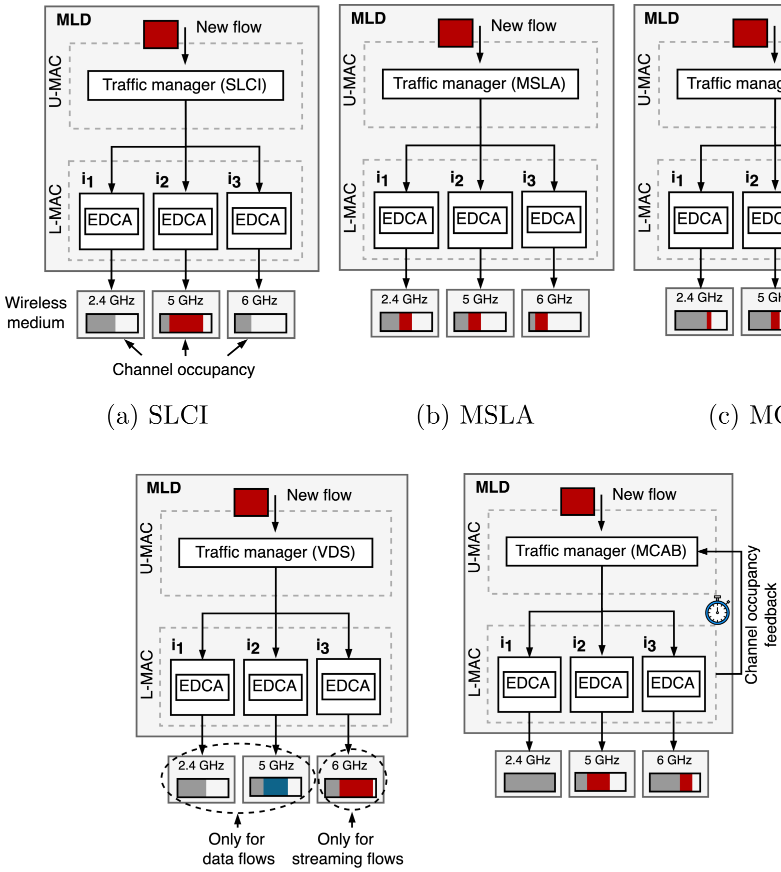
Under a non-dynamic strategy, each flow maintains the same traffic-to-link allocation during its lifetime. That is, upon a flow arrival, the channel occupancy is gathered, and the traffic is distributed either proportionally over multiple interfaces according to their congestion, or fully into the less congested one. We define the following non-dynamic policies

- **Single Link Less Congested Interface (SLCI).** On a flow arrival, pick the less congested interface, and allocate all the traffic to it. The SLCI operation is shown in Figure 3.4a.

- **Multi-Link Same Load to All Interfaces (MSLA).** Upon a new flow arrival, distribute the incoming traffic flow equally between all the enabled interfaces of the receiving station. That is, let ℓ be the bandwidth requirement of the incoming flow, and N the number of enabled interfaces in the destination station. Thus, the traffic allocation per interface is given by: $\ell_i = \ell/N$, with ℓ_i the bandwidth allocated to interface i. The MSLA operation is depicted in Figure 3.4b.

- **Multi Link Congestion-Aware Load Balancing at Flow Arrivals (MCAA).** On a flow arrival, distribute the new incoming flow's traffic accordingly to the observed channel occupancy at the AP, considering the enabled interfaces of the receiving station. Namely, let ρ_i the percentage of available (free) channel airtime at interface i. Then, the fraction of the flow's traffic allocated to interface i is given by $\ell_{i\in\mathcal{J}} = \ell\frac{\rho_i}{\sum_{\forall j\in\mathcal{J}}\rho_j}$, with ℓ being the the bandwidth requirement of the new flow, and $\mathcal{J}$ the set of enabled interfaces at the target station. If there are any other active flows at the AP, their traffic allocation remain the same as it was. The MCAA operation is presented in Figure 3.4c.

Due to their straightforward approach, these non-dynamic policies are well-suited for scenarios where the interfaces' congestion levels remains almost stationary. Their computational cost is low, as only few calculations are done at flow arrivals. Although the previous strategies considered all TIDs to be mapped into the multiple interfaces, the MLO opens up the possibility to perform a link-based traffic separation through the TID-to-link mapping functionality. That is, different TIDs may be mapped to different links, in order to minimize, for instance, access delays for time-sensitive traffic [67, 68]. Besides, such feature may be complemented by



(a) SLCI (b) MSLA (c) MCAA

(d) VDS (e) MCAB

Figure 3.4: Schematics for the different traffic allocation policies

the fact that nodes' spatial distribution may create different contention-free links, specially in the 5 GHz and 6 GHz bands, as a result of favorable radio propagation conditions. Therefore, traffic with higher QoS requirements can be exclusively exchanged through those contention-free links, as long as they exist. In this context, in **paper #4**, we extended the previous policies by adding an additional one, which distinguishes between traffic flows of different types.

- **Video and Data Separation (VDS).** Upon a flow arrival, allocate data flows to the 2.4 GHz or the 5 GHz interface, whereas video flows will be allocated to the one at 6 GHz. Besides, data flows will not be distributed across multiple interfaces, but to a single one (i.e., either the 2.4 GHz or 5 GHz, selecting always the emptiest). The VDS operation is shown in Figure 3.4d.

Dynamic policies

A dynamic strategy is said to able to periodically adjust the traffic-to-link allocation in order to follow channel occupancy changes, and so, taking the most out of the different enabled interfaces. In this regard, a traffic (re)allocation may be triggered by two different events: a new flow arrival or a periodic timer, which wakes up every δ units of time. Under both events, the channel occupancy is gathered to proportionally (re)distribute the traffic load of all active flows to any of the enabled interfaces. It is worth mention that, the dynamic reallocation of traffic is performed by adjusting the interfaces' traffic weights (i.e., traffic percentage associated to each one), which are tracked by the traffic manager at the upper MAC level. Besides, we consider such reallocation to be instantaneous. We define the following dynamic policy

- **Multi-Link Congestion-Aware Load Balancing (MCAB).** On a flow arrival or at every δ units of time, collect the channel occupancy values and sort all flows (including the incoming one) in ascending order, considering the number of enabled interfaces at the destination station (i.e., first the flows with less enabled interfaces). In case two or more flows have the same number of enabled interfaces in the destination station, they are ordered by arrival time. After, start (re)allocating the flows' traffic accordingly to the same procedure as in MCAA. The MCAB operation is presented in Figure 3.4e.

Through its dynamic implementation, the MCAB minimizes the effect of neighboring BSSs actions, as they usually result in abrupt changes in the observed congestion at each link. Therefore, such policy scheme is able to adjust the traffic allocated to each link, exploiting the different traffic activity patterns while maximizing the traffic delivery. However, it is noticeable that the MCAB gain is conditioned to perform multiple operations in shorts amounts of time, which may be impractical in high density areas, as the computational requirements to (re)distribute all flows grows with the number of active users.

In **paper #3**, **paper #4** and **paper #5** we adopt those policies from a flow-level perspective, in order to provide an evaluation on the proposed traffic manager, aiming to maximize the use of the spectrum resources. The main findings on their application can be found in Section 5.2.

3.4 Summary

In this chapter, we presented a big picture towards the evolution of Wi-Fi ecosystems. We introduced new architectural paradigms to enable more flexible and adaptable networks, as the rigidity of nowadays systems may limit the operation of upcoming techniques. In this context, we shown that ML is gaining a lot of attention to allow self-managed networks, as learning from data has become a feasible solution to enhance management operations. At last, we overviewed the MLO feature for the next Wi-Fi generation. There, we presented the main modifications to the standard, pointing out different open issues regarding the MLO operation. In this context, we proposed different policy-based strategies to face the traffic allocation problem, as it may downgrade MLO's performance if not tackled properly.

Chapter 4

METHODOLOGY

In this chapter, we introduce the methodology and assumptions made throughout the evaluation of the mechanisms presented in previous chapters. Besides, we introduce the custom-based simulation made to obtain the results.

4.1 A flow-level perspective of the CSMA/CA

4.1.1 Flows rather than packets

Over the Internet, traffic is typically modelled following either a flow- or a packet-level approach. The former represents traffic as a continuous, or intermittent, stream of packets having some criteria in common (i.e., IP addresses or port numbers) [69]. Hence, flows are considered the minimum transfer unit, which allows to have larger event time scales (i.e., scales of seconds). Also, this representation provides with a high level abstraction on how network resources may be shared among the different participants, aiming to capture, for instance, flow-level dynamics such as the bandwidth allocation [70]. On the other hand, the latter is intended to describe accurately the mechanics of the network at much lower level, making it useful to characterize and capture the different states in the buffers, as well as, queueing and propagation delays.

In this context, throughout this thesis, we consider a flow-level traffic model to perform our evaluation. Such consideration, allows us to study the performance of WLANs considering a flow as a fluid [69], and therefore, showing from a more general perspective how the network resources are shared among the different participants. Through a flow-level model, we aim to take advantage of having larger event time-scales to study large and complex scenarios. Besides, this approach allows us to make many repetitions in an affordable amount of time, and so, getting representative results. However, its main constrain is related to the fact that such methodology is not able to capture interactions that happen at lower temporal scales (i.e., channel access level), as if packet-level was considered.

Following the Internet traffic, flows can be classified into different types. A very basic classification can be done in terms of their traffic requirements. First, there are *streaming flows*, also called rigid, since their bandwidth requirements remain constant during the time that the flow is alive. On the other hand, we find the *elastic flows*, which their main characteristic is their ability to adapt its rate to the network state. Besides, we find that the average holding times for most streaming flows (e.g., video conferencing) is of the order of several minutes. This constitutes a much larger time scale of traffic variations than for elastic flows (e.g., web browsing) which typically last a few seconds [71].

4.1.2 A flow-level CSMA/CA abstraction

To consider the aforementioned flow-level model, we abstracted the channel access operation. While the considered abstraction does not capture low-level details of the PHY and MAC layers operation, it maintains the essence of the CSMA/CA: the fair share of the spectrum resources among contending APs and stations. Basically, the considered abstraction takes into account the aggregate channel load at each AP to calculate the airtime that can be allocated to each station.

To explain the abstraction, let us consider the Figure 4.1, which has represented 2 overlapping BSSs. Both APs are configured with the same channel, and have different STAs associated to them. Moreover, on the left side of Figure 4.1, there are represented different instants of time, which depict the channel load experienced at $\mathrm{AP_B}$. As shown, at instant time $t = 10$, both $\mathrm{STA_{1,B}}$ and $\mathrm{STA_{2,B}}$ become active, requesting some bandwidth from its serving $\mathrm{AP_B}$. Such demand is interpreted as airtime, which is defined as the amount of time needed to satisfy the requested bandwidth requirements of each station. The generalized expression of the airtime for an station i requiring a flow of bandwidth B_i from its serving AP j, and with the length of a data packet L_d, and the bit rate $r_{i,j}$ is given by

$$u_{i,j}(B_i, L_\mathrm{d}, r_{i,j}) = \frac{1}{(1-p_e)} \left\lceil \frac{B_i}{L_\mathrm{d}} \right\rceil (E[\psi]t_e + t_s(r_{i,j}, L_\mathrm{d})), \tag{4.1}$$

where p_e is the packet error probability, the term $\frac{1}{(1-p_e)}$ represents the average number of transmissions per packet, $E[\psi]$ is the average BO duration, t_e is the duration of an empty slot, and $t_s(r_{i,j}, L_\mathrm{d})$ is the duration of a successful packet transmission, which is given by

$$t_s(r_{i,j}, L_\mathrm{d}) = t_\mathrm{RTS} + 3t_\mathrm{SIFS} + t_\mathrm{CTS} + t_\mathrm{DATA}(r_{i,j}, L_\mathrm{d}) + t_\mathrm{ACK} + t_\mathrm{DIFS} + t_e, \tag{4.2}$$

where t_DIFS and t_SIFS are the DIFS and SIFS time, $t_\mathrm{DATA}(r_{i,j}, L_\mathrm{d})$ is the duration of a data packet at the transmission rate used between station i and AP j, and t_ACK correspond to the time that an acknowledgment packet lasts.

If the airtime required is met, as it happens in $t = 10$, both STAs will be satisfied with all the bandwidth requested. On the other hand, at

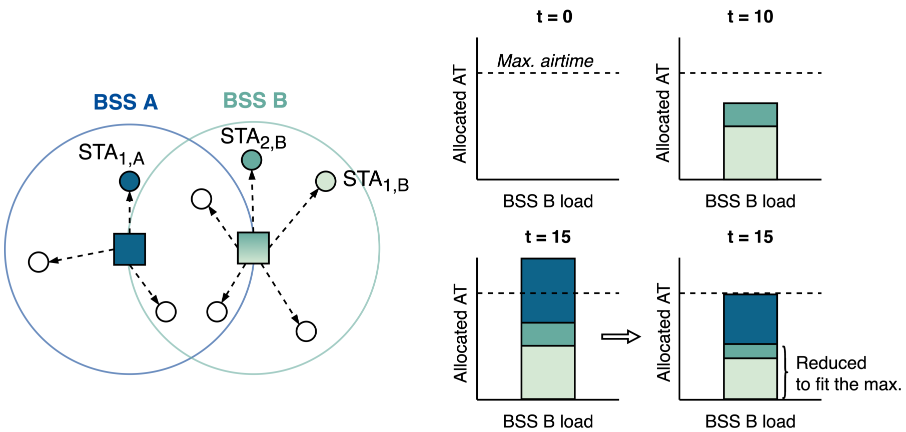


Figure 4.1: CSMA/CA abstraction representation

$t = 15$, $\mathrm{STA}_{1,\mathrm{A}}$ becomes active while the other ones are active too. With that, we observe that the channel load experienced by AP_{B} surpasses the maximum allocable airtime. Indeed, such effect is caused by both APs sharing the same spectrum channel. Hence, the effective channel load perceived at AP_{B} using a channel c is generically expressed as

$$\ell_j^c(t) = \sum_{\forall n \in \mathcal{N}_j^c} \ell_n^c(t) + \sum_{\forall i \in \mathcal{S}_j} u_{i,j}(B_i, L_{\mathrm{d}}, r_{i,j}). \tag{4.3}$$

where $\ell_j^c(t)$ will depend on the airtime required by the own flows of AP j, and $\ell_n^c(t)$, which is the airtime registered due to flows from neighboring APs using the same radio channel.

Since the effective channel load surpasses the maximum allocable time, the airtime of the flows need to be fitted to the maximum. Indeed, this condition implies that the bandwidth requirement can not be met and each flow registers from losses, having unsatisfied stations. The concept of satisfaction in this thesis is key to understand the network performance, as it allows us to quantify how good the flows are performing. We defined the satisfaction of an station i associated to AP j operating in channel c at time t as

$$\Omega_{i,j}(t) = \frac{\min(1, \ell_j^c(t))}{\ell_j^c(t)} \leq 1 \tag{4.4}$$

where $\ell_j^c(t)$ is the channel load as defined in (4.3). Since we consider that all resources are proportionally distributed in our CSMA/CA abstraction, the satisfaction value obtained by stations under the same AP, or APs that having the same set of neighbors, will be the same. Finally, the throughput achieved by a flow from station i, associated to AP j at time t is given by

$$\Gamma_{i,j}(t) = B_i \Omega_{i,j}(t) \tag{4.5}$$

As an example, we assume that $\mathrm{STA}_{1,\mathrm{B}}$ and $\mathrm{STA}_{2,\mathrm{B}}$ require a traffic load of 40% and 30% respectively $t = 10$. The channel load perceived by both APs, then, adds up to 70%, lower than the maximum 100%, and therefore, making stations to be satisfied as they receive the airtime allocation that they need. On the contrary, at $t = 15$, we consider that $\mathrm{STA}_{1,\mathrm{B}}$ and $\mathrm{STA}_{2,\mathrm{B}}$ still requires 40% and 30% respectively, but $\mathrm{STA}_{1,\mathrm{A}}$

becomes active requesting a traffic need up to 60%. This higher requirement of $\text{STA}_{1,\text{A}}$ makes both APs to enter in saturation, since the total channel load raises up to 130%. As a result, the satisfaction experienced by the three stations scores a value of 76.9%. Essentially, this value indicates that only the 76.9% of the required load of the stations (i.e., 30.77% 23.08%, and 46.15% for $\text{STA}_{1,\text{B}}$, $\text{STA}_{2,\text{B}}$ and $\text{STA}_{1,\text{A}}$, respectively) will be allocated.

4.2 A custom flow-level simulator for high density networks

The need for a flexible, low complex and scalable flow-level simulator lead the creation of a new tool to provide with an easy environment to get meaningful results. The Neko simulator[1] is designed to allow users perform simulations of large 802.11 deployments getting relevant results within a reasonable amount of time. To do so, the simulator is based on the COST libraries [72], which allows to perform different synchronous and asynchronous events that characterize for instance, creation of new flows, and activation or deactivation of stations. Under its current state, the simulator supports the implementation of MABs for a single link WLANs, either to optimize the channel allocation, or the AP selection, following the procedure described in Algorithm 1. To that end, agents are able to collect meaningful information of their environment to perform the decision-making process. Additionally, it fully supports the MLO's asynchronous transmission mode, in which multiple nodes can transmit and/receive data from different interfaces. However, the fact of considering a the CSMA/CA abstraction represents that the Neko simulator is not able low-level interactions such as the case where two or more interfaces can be aggregated at channel access level.

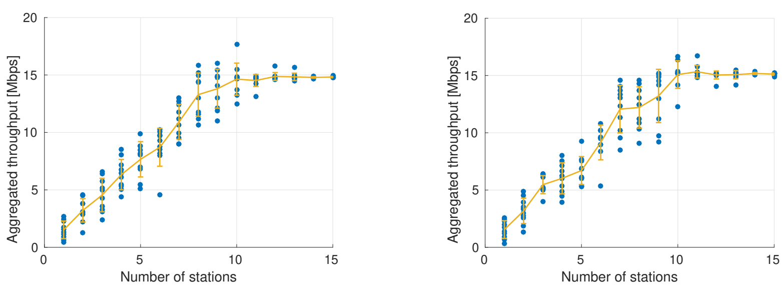
In order to ensure the accuracy of the presented CSMA/CA abstraction, we performed a performance validation against the well-known ns-3 [73] simulator. In this regard, we aimed to prove that the assumed abstraction reproduces its actual operation. To that end, we emulate the same conditions on both simulators, over a reference deployment with

[1] `https://github.com/wn-upf/Neko`

2 overlapping BSSs that share the same radio channel. The results are shown in Figure 4.2 Figure 4.3. From the figure, we observe that, until 9 deployed STAs, the abstraction is able to reproduce the same performance as the one obtained in ns-3 for both APs. However, we find significant differences when reaching a saturation regime (i.e., the limit reached by the system throughput as the offered load increases) [74] with 10 deployed STAs. While the ns-3 keeps their performance strictly to the saturation limit, the CSMA/CA abstraction may have higher values. For instance, for the AP_A in Figure 4.3a, we observe a value near 20 Mbps when having 13 STAs. Such phenomena is caused by the CSMA/CA abstraction, which shares among the different participants the losses proportionally to their traffic requirements. Consequently, for the same point in Figure 4.3b, we observe a low near 10 Mbps. As a result, the average between the values for both APs, when using the Neko simulator, is the same as the average obtained from ns-3.

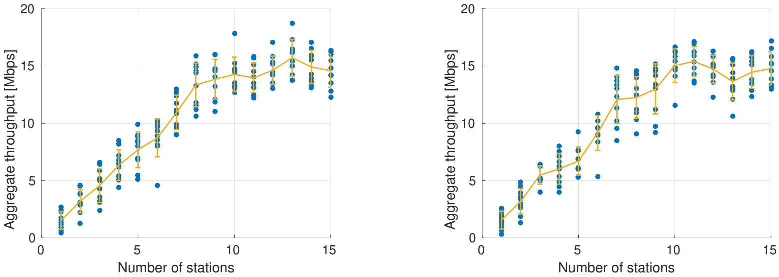
4.3 Summary

This chapter provided an overview on the methodology used throughout this thesis. First, we justify our election to model the Internet traffic through a flow-level perspective, pointing out the main benefits and drawbacks of this high-level approach. Then, we introduced the CSMA/CA abstraction for a flow-level traffic modelling, characterizing how the air-



(a) AP_A (b) AP_B

Figure 4.2: Performance evaluation obtained using ns-3



(a) $\mathrm{AP_A}$ (b) $\mathrm{AP_B}$

Figure 4.3: Performance evaluation obtained using Neko

time allocation is managed. At last, we presented the developed simulation tool, performing its validation against a well-known simulator and demonstrating the accuracy of the abstraction. Next, we provide the performance evaluation of the mechanisms presented in Chapter 3, and the derived main findings.

Chapter 5

PERFORMANCE OPTIMIZATION FOR NG WLANS

In this Chapter, we present the main findings of this thesis, as a result from the analysis on the multiple evaluations provided through the different papers. First, we focus in **paper #2** that unravels the potential gains that MABs may introduce under the adversarial HD deployments of future WLANs. In this context, we observe the feasibility of its applications addressing multiple scenarios, while tackling the benefits of following a reinforcement learning approach. Later, we evaluate the introduction of the upcoming MLO feature for the 802.11be amendment. Then, **paper #3**, **paper #4** and **paper #5** introduce a set of policies to shed some light on how to perform an effective traffic allocation.

5.1 Dynamic channel allocation and AP selection

In **paper #2**, we assessed the application of MABs under an adversarial setting. The adversarial environment is developed through the fact that rewards experienced by actions depend on how the other players behave. So, from a point of view of single player, its opponents will have control over its rewards. For instance, a given AP may experience a bad channel reward if it changes to the channel used by its neighbor, or if its neighbor changes to its operating channel. Also, it is important to note that players belonging to different groups may interact during the decision-making process. Hence, the exemplified AP will also get a negative reward if all the stations select it as their serving AP. Indeed, an interesting contribution of the work presented in **paper #2** is the consideration on the interactions between two different types of players.

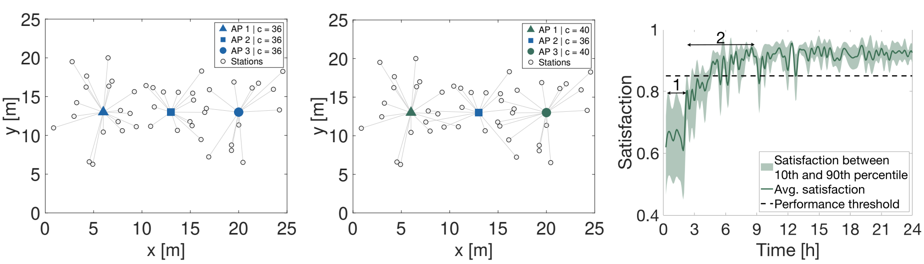
Players are naturally classified into APs and STAs, which rewards are defined accordingly to the optimization problem. For STAs, their objective will be to maximize the satisfaction (defined in Equation 4.4) experienced from a given AP. On the contrary, APs will try to minimize its channel occupancy by trying to maximize their channel reward. The channel reward fo an AP j that uses a channel c at time t is given by

$$\Psi_j^c(t) = \max(0, 1 - \ell_j^c(t)) \leq 1, \tag{5.1}$$

where $\ell_j^c(t)$ is the one defined in Equation 4.3.

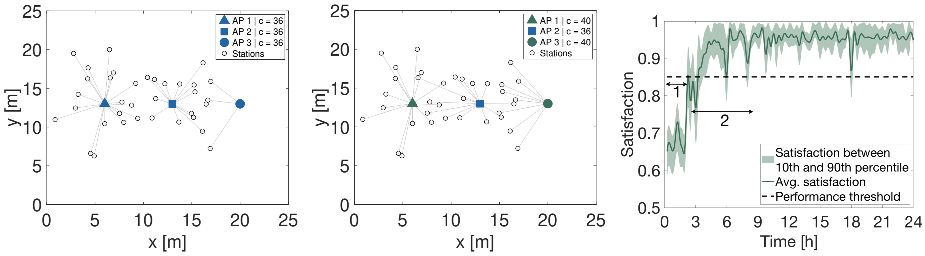
To combat the instability of adversarial MABs, we limit the actions of STAs. In this context, we set a quality threshold (i.e., $\mathrm{RSSI_{th}}$) which needs to be met for an AP to become eligible. With that, not only limit the action state, reaching solutions in a faster pace, but we avoid an excessive data rate degradation that could cause to have unnecessary and unrealistic high airtime values.

We first take the evaluation of the DCA and DPAS mechanism over a controlled scenario, evaluating two different cases. On one hand, Figure 5.1a shows a balanced scenario with 3 BSSs having the same number



(a) (b) (c)

Figure 5.1: Evaluation of MABs over a controlled scenario



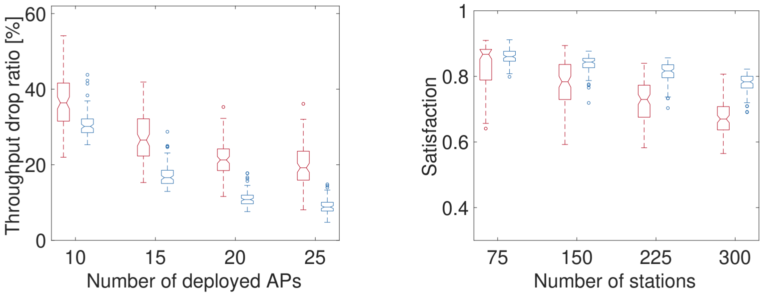
(a) (b) (c)

Figure 5.2: Evaluation of MABs over an unbalanced controlled scenario

of stations. Moreover, all 3 APs are configured with the same radio channel. After executing the simulation, we observe through the satisfaction curve in Figure 5.1c how the learning process (i.e., step marked as 2) is carried successfully until convergence is achieved. Although the gains are significant, they are mainly due to the correct channel allocation of the APs, which allows to avoid overlapping between neighboring BSSs. Hence, under balanced situations the DAPS mechanisms can not provide any major improvement as Figure 5.1b shows that the number of stations keeps more or less the same.

On the other hand, we assess an unbalanced station distribution across the different BSSs, as shown in Figure 5.2a. Now, the satisfaction curve depicted in Figure 5.2c has a more prominent slope than the one in the previous case, as here, both mechanisms are working jointly to enhance the network performance. Hence, Figure 5.2b shows how the most congested APs are alleviated by the DAPs mechanism, which allows STAs to migrate to the less congested APs, getting better satisfaction results but at the cost of losing some data rate. Such characterization, then, reveals a significant finding as the DCA is essential to avoid frequency overlaps, but the DAPS mechanism is able to provide significant performance gains under unbalanced conditions.

Now, we take the evaluation of joint operation between the DCA and DAPS to random deployments. Here, we evaluate i) AP densification, and ii) growing user density cases. For the former, the user density is kept the same but the number of the APs deployed over the network is sequentially increased. For the latter, the AP number is kept constant, while the



(a) AP densification (b) Growing user density

Figure 5.3: Evaluation of MABs over AP random scenarios

number of users is increased. Indeed, the introduction of randomness into the scenarios allows us to create unbalance conditions to show the potential of both mechanisms under large network scenarios. Figure 5.3 shows the performance obtained for both cases. In red there are depicted results from considering a random channel selection, as well as, a traditional SSF AP selection, whereas in blue are depicted the results obtained from the joint operation of DCA and DAPS.

Although network densification can be a good solution to tackle network congestion, Figure 5.3a shows that it does not improve the performance by itself. Analyzing the throughput drop ratio in the case of 25 APs, we notice that the static approach still performs badly, as the 75th percentile of the measurements surpasses a value of 20%. This effect is associated to the fact that the such approach is very sensitive to the specific topology of each scenario. As a result, we obtain a high variability in the results, as the different whiskers on the boxplots show. Comparing the DCA and DAPs performance against the static scheme, we find that this dependency on the scenario's topology, and so the variance in the results, is highly reduced using the adaptive MABs. Hence, the benefits of using both mechanisms are very relevant, since the network is able to better manage the spectrum resources, while achieving a better performance and removing the dependency on the network topology.

Under high user density conditions, we also observe that despite the downtrend of the boxes, the difference between the static and the adaptive MABs approach gets higher, as the network gets denser. For instance, for

150 STAs the gain is around 7%, whereas for the 225 and 300 STAs the gain gets up to 11% and 16%, respectively. This effect shows us that the adaptive MABs approach is capable to support a larger number of users before downgrading significantly the performance. Again, the low variance results observed in all figures show that in a wide diversity of scenarios the output values are very constant in front of the static approach, which presents a high variability even in scenarios with few APs as indicated by the larger whiskers range. Also, it can be observed that, when applying the DCA and DAPS, agents are able to successfully learn even in complex and challenging scenarios, such as the case of having 300 STAs.

At last, we evaluate the MABs under changing conditions. To do so, we employ the same controlled scenario depicted in Figure 5.2a, triggering a sudden reconfiguration on the central AP. Figure 5.4 show the overall satisfaction evolution of the STAs. As it can be noted, at $t = 12h$, the reconfiguration is triggered, and so, all STAs' agents start to perceive negative rewards as the central AP overlaps with the other ones. However, the MABs are able to take into account such negative rewards into the decision-makin progress, allowing APs and STAS to explore new actions until reaching a new stable state. Hence, this mechanisms are provide a good solution to face network malfunctions or unexpected situations, providing the necessary mechanisms to enable a reactive network.

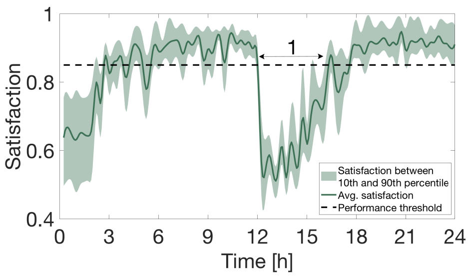
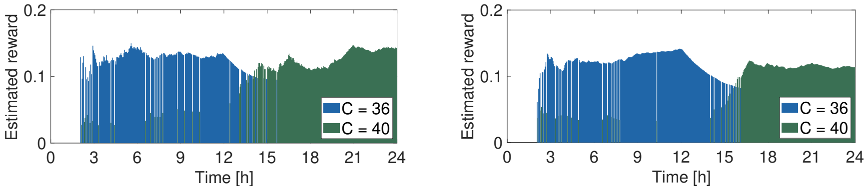


Figure 5.4: Satisfaction evolution under changing conditions

One thing to notice from Figure 5.4 is the time spent during the learning process. Indeed, one consideration to take into account when using learning algorithms is the age of information (AOI) as it has direct impact on the time spent during the learning phase. It is important to identify the existing trade-off between still valid and outdated information. In very dy-



(a) Sliding window interval 60s

(b) Sliding window interval 540s

Figure 5.5: Sliding window effect over MABs

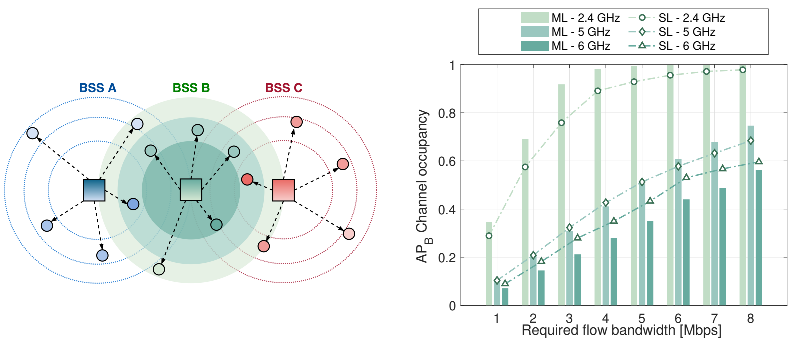
namic scenarios, keeping track of old observations can lead agents to take decisions based on information that is outdated. However, not considering enough past data will reduce the ability to select a proper new action, as agents may lose useful information. To tackle this trade-off, we used the concept of sliding window, which is intended to filter the useful information from the outdated one. For such purpose, we evaluated different window sizes, realizing that small windows sizes are able to react fast to changes from other players, since the reward evaluation is only averaged over a small set of reward entries. However, it makes agents to become more vulnerable to others' decisions, and any random exploration by an agent may lead to an action change in all the others. Indeed, this issue can be corroborated on Figure 5.5a as frequent explorations are performed even after reaching a convergence state at $t = 6$ h.

On the contrary, larger windows help to control and minimize the impact of the other agents on its own behavior, as it can be seen in Figure 5.5b, in which smoother transitions reveal that agents are more robust against sporadic changes. Although having a larger window size helps out agents to overcome the case of intermittent bad performances, it costs agent reaction time. We refer to agents' ability to detect and avoid an action that has been repeatedly performing bad. Therefore, setting a conservative approach in order to provide robustness to agents may lead to unfeasible large reconfiguration times. This issue is shown in Figure 5.5b, in which agents require much more time to start exploring new actions, and finally changing to another one.

5.2 Traffic management in MLO enabled IEEE 802.11be WLANs

In the quest to provide a good resource management for MLO, we assessed the different flow-based policies presented in Section 3.3.2. However, first we tackle the need for a traffic manager, in order to provide a clear perspective and justifying its application. If considering the scenario depicted Figure 5.6a, we observe that there are 3 AP MLDs, which have three different enabled interfaces. The inner ring correspond to the 6 GHz interface, while the outer ones correspond to the 5 GHz and 2.4 GHz bands, respectively. Figure 5.6b shows the channel occupancy observed by AP_B when considering an equally distributed MLO traffic allocation, and when assessing a Multi-Band Single Link (MB-SL) deployment.

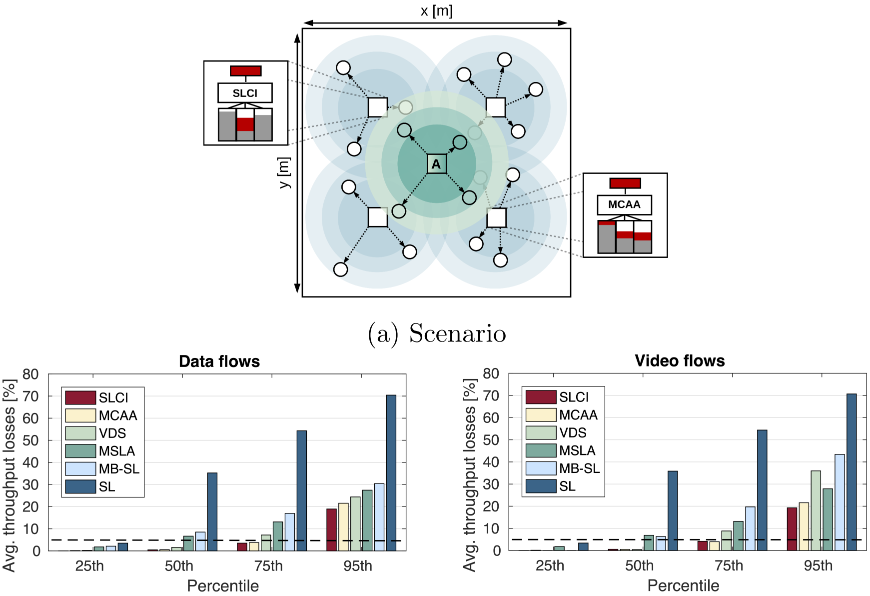
As it can be observed, the channel occupancy values under a MLO feature, resembles the ones obtained when the legacy MB-SL is enabled. Thus, there is no potential gain on the MLO application, as it does not leverage the fact of having two downlink contention-free links (i.e., only on the 2.4GHz link neighboring BSSs are detected). That is, there is no difference between spreading the same proportion the traffic over the different links through MLO, than balancing and/or steering clients across links, like current MB-SL does. We observe, then, that without a proper traffic allocation policy, MLO underperforms.



(a) Scenario representation (b) Channel occupancy

Figure 5.6: Comparison between legacy SL and MLO

Under the aforementioned analysis, the majority of the proposed policies in Section 3.3.2 are designed taking into account the channel occupancy at each interface, in order to provide with a congestion-aware implementation on the traffic manager. That is, before each allocation, the instantaneous occupancy is gathered, and then, traffic is allocated accordingly. A comparison between the different MLO policies, and the legacy single-link (SL) and MB-SL is provided in Figure 5.7b and Figure5.7c. The scenario considered to obtain such results is depicted in Figure 5.7a, in which a central AP applies one of the policies, whereas the neighboring APs select either the MCAA or SLCI with the same probability.



(a) Scenario

(b) Histogram for Data flows

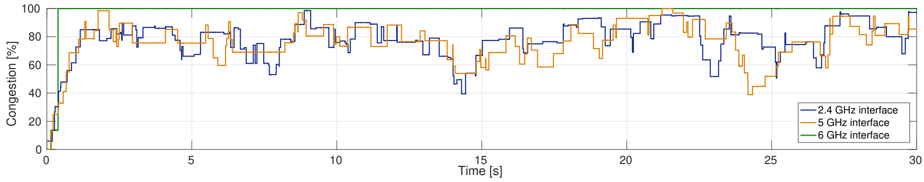
(c) Histogram for Video flows

Figure 5.7: Comparison between policies and legacy operations

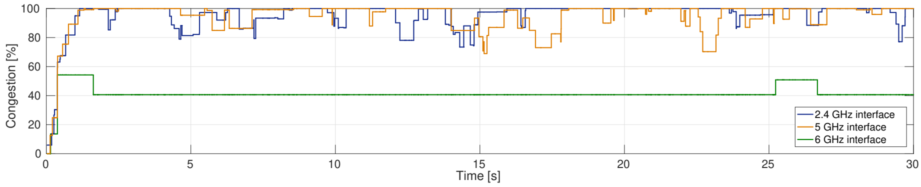
All the congestion-aware policies are able to provide much better results under the same conditions than the legacy approaches. Specially, we observe that the SLCI and MCAA congestion-aware policies are able to overcome such negative performance in 75% of the scenarios because their ability to balance the traffic between all the active links. Also, it is noticeable for video flows that when using the VDS, the 5% worst case raises up to throughput losses nearly 40%, performing even worse than the MSLA. Such results reveal a critical drawback of the VDS policy: the traffic separation in VDS may suffer from severe performance problems in conditions with high number of neighboring BSSs overlaps, or high traffic scenarios.

Comparing legacy approaches, we observe the advantage of adding more bands to the system through the MB-SL, as the stations can be spread across them, reducing also their congestion levels if compared to the SL. Although the MB-SL performance is better, it barely keeps the throughput losses below an acceptable 5% value for both flow types in only 25% of the scenarios. Also, compared to the legacy approaches, MLO is shown to be able to perform better in all the evaluated scenarios independently. In fact, it is noticeable that with SL and MB-SL only the 25% of the considered scenarios achieve average throughput losses below 5%, which is increased up to the 75% with either SLCI and MCAA. Those results, prove that the MLO framework will be a relevant new feature to Wi-Fi, enabling currently unsuitable scenarios with SL and MB-SL.

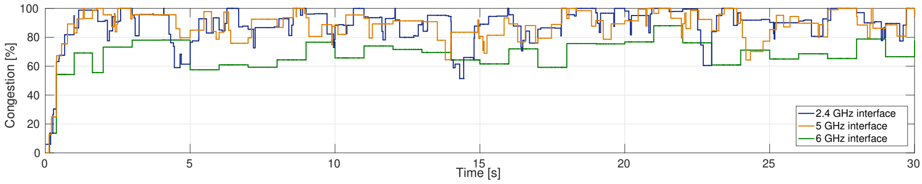
To assess the problems observed with video flows, we considered the implementation of a dynamic policy, which is able to minimize the principal constraint of the non-dynamic ones: the static allocation upon a flow arrival. To exemplify such issue we provide a detailed insight on the channel occupancy evolution for each AP_A's interface during the first 30 s. Figure 5.8a and Figure 5.8b expose the main drawbacks of SLCI and MCAA, respectively, as the temporal evolution of the congestion reveals how unbalanced the interfaces are. First, the SLCI overloads the 6 GHz link by placing the whole video flow in it, while there is still room for some traffic in the other interfaces. On the contrary, the MCAA does not leverage the fact of having empty space at the 6 GHz interface, which makes the proportional parts of the flow allocated to the 2.4 GHz and 5 GHz links to suffer from congestion. Such inefficient operation from the non-dynamic policies is shown in Figure 5.8c to be overcomed by the MCAB, as it reveals a more balanced use of the interfaces. However, we also observe that most of the time the congestion values for the 6 GHz interface are lower than for the other two. Such effect is related to the unequal number of neighboring nodes detected at each band. As a result, even if most of the traffic is allocated to this interface, it still manages to



(a) SCLI



(b) MCAA



(c) MCAB

Figure 5.8: Congestion distribution per interface, and per policy type

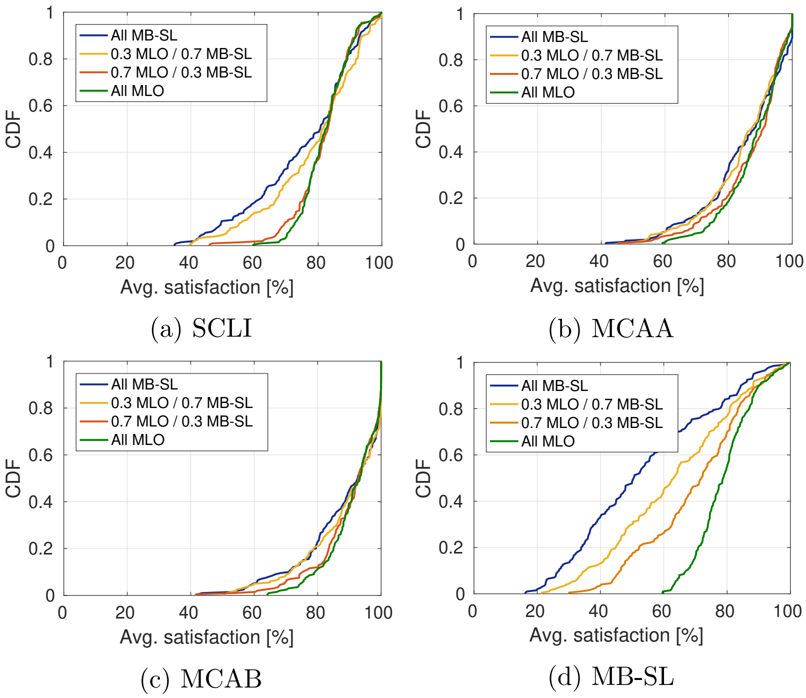
provide traffic with fewer congestion episodes.

At last, Wi-Fi's constant evolution makes newer devices, which implement up-to-date specifications, to coexist with others with less capabilities. As a result, last generation devices may decay in performance due to its coexistence with legacy ones. To assess if MB-SL BSSs affects the performance of MLO ones, we analyze four different cases in which we increment the fraction of MLO BSSs around the central one from 0, to 0.3, 0.7, and 1.

Figures 5.9a, 5.9b, and 5.9c show the CDF of the flow average satisfaction obtained when using each policy. Regardless of the policy used, the central BSS_A experiences a negative trend when it is surrounded by more legacy BSSs, as the results show lower satisfaction values when so. Although the MCAA and MCAB experience low gains when increasing the number of MLO BSSs, the SLCI presents a 17% improvement for the

25th percentile, when comparing the performance results between the best and the worst (i.e., all MLO and all MB-SL, respectively) cases. Such an improvement is caused by the higher link availability from the neighboring BSSs to allocate traffic, which also avoid to overload the interfaces by the use of congestion-aware policies.

At last, Figure 5.9d shows, as well, the average satisfaction when BSS_A is set as a legacy MB-SL with the aim to observe if the presence of MLO devices will benefit legacy ones. As previously, we incremented the fraction of MLO BSSs from 0, to 0.3, 0.7, and 1. Figure 5.9a reveals that legacy MB-SL BSSs can benefit from the fact of having MLO BSSs around them, as the improvement is highly noticeable. In fact, we observe that between the best and worst cases the satisfaction increases by a 40% for half of the scenarios evaluated. Then, from the perspective of a legacy BSS, the adoption of the MLO represents also a performance improvement for itself.



(a) SCLI (b) MCAA

(c) MCAB (d) MB-SL

Figure 5.9: Coexistence performance per policy type, and MB-SL.

5.3 Summary

As a summary of this Chapter, we point out the main findings obtained from our work

- The MABs framework constitute a great candidate to enable future HD deployments with reactive mechanisms to solve unexpected situations or system malfunctions. Also, it avoids any topology dependency (i.e., spatial distribution), allowing different scenarios currently unsupported by static approaches.
- The joint operation of the DCA and DAPS has been proven to provide with great results, even in complex scenarios with high number of users. Additionally, under unbalanced situations, DAPS mechanism has been found to more useful, as it allows to alleviate the most congested APs keeping up with the network performance.
- The correct characterization of the AOI plays a key role on the learning procedure of the agents. Hence, the trade-off between smaller o larger times depends on the taxonomy of the problem, as well as, the dynamic conditions of the scenario.
- Although MLO is a promising feature behind the 802.11be amendment, the addition of a traffic manager within the upper MAC layer can improve even further the potential gains of such feature. Indeed, we observed that dummy traffic allocation techniques may perform equally as existing MB-SL techniques leading MLO to underperform.
- Congestion-aware policies are basic to provide with an efficient traffic management, as the knowledge on the traffic conditions allows to provide with an efficient mechanism to allocate traffic. However, we observe that the SLCI provides a more robust performance under HD conditions, as the other policies tend to suffer from actions taken by the surrounding BSSs.
- The adoption of MLO will become highly relevant to Wi-Fi, enabling unsuitable scenarios of nowadays standard. In this context, coexistence issues may not represent a critical concern, as the registered results show low degradation on the satisfaction. Then, it

is found that the higher link availability for 802.11be capable devices, jointly with a congestion-aware mechanism, should minimize the performance issues on both types of nodes.

Chapter 6

CONCLUDING REMARKS

In this thesis, we conducted an investigation on different solutions that are likely to be included under the next generation of Wi-Fi networks. For that purpose, we first provided an overview on the basic IEEE 802.11 operation, pointing out that the uncoordinated nature of Wi-Fi may be difficult to manage over dense scenarios. Later, we analyzed the performance challenges that may constrain Wi-Fi deployments over HD scenarios, providing different literature on those topics. With that, we presented the research directions taken by the academia in order to solve such problems. However, we noticed that static approaches may not become feasible to be implemented, as the high and non-stationarity dynamics of the wireless medium need for elastic and adaptable underlying systems. A new Wi-Fi ecosystem, then, poses as solution to face either unplanned network configurations or an unbalanced user distribution. In this context, we find that new softwarized architectures will provide the sufficient flexibility to react to changes, while intelligent management techniques will know how to react to them. Also, in regards of throughput provisioning, the standard continues its evolution with the proposal of new features. Under this thesis, we evaluated the MLO observing that such feature will become highly relevant in a near-future.

The main findings of our work confirm that the intelligent MAB approach of the Dynamic CA (DCA) and Dynamic AP Selection (DAPS) is a very feasible solution to enhance network resource management processes, while having the ability to combat unexpected situations. Under an adversarial setting, which may be typically found in home scenarios, we can conclude that the performance of MABs overcomes different limitations of traditional static approaches, removing, for instance, their high dependency on the scenario conditions. In such context, not only throughput improvements have been observed, but the capacity to support more users. Besides, the trade-off between outdated and useful data need to be taken into consideration, as the results show that depending on the taxonomy of the problem (i.e., scenario considered), accounting for high amounts of data can lead to slow reaction times.

In regards of the MLO evaluation, it has been found that MLO's performance is tied to the selected traffic allocation strategy. For instance, we observed that a dummy strategy may cause to equally perform, or even underperform, the current MB-SL solution. In this context, we assessed different traffic allocation policies based on a congestion-aware implementation, which resulted to be the most effective when coping with different levels of network congestion. However, we found that their static approach made allocation decisions highly vulnerable to the different actions taken by neighboring BSSs. Specifically, for long-lasting flows, such condition was exacerbated. Hence, we implemented a dynamic policy to overcome such issue, observing that long-lasting flows' performance was improved through a better load distribution over the available resources. At last, we assessed coexistence issues, showing that the greater link-availability from MLO devices, as well as a proper allocation strategy, eases the congestion perceived by legacy nodes, improving their performance.

We left as future work the potential integration of MABs over MLO, in order to study the viability to allocate traffic over each interface through intelligent decisions. In this context, such research line seems to have high potential, as traffic allocations can be autonomously adapted to the variations on the network congestion. Additionally, we also consider as future work the redesign of the traffic management module as part of an end-to-end SDN solution, closely working with an external controller to properly allocate traffic flows to interfaces in a multi-AP WLAN. In this sense, evaluating centralized multi-AP MLO policies could even provide more relevant insights.

Funding Sources

This work has been partially supported by:

- The Spanish Ministry of Economy, Industry and Competitiveness, under the Maria de Maeztu Units of Excellence Programme (MDM-2015-0502).
- The Spanish Government under grant WINDMAL PGC2018-099959-B-I00 (MCIU/AEI/FEDER,UE)
- The Catalan Government under grant 2017-SGR-1188.

Bibliography

[1] Cisco Visual Networking. Cisco visual networking index: Global mobile data traffic forecast update, 2017–2022. *Cisco white paper*, 2019.

[2] IEEE P802.11be/D1.0 Draft Standard for Information technology— Telecommunications and information exchange between systems Local and metropolitan area networks— Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 8: Enhancements for extremely high throughput (EHT), May 2021.

[3] Feilu Liu, Jian Lin, Zhifeng Tao, Thanasis Korakis, Elza Erkip, and Shivendra Panwar. The hidden cost of hidden terminals. In *2010 IEEE International Conference on Communications*, pages 1–6. IEEE, 2010.

[4] Zhenzhe Zhong, Parag Kulkarni, Fengming Cao, Zhong Fan, and Simon Armour. Issues and challenges in dense wifi networks. In *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 947–951. IEEE, 2015.

[5] Boris Bellalta. Ieee 802.11 ax: High-efficiency wlans. *IEEE Wireless Communications*, 23(1):38–46, 2016.

[6] Martin Heusse, Franck Rousseau, Gilles Berger-Sabbatel, and Andrzej Duda. Performance anomaly of 802.11 b. In *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, volume 2, pages 836–843. IEEE, 2003.

[7] Robert Akl and Anurag Arepally. Dynamic channel assignment in ieee 802.11 networks. In *2007 IEEE international conference on portable information devices*, pages 1–5. IEEE, 2007.

[8] Hui Luo and NK Shankaranarayanan. A distributed dynamic channel allocation technique for throughput improvement in a dense wlan environment. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages V–345. IEEE, 2004.

[9] Alex Hills. Large-scale wireless lan design. *IEEE Communications Magazine*, 39(11):98–107, 2001.

[10] Andreas Eisenblatter, Hans-Florian Geerdes, and Iana Siomina. Integrated access point placement and channel assignment for wireless lans in an indoor office environment. In *2007 IEEE international symposium on a world of wireless, mobile and multimedia networks*, pages 1–10. IEEE, 2007.

[11] Murali Achanta. Method and apparatus for least congested channel scan for wireless access points, April 6 2006. US Patent App. 10/959,446.

[12] Ming Yu, Hui Luo, and Kin K Leung. A dynamic radio resource management technique for multiple aps in wlans. *IEEE Transactions on Wireless Communications*, 5(7):1910–1919, 2006.

[13] Marko Hoyhtya, Sofie Pollin, and Aarne Mammela. Classification-based predictive channel selection for cognitive radios. In *2010 IEEE International Conference on Communications*, pages 1–6. IEEE, 2010.

[14] Dusit Niyato and Ekram Hossain. Cognitive radio for next-generation wireless networks: An approach to opportunistic channel selection in ieee 802.11-based wireless mesh. *IEEE Wireless Communications*, 16(1):46–54, 2009.

[15] Yuhua Xu, Alagan Anpalagan, Qihui Wu, Liang Shen, Zhan Gao, and Jinglong Wang. Decision-theoretic distributed channel selection for opportunistic spectrum access: Strategies, challenges and solutions. *IEEE Communications Surveys & Tutorials*, 15(4):1689–1713, 2013.

[16] Arunesh Mishra, Vivek Shrivastava, Dheeraj Agrawal, Suman Banerjee, and Samrat Ganguly. Distributed channel management in uncoordinated wireless environments. In *Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 170–181. ACM, 2006.

[17] Olivier Jeunen, Patrick Bosch, Michiel Van Herwegen, Karel Van Doorselaer, Nick Godman, and Steven Latré. A machine learning approach for ieee 802.11 channel allocation. In *2018 14th International Conference on Network and Service Management (CNSM)*, pages 28–36. IEEE, 2018.

[18] Sergio Barrachina-Muñoz, Francesc Wilhelmi, and Boris Bellalta. Online primary channel selection for dynamic channel bonding in high-density wlans. *IEEE Wireless Communications Letters*, 2019.

[19] Wang Lingzhi, Hua Cunqing, Zheng Rong, and Ni Rui. Online channel selection and user association in high-density wifi networks. In *2015 IEEE International Conference on Communications (ICC)*, pages 1571–1576. IEEE, 2015.

[20] Francesc Wilhelmi, Cristina Cano, Gergely Neu, Boris Bellalta, Anders Jonsson, and Sergio Barrachina-Muñoz. Collaborative spatial reuse in wireless networks via selfish multi-armed bandits. *Ad Hoc Networks*, 88:129–141, 2019.

[21] Vivek P Mhatre, Konstantina Papagiannaki, and Francois Baccelli. Interference mitigation through power control in high density 802.11 wlans. In *IEEE INFOCOM 2007-26th IEEE International Conference on Computer Communications*, pages 535–543. IEEE, 2007.

[22] Aditya Akella, Glenn Judd, Srinivasan Seshan, and Peter Steenkiste. Self-management in chaotic wireless deployments. In *Proceedings of the 11th annual international conference on Mobile computing and networking*, pages 185–199, 2005.

[23] Byoung Hoon Jung, Hu Jin, and Dan Keun Sung. Adaptive transmission power control and rate selection scheme for maximizing energy efficiency of ieee 802.11 stations. In *2012 IEEE 23rd International*

Symposium on Personal, Indoor and Mobile Radio Communications-(PIMRC), pages 266–271. IEEE, 2012.

[24] Francesc Wilhelmi, Sergio Barrachina-Muñoz, Cristina Cano, Ioannis Selinis, and Boris Bellalta. Spatial reuse in ieee 802.11 ax wlans. *Computer Communications*, 170:65–83, 2021.

[25] Francesc Wilhelmi, Sergio Barrachina-Muñoz, and Boris Bellalta. On the performance of the spatial reuse operation in ieee 802.11 ax wlans. In *2019 IEEE Conference on Standards for Communications and Networking (CSCN)*, pages 1–6. IEEE, 2019.

[26] Heyi Tang, Liu Yang, Jiang Dong, Zhonghong Ou, Yong Cui, and Jianping Wu. Throughput optimization via association control in wireless lans. *Mobile Networks and Applications*, 21(3):453–466, 2016.

[27] Phillip B Oni and Steven D Blostein. Decentralized ap selection in large-scale wireless lans considering multi-ap interference. In *2017 International conference on computing, networking and communications (ICNC)*, pages 13–18. IEEE, 2017.

[28] Issam Jabri, Nicolas Krommenacker, Thierry Divoux, and Adel Soudani. Ieee 802.11 load balancing: an approach for qos enhancement. *International Journal of Wireless Information Networks*, 15(1):16–30, 2008.

[29] Yigal Bejerano, Seung-Jae Han, and Li Li. Fairness and load balancing in wireless lans using association control. In *Proceedings of the 10th annual international conference on Mobile computing and networking*, pages 315–329, 2004.

[30] Dawei Gong and Yuanyuan Yang. On-line ap association algorithms for 802.11 n wlans with heterogeneous clients. *IEEE Transactions on Computers*, 63(11):2772–2786, 2013.

[31] Tao Lei, Xiangming Wen, Zhaoming Lu, and Yue Li. A semi-matching based load balancing scheme for dense ieee 802.11 wlans. *IEEE Access*, 5:15332–15339, 2017.

[32] Eduard Garcia, Rafael Vidal, and Josep Paradells. Cooperative load balancing in ieee 802.11 networks with cell breathing. In *2008 IEEE Symposium on Computers and Communications*, pages 1133–1140. IEEE, 2008.

[33] Chia-Ying Lin, Wan-Ping Tsai, Meng-Hsun Tsai, and Yun-Zhan Cai. Adaptive load-balancing scheme through wireless sdn-based association control. In *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)*, pages 546–553. IEEE, 2017.

[34] Estefania Coronado, Jose Villalon, and Antonio Garrido. Wi–balance: Sdn–based load–balancing in enterprise wlans. In *2017 IEEE Conference on Network Softwarization (NetSoft)*, pages 1–2. ieee, 2017.

[35] Nahida Kiran, Yin Changchuan, and Zaid Akram. Ap load balance based handover in software defined wifi systems. In *2016 IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, pages 6–11. IEEE, 2016.

[36] Israat Tanzeena Haque and Nael Abu-Ghazaleh. Wireless software defined networking: A survey and taxonomy. *IEEE Communications Surveys & Tutorials*, 18(4):2713–2737, 2016.

[37] Nachikethas A Jagadeesan and Bhaskar Krishnamachari. Software-defined networking paradigms in wireless networks: A survey. *ACM Computing Surveys (CSUR)*, 47(2):1–11, 2014.

[38] Rashid Mijumbi, Joan Serrat, Juan-Luis Gorricho, Niels Bouten, Filip De Turck, and Raouf Boutaba. Network function virtualization: State-of-the-art and research challenges. *IEEE Communications surveys & tutorials*, 18(1):236–262, 2015.

[39] Jon Matias, Jokin Garay, Nerea Toledo, Juanjo Unzilla, and Eduardo Jacob. Toward an sdn-enabled nfv architecture. *IEEE Communications Magazine*, 53(4):187–193, 2015.

[40] Junfeng Xie, F Richard Yu, Tao Huang, Renchao Xie, Jiang Liu, Chenmeng Wang, and Yunjie Liu. A survey of machine learning

techniques applied to software defined networking (sdn): Research issues and challenges. *IEEE Communications Surveys & Tutorials*, 21(1):393–430, 2018.

[41] Raja Karmakar, Samiran Chattopadhyay, and Sandip Chakraborty. A deep probabilistic control machinery for auto-configuration of wifi link parameters. *IEEE Transactions on Wireless Communications*, 19(12):8330–8340, 2020.

[42] Juan Sánchez-González, Jordi Pérez-Romero, Ramon Agustí, and Oriol Sallent. On learning mobility patterns in cellular networks. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pages 686–696. Springer, 2016.

[43] David D Clark, Craig Partridge, J Christopher Ramming, and John T Wroclawski. A knowledge plane for the internet. In *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 3–10. ACM, 2003.

[44] Szymon Szott, Katarzyna Kosek-Szott, Piotr Gawłowicz, Jorge Torres Gómez, Boris Bellalta, Anatolij Zubow, and Falko Dressler. Wifi meets ml: A survey on improving ieee 802.11 performance with machine learning. *arXiv preprint arXiv:2109.04786*, 2021.

[45] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. Overview of supervised learning. In *The elements of statistical learning*, pages 9–41. Springer, 2009.

[46] Pádraig Cunningham, Matthieu Cord, and Sarah Jane Delany. Supervised learning. In *Machine learning techniques for multimedia*, pages 21–49. Springer, 2008.

[47] Zoubin Ghahramani. Unsupervised learning. In *Summer school on machine learning*, pages 72–112. Springer, 2003.

[48] Artúr István Károly, Róbert Fullér, and Péter Galambos. Unsupervised clustering for deep learning: A tutorial survey. *Acta Polytechnica Hungarica*, 15(8):29–53, 2018.

[49] Mance E Harmon and Stephanie S Harmon. Reinforcement learning: A tutorial. 1997.

[50] Abhijit Gosavi. Reinforcement learning: A tutorial survey and recent advances. *INFORMS Journal on Computing*, 21(2):178–192, 2009.

[51] Shuyan Hu, Xiaojing Chen, Wei Ni, Ekram Hossain, and Xin Wang. Distributed machine learning for wireless communication networks: Techniques, architectures, and applications. *IEEE Communications Surveys & Tutorials*, 23(3):1458–1493, 2021.

[52] Aleksandrs Slivkins et al. Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning*, 12(1-2):1–286, 2019.

[53] Omar Besbes, Yonatan Gur, and Assaf Zeevi. Stochastic multi-armed-bandit problem with non-stationary rewards. *Advances in neural information processing systems*, 27, 2014.

[54] Peter Auer and Ronald Ortner. Ucb revisited: Improved regret bounds for the stochastic multi-armed bandit problem. *Periodica Mathematica Hungarica*, 61(1-2):55–65, 2010.

[55] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of IEEE 36th annual foundations of computer science*, pages 322–331. IEEE, 1995.

[56] Djallel Bouneffouf, Irina Rish, and Charu Aggarwal. Survey on applications of multi-armed and contextual bandits. In *2020 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE, 2020.

[57] Shipra Agrawal and Navin Goyal. Further optimal regret bounds for thompson sampling. In *Artificial intelligence and statistics*, pages 99–107, 2013.

[58] William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.

[59] Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.

[60] David López-Pérez, Adrian Garcia-Rodriguez, Lorenzo Galati-Giordano, Mika Kasslin, and Klaus Doppler. Ieee 802.11 be extremely high throughput: The next generation of wi-fi technology beyond 802.11 ax. *IEEE Communications Magazine*, 57(9):113–119, 2019.

[61] Marc Carrascosa-Zamacois, Lorenzo Galati-Giordano, Anders Jonsson, Giovanni Geraci, and Boris Bellalta. Performance and coexistence evaluation of ieee 802.11 be multi-link operation. *arXiv preprint arXiv:2205.15065*, 2022.

[62] Abhishek Patil, George Cherian, Alfred Asterjadhi, and Duncan Ho. Multi-link operation: Design discussion. IEEE 802.11 Documents. 2019. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=823&is_group=00be`.

[63] Christoph Paasch and Olivier Bonaventure. Multipath tcp. *Communications of the ACM*, 57(4):51–57, 2014.

[64] Sébastien Barré, Christoph Paasch, and Olivier Bonaventure. Multipath tcp: from theory to practice. In *International conference on research in networking*, pages 444–457. Springer, 2011.

[65] Sharan Naribole, Wook Bong Lee, Srinivas Kandala, and Ashok Ranganath. Simultaneous transmit-receive multi-channel operation in next generation wlans. In *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–8. IEEE, 2020.

[66] Mark Hamilton and Joseph Levy. Mld architecture part 2. IEEE 802.11 Documents. 2021. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1111&is_group=00be`.

[67] Toni Adame, Marc Carrascosa-Zamacois, and Boris Bellalta. Time-sensitive networking in ieee 802.11 be: On the way to low-latency wifi 7. *Sensors*, 21(15):4954, 2021.

[68] Guillermo Lacalle, Iñaki Val, Oscar Seijo, Mikel Mendicute, Dave Cavalcanti, and Javier Perez-Ramirez. Analysis of Latency and Reliability Improvement with Multi-Link Operation over 802.11. In *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, pages 1–7. IEEE, 2021.

[69] S Ben Fred, Thomas Bonald, Alexandre Proutiere, Gwénaël Régnié, and James W Roberts. Statistical bandwidth sharing: a study of congestion at flow level. *ACM SIGCOMM Computer Communication Review*, 31(4):111–122, 2001.

[70] Ramji Venkataramanan, Min wook Jeong, and Balaji Prabhakar. A flow- and packet-level model of the internet.

[71] Frank Delcoigne, Alexandre Proutiere, and Gwénaël Régnié. Modeling integration of streaming and data traffic. *Performance Evaluation*, 55(3-4):185–209, 2004.

[72] Gilbert Chen and Boleslaw K Szymanski. Cost: A component-oriented discrete event simulator. In *Proceedings of the Winter Simulation Conference*, volume 1, pages 776–782. IEEE, 2002.

[73] George F Riley and Thomas R Henderson. The ns-3 network simulator. In *Modeling and tools for network simulation*, pages 15–34. Springer, 2010.

[74] Giuseppe Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *IEEE Journal on selected areas in communications*, 18(3):535–547, 2000.

Chapter 7

PUBLICATIONS

Combining Software Defined Networks and Machine Learning to enable Self Organizing WLANs

Álvaro López-Raventós, Francesc Wilhelmi, Sergio Barrachina-Muñóz and Boris Bellalta

Abstract

Next generation of wireless local area networks (WLANs) will operate in dense, chaotic and highly dynamic scenarios that in a significant number of cases may result in a low user experience due to uncontrolled high interference levels. Flexible network architectures, such as the software-defined networking (SDN) paradigm, will provide WLANs with new capabilities to deal with users' demands, while achieving greater levels of efficiency and flexibility in those complex scenarios. On top of SDN, the use of machine learning (ML) techniques may improve network resource usage and management by identifying feasible configurations through learning. ML techniques can drive WLANs to reach optimal working points by means of parameter adjustment, in order to cope with different network requirements and policies, as well as with the dynamic conditions. In this paper we overview the work done in SDN for WLANs, as well as the pioneering works considering ML for WLAN optimization. Finally, in order to demonstrate the potential of ML techniques in combination with SDN to improve the network operation, we evaluate different use cases for intelligent-based spatial reuse and dynamic channel bonding operation in WLANs using Multi-Armed Bandits.

1 Introduction

In recent years, IEEE 802.11-based WLANs, commonly known as Wi-Fi networks, have experienced a remarkable growth in terms of traffic consumption. According to [1], in 2016 more traffic was offloaded from cellular networks onto Wi-Fi than remained on them. Moreover, they expect by 2021 that the 63 % of total mobile data traffic will be offloaded onto Wi-Fi network as a consequence of an increased use of portable and handheld devices. In this context, network capacity needs to be targeted to cope with the expected data traffic. Thus, efforts are focused in network densification as the spectrum scarcity and the high spectral efficiency achieved by current technologies are limiting factors [2].

Regarding dense deployments, there exist some potential issues in regards of performance degradation. Existing channel access protocols, such as carrier sense multiple access (CSMA), have been designed to operate efficiently in non-dense scenarios, and they may become a bottleneck when pushed further. In dense WLANs, due to the great number of contending nodes using CSMA, we could find three well-known performance issues. We refer to the hidden and exposed node problems, and to the flow starvation. In terms of performance, the appearance of these issues can cause a remarkable degradation of the experienced throughput due to different factors, such as a large number of collisions or wasting useful time slots. Moreover, some solutions like request-to-send and clear-to-send (RTS/CTS) mechanisms that are intended to avoid the hidden and exposed node problems, can lead to an excessive control packet overhead, which may negatively affect the overall performance, too. Apart from the above-mentioned, other concerns are related with chaotic deployments since they lead to have an excessive co-channel and adjacent channel interference (CCI/ACI) levels, directly caused by the lack of frequency planning and inefficient power configuration choices.

To cope with the aforementioned challenges, the software-defined networking (SDN) paradigm can be applied to Wi-Fi networks in order to enable a more efficient and flexible network control and management. The main concept behind SDN is that it proposes to decouple the control and the data planes into different layers, with a central controller performing configuration changes with a global view of the network state. As a result, control processes are removed from forwarding devices, which stand as simple programmable nodes that directly depend on the controller's instructions. In consequence, networks can be adjusted dynamically according to the knowledge extracted from statistics, which are collected at the central entity. This specific characteristic of SDN is quite relevant for wireless environments due to their non-stationary conditions (i.e., users moving, diverse traffic requirements and changing channel conditions). Having a dynamic and centralized control design, the overall performance of the network can be improved and interferences, unbalanced situations or system failures mitigated. In this regard, network management and data analytics play a key role in order to increase network efficiency. For instance, network information such as the signal-to-interference-plus-noise ratio (SINR), the received signal strength indicator (RSSI), the total number of active users and throughput rates can be easily collected. Thus, network optimization needs to exploit this useful data and the use of learning algorithms can lead this processes. This envision open up new research directions and so, we focus our studies in the joint integration of machine learning and SDN for wireless optimizations.

The rest of the paper is organised as follows: in Section 2, we present a general overview about the SDN paradigm, reviewing current implementations of the SDN architecture into wireless networks. Then, we point out different features to be taken as future research directions. After, in Section 3, we discuss an architecture involving wireless SDN and ML solutions, together with an overview of different management functionalities. Later, in Section 4, we perform a proof of concept with the aim to demonstrate how ML can enable self-organizing WLANs. Through different use cases, we evaluate the usage of ML over SDN-controlled WLANs with the aim to to find the best configuration according to a Max-min fairness policy. To do so, we exploit the Multi-Armed Bandits (MABs) framework to empower a collaborative behavior between them. Finally, conclusions are stated in Section 5.

2 SDN: From wired to wireless

2.1 SDWN through SDN

SDN is a novel network architecture paradigm that is dynamic, manageable, cost-effective and adaptable. Moreover, SDN decouples network control and forwarding functions into different planes, allowing the underlying infrastructure to be abstracted from application and network services. In consequence, unlike distributed architectures, in which forwarding devices listen for events from their neighbors and make decisions based on a local view, the network infrastructure (i.e., switches and routers) just act as packet forwarding devices. In addition, SDN empowers programmability and network function virtualization (NFV) at the controller, allowing network administrators to have flexibility and a fine-grained control over the entire network. Thus, SDN reduces capital and operational expenditures (CapEX and OpEX, respectively), while enabling inno-

Table 1: Taxonomy of the related work presented

Name	Application development	Southbound communication	VAP / LVAP	Separated MAC	End-user modification
OpenRoads	✓	OpenFlow + SNMP	✓	✗	✗
Odin	✓	OpenFlow + Proprietary	✓	✓	✗
CloudMAC	✓	OpenFlow	✓	✓	✗
Ætherflow	✓	Extended OpenFlow	✗	✗	✗
COAP	✓	Extended OpenFlow	✗	✗	✗
Ethanol	✓	OpenFlow + Proprietary	✓	✓	✗
Aeroflux	✓	OpenFlow + Proprietary	✓	✓	✗
OpenSDWN	✓	OpenFlow + REST	✓	✗	✗
BeHop	✓	OpenFlow + Proprietary	✓	✗	✗
EmPOWER	✓	OpenFlow + Proprietary	✓	✗	✗

vation.

Typically, the SDN architecture is divided into three different layers, which can be found in literature as infrastructure, control and application layers. The first one contains the different network elements that follow the rules provided by the controller. The second one involves the controller, which is in charge of configuring the devices as well as to the different services. The last one contains the network applications which define the different policies to be applied over the network. Communication between layers is done by means of the northbound and southbound interfaces. The former is based on APIs (e.g. REST) that are intended to application development, while the latter is based on standard protocols such as OpenFlow, Simple Network Management Protocol (SNMP) or Control And Provisioning of Wireless Access Points (CAPWAP). However, none of this protocols are intended to wireless communications and therefore, as currently defined, they cannot control layer 2 traffic over wireless networks nor report measurements of the wireless medium. To overcome with that issue, modifications by means of extending the current protocols, or even the use of proprietary ones should be adopted to enable the control of wireless devices.

Although SDN needs to be clearly reconditioned in order to be used in wireless networking, the previously described features have pushed the trend to adopt SDN for WLANs. In this context, the concept of software-defined wireless networking (SDWN) appears with a clear aim to improve the management of wireless networks and so, SDWN has become an emerging research branch of SDN. Many publications have focused on identifying the concerns and applications of SDWN, as well as suggesting different network architectures. SDWN solutions go from extending the OF protocol with new messages, to the implementation of applications on top of OF controllers that have their own proprietary control messages. Next, in 2.2, we review different architecture solutions proposed for SDWN.

2.2 Overview of proposals for SDWN

To begin with, OpenRoads [3] was the first project focusing in SDN for WLANs. Moreover, it also introduced a testbed to control mobility between Wi-Fi and WiMax base stations. OpenRoads consists on a three layer-based architecture that is divided into physical, slicing and control layer. The physical layer is made of all the devices that are OF-enabled. The control layer is in charge of network orchestration and device configuration. Finally, the slicing layer intercepts OF protocol messages to support the slicing layer according to the network administrator policy. Thus, different network administrators can operate over the same physical network, since the slicing layer divides it into multiple logical networks. From here, other solutions such as Odin [4] came up. The Odin's architecture is composed by an Odin master (running on the OpenFlow controller), and an Odin agent (running on the APs). The Odin master communicates to the switches and the APs by means of the OpenFlow protocol, in order to control the wired connections, whereas it uses a custom protocol to communicate to each Odin agent, with the aim of collecting different network statistics (i.e., RSSI, SINR, etc.). As a result, the network is able to manage mobility, load balancing and interference in wireless connections. In addition, time-critical operations (e.g. ACKs) are performed by the APs, and non-time-critical operations are handled by the controller. Regarding client-AP association, Odin implements the concept of logical virtual access point (LVAP), which are client-specific. So, each user receives a unique BSSID to be connected to. This implementation allows client isolation as well as performing a hand-off process without triggering any re-association mechanism, since LVAPs can be removed from one AP and transferred to another. However, the hand-off in Odin is still performed based only on the RSSI, which could lead to load imbalance situations. Similar to Odin, OpenSDWN [5] is a framework that introduces a more detailed wireless data-path transmission control, enabling user-service differentiation by identifying and classifying flow types. To do so, OpenSDWN uses per-client middle-boxes, called virtual middle-boxes (vMB), that can be migrated from one AP to another. Therefore, network functionalities are migrated to destination APs as the user performs a hand-off. From Odin, OpenSDWN inherits Odin's LVAPs concept as well as the mobility method and user isolation. Later on, BeHop [6] and Ethanol [7] appeared as other solutions in the SDWN context, which took the same basis as Odin. First, BeHop architecture consists of a central controller, a set of APs forming the data plane, and a network monitor and data collector. Each BeHop AP acts as an OpenFlow switch that contains per-client virtual APs (VAPs), and a client table to track the user information (e.g., client-VAP mapping) and the network status information (e.g., channel and power allocation). Here, the network control is performed through a BeHop own proprietary API used for channel and power allocation purposes. Moreover, through a dedicated interface, the controller is able to access the data stored in the network monitor, in order to take advantage of it and enhance network management. Regarding Ethanol, it consists of two types of devices, the controller and the Ethanol-based APs, or Ethanol agents. Ethanol uses its own proprietary code to gather link information from the APs (e.g., SINR or bit rate) in order to provide the controller with statistics for network managing. Open research directions in Ethanol aim to guarantee security and quality of service (QoS) through traffic shaping. At last, EmPOWER [8] is an SDWN programming architecture that provides a set of Python based APIs, which model the fundamentals of wireless management. The aim of this architecture is to reduce complexity by applying four abstractions, each of one addressing a different control aspect such as: the state management, resource allocation, network monitoring, and network reconfiguration. Communication between wireless terminals and the controller is done by a proprietary protocol, whereas OpenFlow is used for managing the switching operations. Regarding time-critical actions, CloudMAC [9] proposed to break down the MAC operations by offloading them into different devices. Therefore, physical APs are in charge of time-critical MAC operations, whereas virtual APs (VAPs) are in charge of MAC generation. Besides, communication between them is performed through a layer 2 tunneling. The rest of the architecture is composed by an OpenFlow switch, which is used to forward packets between APs and VAPs, and an OpenFlow controller that orchestrates the network according to the user-defined policies. Similarly, Aeroflux [10] also promotes a separation between MAC features by implementing a 2-tier control plane. Here, the global control plane (GC) handles non-real time tasks such as authentication and load balancing, whereas the near-sighted control plane (NSC) is located closer to the APs to manage time-critical operations such as rate control and power adjustment. Then, this architecture emphasizes that control plane delays need to be short.

In contrast to the previous works reviewed, Ætherflow [11] and COAP (Coordination framework for Open APs) [12] extended the OpenFlow protocol in order to manage the communication between the controller and the APs. In consequence, both techniques simplify the data plane programmability as there is no need of extra software implementation. Thus, the extended OpenFlow protocol by itself comprises all the required messages to allow the controller gather different network statistics such as RSSI, SINR, bitrate or airtime usage.

2.3 SDWN applications for wireless networking

In the previous section we presented a set of different proposals. Most of them only propose or implement mechanisms to enhance mobility. However, here we present other functionalities that can be implemented:

- **Spatial reuse:** power control mechanisms are essential in order to reduce interference. In SDWN environments, thanks to the centralised control plane, power control mechanisms can be applied to avoid unnecessary overlaps between WLANs. In addition, the set-up of different clear channel assessment (CCA) levels could enhance the spatial reuse.

- **Dynamic channel allocation (DCA):** by gathering channel statistics in the controller, SDWN can perform dynamic channel allocation to minimize co-channel interference between WLANs.

- **Dynamic channel bonding (DCB):** the use of channel bonding based on the spectrum occupancy of neighboring WLANs can be performed as a solution to increase throughput rates and reduce interference between nodes, allocating different channel widths to each WLAN based on its traffic demands and capabilities.

- **Multi-AP communication:** by taking advantage of network programmability, multiple connections per user to different APs could be easily managed. The controller would be in charge of deciding whether or not the use of multiple simultaneous connections improve user and network performance, as well as to take actions by installing new forwarding rules in the forwarding devices.

- **Multiple input multiple output (MIMO) and multi-user (MU) MIMO coordination:** This application is more related with a joint SDN and SDR framework. However, the programmability of SDN creates a great opportunity for SDR to be applied and therefore, techniques such as interference coordination and alignment can be implemented in order to reduce and mitigate interfering signals. Coordination of such techniques can lead future WLANs to a new level of complexity, but with high performance gains.

3 Towards intelligent networking

The SDWN paradigm is extremely flexible as networks can be dynamically reconfigured to handle new states. Thus, the introduction of machine learning techniques constitute a potential solution to achieve higher gains in terms of network performance. By using different techniques, patterns can be extracted from data sets, or learned through interacting with the environment. Therefore, the knowledge extracted from past observations can be applied to update the behavior of the network. Existing machine learning algorithms are generally classified into three different categories depending on how the learning process is done. Supervised learning (SL) algorithms are trained using labeled examples. By comparing the predicted output with the labeled ones, these algorithms update the model accordingly to the error measured. On the other hand, unsupervised learning (USL) algorithms are used against data that has no historical labels. Thus, USL algorithms try to focus on arranging samples into different groups. Last, reinforcement learning (RL) algorithms, which through trial and error, try to find the actions that yield the greatest rewards.

The inclusion of machine learning into networking motivated the consideration of a new architectural division due to the fact that this kind of algorithms does not belong to data nor control planes. The new architectural division is the knowledge plane (KP), which was proposed in [13], and which intends to place machine learning techniques over the network architecture scheme. The KP is responsible for learning the behavior of the network, and the decision-making process. Basically, the KP processes the statistics collected by the control plane, transforms them into knowledge via machine learning algorithms, and uses that knowledge to make decisions. Hence, in the context of SDN networks, the KP participates actively in the network orchestration due to its interaction with the controller, which configures the network according to KP's instructions. In the literature, the joint consideration of SDN and machine learning techniques can be found as Knowledge-Defined Networking (KDN) [14]. This new paradigm consists in combining data, control and knowledge planes to provide automated network control. Figure 1 depicts an architecture that merges both KP and SDWN concepts to have a flexible wireless environment. Here, the SDN paradigm is identified in how the network is orchestrated since the control plane is managed by a controller that communicates and requests different information from the APs through OpenFlow. In regards to the machine learning related functionalities, a dedicated server, in which data is stored and machine learning algorithms executed, it is connected to the controller to take full advantage of network statistics to take decisions. Through the results from the machine learning algorithms, the decision-making process according to the knowledge obtained can be driven directly by the KP in an autonomous way based on a set of predefined requirements. On top of the controller, network applications are executed in order to give the directives to the controller for managing the network. In this context, some applications already done are:

- **Traffic prediction and classification:** both features were the earliest machine learning applications in the networking field. In this context, traffic classification is done in order to ensure QoS as well as quality of ex-

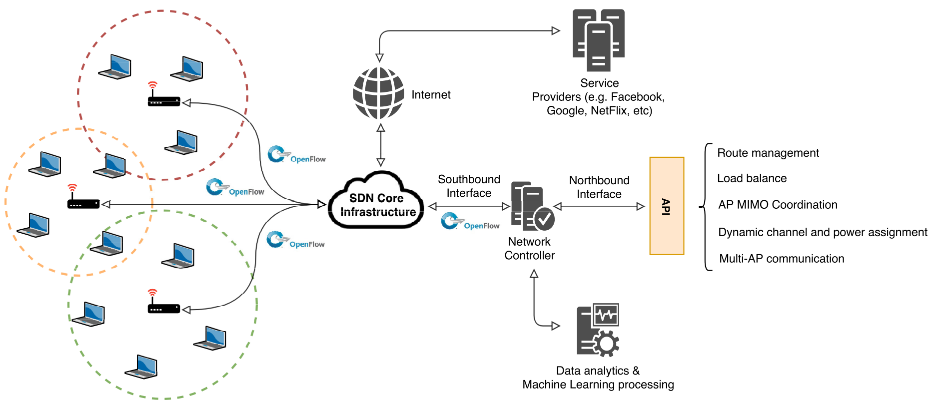


Figure 1: SDWN architecture with knowledge plane

perience (QoE). Thus, statistics gathered by the controller can be used to classify data flows into different QoS-categories. On the other hand, traffic prediction is used to forecast the total amount of traffic expected. As an example, in [15], neural networks (NN) are used to perform traffic prediction by using flow level statistics together with a learning window of past time intervals, which repetitively trains the algorithm in order to characterize and predict the network behavior. Traffic prediction solutions may lead to have proactive systems in which different actions can be triggered before traffic imbalances happen. For instance, some actions could lead to a reconfiguration of the spectrum allocation in order to provide more bandwidth to a group of WLANs, or trigger load balance mechanisms.

- **Routing:** Regarding to the management of the wired part of the network, routing strategies have been tackled such as in [16], in which is proposed a network congestion prevention mechanism based on the Q-learning algorithm. In case of detecting congestion between a link pair, the algorithm recomputes the reward matrix accordingly to the inputs, in order to search a new route. As the authors proved, in comparison with Dijkstra's algorithm, Q-learning based routing provides better results.

- **Security:** This is one of the most important factors that SDN architectures must face. The centralized nature of the control plane has many benefits, but it is a risky approach in terms of security, as all the network control is placed in a single point. For instance, current attacks such as denial of service (DoS) can be potentially critical, since the control plane is no longer distributed, and so the entire network can be compromised. In this context, machine learning can help to achieve a good level of security due to its ability to automatically find correlations in data. Deep learning techniques, such as the ANN proposed in [17], are good mechanisms to detect any anomaly by just analyzing few per flow statistics. So, the algorithm compares any incoming traffic with the previous ones and raises an alert when the deviation between them is greater than a certain threshold. In consequence, attacks such as DoS can be detected and mitigated.

- **Spatial reuse and channel bonding:** these are two techniques that are gaining attention since last IEEE 802.11ax amendment supports both of them. The former is based on the application of different techniques such as transmission power control (TPC) and CCA adjustment in order to control the potential drawbacks of uncoordinated deployments. The later refers to a technique in which two or more adjacent channels, within a given frequency band, are temporally combined to increase throughput and data transfer between devices. The application of this techniques have opened a new set of challenges in wireless environments and so, different works attempted to enhance the network performance by their application. First, in the work done in [18], MABs are used for finding the AP configuration that maximizes the aggregate throughput. There, the authors analyze different policies, in which the nodes' learning process is done by means of exploiting and exploring the medium. In regards of DCB, the work done in [19], assesses the problem for dense WLANs by evaluating different DCB policies. There, the authors show, through analytical results, that always selecting the widest available bandwidth is counterproductive in the long term. Moreover, authors conclude that, in non-fully overlapping scenarios, the optimal solution is to apply different policies depending on the context of each WLAN, and therefore they must be on-line learned.

4 Performance evaluation

In order to assess the integration of the KP, we have studied the application of machine learning algorithms to tackle the spatial reuse and channel bonding issues. To do so, we have considered an SDWN composed of different WLANs, whose APs' power and channel configurations are defined by the ML server, and then advertised by the controller. By including intelligent operations, we expect to increase the network performance, aiming to find the best configuration according to a policy, while empowering a collaborative behavior[1]. In this context,

[1] All the simulations have been performed using the SFCTMN framework developed in [19] and the learning package used in [18].

Algorithm 1: Implementation of Thompson Sampling for WLANs

Input: $\mathcal{A}$: set of possible actions in $\{1, ..., K\}$

initialize: $t = 0$, for each arm $k \in \mathcal{A}$, set $\hat{r}_k = 0$ and $n_k = 0$

while *active* **do**

  For each arm $k \in \mathcal{A}$, sample $\theta_k(t)$ from normal distribution $\mathcal{N}(\hat{r}_k, \frac{1}{n_k+1})$

  Play arm $k = \underset{1,...,K}{\text{argmax}}\ \theta_k(t)$

  Observe the reward $r_{k,t}$

  $\hat{r}_{k,t} \leftarrow \frac{\hat{r}_{k,t} n_{k,t} + r_{k,t}}{n_{k,t}+2}$

  $n_{k,t} \leftarrow n_{k,t} + 1$

  $t \leftarrow t + 1$

end

Table 2: Simulation parameters and action mapping

Parameter	Description	Value
C	Set of channels	1 / 2 / 3 / 4
P_{tx}	Set of transmit power values	1 dBm / 20 dBm
f	Central frequency	5 GHz
B	Bandwidth	20 MHz
SUSS	Spatial streams per user	1
G_{tx}	Transmitting gain	0 dBi
G_{rx}	Reception gain	0 dBi
P_n	Noise level	-95 dBm
CCA	Clear channel assessment	-62 dBm
Action number	**Transmission power**	**Channel number**
1	1 dBm	[36,40]
2	1 dBm	[44,48]
3	1 dBm	[36,40,44,48]
4	20 dBm	[36,40]
5	20 dBm	[44,48]
6	20 dBm	[36,40,44,48]

the problem is modelled through the multi-armed bandits (MABs) framework by defining a set of K configurations, which correspond to any combination of channel range and transmit power that each WLAN can select (refer to Table 2). Moreover, as an action-selection strategy, we use Thompson sampling (TS) algorithm, since it has been shown to provide better results than other well-known algorithms such as Upper Confidence Bound (UCB) for similar problems in WLANs [18]. The TS algorithm is a Bayesian algorithm that constructs a probabilistic model of the rewards observed by each configuration. After selecting an arm to play, TS observes the reward, and updates its prior belief in a way that the probability of a particular arm being optimal matches with the probability of each arm being selected. In practice, this is done by sampling each arm from its posterior distribution, and selecting the one that returns the maximum expected reward. Accordingly, it randomly selects the probabilistic optimal configuration. Algorithm 1 shows in detail the implementation of TS for this use case.

Regarding the reward function, we define a common goal for all the WLANs, which refers to maximize the minimum throughput. To allow a collaborative behavior, the resulting throughput of each WLAN, which is obtained by means of the Shannon capacity, is passed to the ML server. However, note that even if the rewards are known, actions are selected independently for each WLAN, as no other information regarding the configurations of the neighboring WLANs is informed. The Shannon capacity expression is shown in 1:

$$C = B \cdot \log_2(1 + \mathrm{SINR}) \tag{1}$$

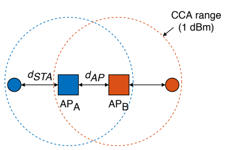
where B is the channel bandwidth, and the SINR is the signal-to-interference-plus-noise ratio given by $\mathrm{SINR} = \frac{P_s}{P_n + P_i}$. Here, the P_n and P_i refer to the noise and interference levels respectively, whereas the P_s refers to the signal level received at the AP, which is calculated through the path loss model proposed in [20] that is given in 2. This path loss model is simple but accurate, and it is used for 5GHz systems in indoor environments:

$$L_{prop}(d) = \mathrm{FSL} + \alpha \cdot d \tag{2}$$

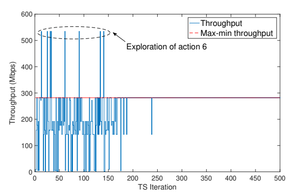
where FSL are the well-known free space losses at distance d, and $\alpha = 0.44$ dB/m is the constant attenuation per unit of path length. The different simulation parameters taken into account are described in table 2.

4.1 Full overlapping WLANs

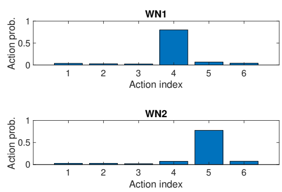
In this first scenario, which is presented in Figure 2a, we consider 2 WLANs that fully overlap. The parameters d_{STA} and d_{AP} are set to 5 m. Either choosing 1 dBm or 20 dBm, both APs will be inside the CCA range of its neighbor. At the end of the simulation, we observe that both WLANs reached the optimal configuration. The two WLANs selected the maximum transmission power, and a different channel scheme as it can be seen in Figure 2c. Besides, we can see that actions containing the whole set of channels have been explored but discarded, as they were only beneficial for one WLAN in detriment of the other. Moreover, in Figure 2b, we can observe that the max-min throughput converges into a collaborative solution before iteration 200, discarding selfish decisions. Regarding the transmission power, both networks decide to use the maximum allowable as using the lower value does not reduce the contention between the two WLANs.



(a)



(b)

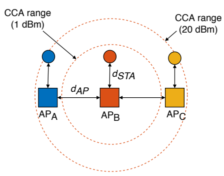


(c)

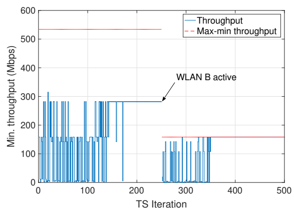
Figure 2: Use case with 2 WLAN. (a) Scenario considered. (b) Evolution of the throughput experienced by WLAN A. (c) Histogram of the probabilities for each action

4.2 Partial overlapping WLANs

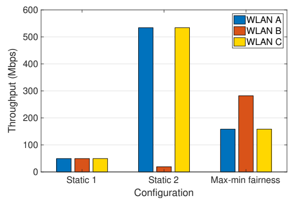
In this scenario, we want to tackle a non-stationary scenario by simulating changing conditions. For this purpose, we deploy three partially overlapping WLANS (Figure 3a), which activation time is different. Then, WLAN A and C are activated since the beginning, whereas WLAN B is activated at iteration 250. The parameters d_{STA} and d_{AP} are set to 5 m. Figure 3b shows the obtained results, and how TS fails at reaching the best possible configuration for the first 250 iterations, so WLANs A and C end up choosing different channel ranges in order to avoid interference. In this particular case, the optimal configuration is not found since it requires both WLANs to choose the optimal action simultaneously (i.e., minimum transmit power and the entire channel range). Moreover, in case that only one of the WLANs chooses the optimal one, it becomes vulnerable if the other WLAN uses maximum transmit power, thus leading to a low collaborative reward. On the other hand, when WLAN B becomes active, the three WLANs are able to choose the optimal configuration. Note that at iteration 250, the previous knowledge is discarded since the network state has changed. In Figure 3c, we have performed a comparison between applying learning, and leaving WLANs with an static configuration. We show that this kind of techniques can minimise the appearance of problems such as the flow starvation. For instance, from the scenario presented in Figure 3a, we can see that WLAN B will suffer flow starvation as the other WLANs will get most of the time to transmit. If we do not apply any mechanism and we explore the different available actions, we find that applying a conservative action (i.e., minimum power and minimum bandwidth) will lead to downgrade the performance of the three WLANs but maintaining the fairness. On the contrary, if we apply and aggressive solution, WLAN B barely



(a)



(b)



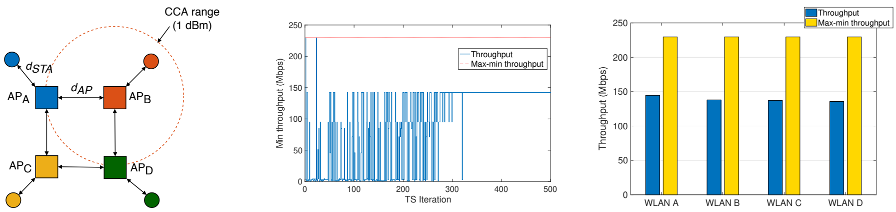
(c)

Figure 3: Use case with 3 WLAN. (a) Scenario considered. (b) Minimum throughput evolution. (c) Throughput per WLAN corresponding to different action settings. Static 1: All WLAN select action 1 (conservative). Static 2: All WLAN select action 6 (aggressive). Optimal configuration: WLANs A&C select action 1, whereas WLAN B selects action 5.

transmits. As a result, none of the previous solutions solves the situation without diminishing the performance of several WLANs, nor making the network unfair.

4.3 Grid scenario

Lastly, we have studied the behaviour of the proposed solution in a grid scenario, which is depicted in Figure 4a. The parameters d_{STA} and d_{AP} are set as $\sqrt{8}$ m and 5 m respectively. Here, we intend to see the interactions between multiple neighbors, and how the decisions of others affect the action-selection process. For this scenario, finding the optimal configuration in a decentralized way is unlikely to occur, since it requires that all WLANs choose the optimal action simultaneously. Therefore, there is a narrow window of possibilities for that to happen. In case that only one of the WLANs chooses the optimum, it becomes vulnerable and so leading to a low collaborative reward. However, as shown in Figure 4b, the learning algorithms reach a solution that is fair. Note that as the number of nodes increases, the convergence time increases too. So, the more nodes we have, the later we converge into a solution when considering a collaborative reward. Figure 4c shows a comparison among optimal and achieved throughput per WLAN.



(a) (b) (c)

Figure 4: Use case with 4 WLANs. (a) Scenario considered. (b) Minimum throughput evolution. (c) Throughput histogram.

5 Conclusions

In this article, we have shown that new networking paradigms, such as the presented SDN and SDWN, are grabbing attention from academia and research institutions, with a clear aim to be used in next generation of WLAN deployments. Besides, big data mining and machine learning techniques are also raising attention due to their ability to use collected information for improving network management. In this regard, we have performed different study cases to analyse the behavior of ML over wireless networks for management purposes. ML and SDWN can be perfectly combined to achieve better performance, as the results obtained prove that there is a clear improvement over the pre-defined configurations. However, further research must be carried out in order to quantify the different drawbacks and trade-offs that exist, such as the negative effects that greater network delays can have in the overall network performance.

Acknowledgment

This work has been partially supported by the Spanish Ministry of Economy and Competitiveness under the Maria de Maeztu Units of Excellence Programme (MDM-2015-0502), by the Catalan Government under grant 2017-SGR-1188, by the Spanish Government under grant PGC2018-099959-B-I00 (MCIU/AEI/FEDER,UE), and by a Gift from the Cisco University Research Program (CG#890107, Towards Deterministic Channel Access in High-Density WLANs) Fund, a corporate advised fund of Silicon Valley Community Foundation. The work done by S. Barrachina-Muñoz is supported by a FI grant from the Generalitat de Catalunya.

References

[1] Cisco Visual Networking Index. Cisco visual networking index: Global mobile data traffic forecast update, 2016-2021. *White Paper*, March, 2017.

[2] B. Bellalta. IEEE 802.11ax: High-efficiency WLANs. *IEEE Wireless Communications*, 23(1):38–46, 2016.

[3] Kok-Kiong Yap, Masayoshi Kobayashi, Rob Sherwood, Te-Yuan Huang, Michael Chan, Nikhil Handigol, and Nick McKeown. Openroads: Empowering research in mobile networks. *ACM SIGCOMM Computer Communication Review*, 40(1):125–126, 2010.

[4] Lalith Suresh, Julius Schulz-Zander, Ruben Merz, Anja Feldmann, and Teresa Vazao. Towards programmable enterprise wlans with odin. In *Proceedings of the first workshop on Hot topics in software defined networks*, pages 115–120. ACM, 2012.

[5] Julius Schulz-Zander, Carlos Mayer, Bogdan Ciobotaru, Stefan Schmid, and Anja Feldmann. Opensdwn: programmatic control over home and enterprise wifi. In *Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research*, page 16. ACM, 2015.

[6] Yiannis Yiakoumis, Manu Bansal, Adam Covington, Johan van Reijendam, Sachin Katti, and Nick McKeown. Behop: A testbed for dense wifi networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 18(3):71–80, 2015.

[7] Henrique Moura, Gabriel VC Bessa, Marcos AM Vieira, and Daniel F Macedo. Ethanol: Software defined networking for 802.11 wireless networks. In *Integrated Network Management (IM), 2015 IFIP/IEEE International Symposium on*, pages 388–396. IEEE, 2015.

[8] Roberto Riggio, Mahesh K Marina, Julius Schulz-Zander, Slawomir Kuklinski, and Tinku Rasheed. Programming abstractions for software-defined wireless networks. *IEEE Transactions on Network and Service Management*, 12(2):146–162, 2015.

[9] Jonathan Vestin, Peter Dely, Andreas Kassler, Nico Bayer, Hans Einsiedler, and Christoph Peylo. Cloudmac: towards software defined wlans. *ACM SIGMOBILE Mobile Computing and Communications Review*, 16(4):42–45, 2013.

[10] Julius Schulz-Zander, Nadi Sarrar, and Stefan Schmid. Aeroflux: A near-sighted controller architecture for software-defined wireless networks. In *Presented as part of the Open Networking Summit {ONS}*, 2014.

[11] Muxi Yan, Jasson Casey, Prithviraj Shome, Alex Sprintson, and Andrew Sutton. Aetherflow: principled wireless support in sdn. In *2015 IEEE 23rd International Conference on Network Protocols (ICNP)*, pages 432–437. IEEE, 2015.

[12] Ashish Patro and Suman Banerjee. Coap: A software-defined approach for home wlan management through an open api. *ACM SIGMOBILE Mobile Computing and Communications Review*, 18(3):32–40, 2015.

[13] David D Clark, Craig Partridge, J Christopher Ramming, and John T Wroclawski. A knowledge plane for the internet. In *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 3–10. ACM, 2003.

[14] Albert Mestres, Alberto Rodriguez-Natal, Josep Carner, Pere Barlet-Ros, Eduard Alarcón, Marc Solé, Victor Muntés-Mulero, David Meyer, Sharon Barkai, Mike J Hibbett, et al. Knowledge-defined networking. *ACM SIGCOMM Computer Communication Review*, 47(3):2–10, 2017.

[15] Abdelhadi Azzouni and Guy Pujolle. Neutm: A neural network-based framework for traffic matrix prediction in sdn. *arXiv preprint arXiv:1710.06799*, 2017.

[16] Seonhyeok Kim, Jaehyeok Son, Ashis Talukder, and Choong Seon Hong. Congestion prevention mechanism based on q-leaning for efficient routing in sdn. In *Information Networking (ICOIN), 2016 International Conference on*, pages 124–128. IEEE, 2016.

[17] Tuan A Tang, Lotfi Mhamdi, Des McLernon, Syed Ali Raza Zaidi, and Mounir Ghogho. Deep learning approach for network intrusion detection in software defined networking. In *Wireless Networks and Mobile Communications (WINCOM), 2016 International Conference on*, pages 258–263. IEEE, 2016.

[18] Francesc Wilhelmi, Cristina Cano, Gergely Neu, Boris Bellalta, Anders Jonsson, and Sergio Barrachina-Muñoz. Collaborative spatial reuse in wireless networks via selfish multi-armed bandits. *Ad Hoc Networks*, 88:129–141, 2019.

[19] Sergio Barrachina-Muñoz, Francesc Wilhelmi, and Boris Bellalta. Dynamic channel bonding in spatially distributed high-density WLANs. *(In press) IEEE Transactions on Mobile Computing*, 2019.

[20] Jonas Medbo and J-E Berg. Simple and accurate path loss modeling at 5 ghz in indoor environments with corridors. In *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, volume 1, pages 30–36. IEEE, 2000.

Concurrent Decentralized Channel Allocation and Access Point Selection using Multi-Armed Bandits in multi BSS WLANs

Álvaro López-Raventós and Boris Bellalta

Abstract

Enterprise Wireless Local Area Networks (WLANs) consist of multiple Access Points (APs) covering a given area. In these networks, interference is mitigated by allocating different channels to neighboring APs. Besides, stations are allowed to associate to any AP in the network, selecting by default the one from which receive higher power, even if it is not the best option in terms of the network performance.

Finding a suitable network configuration able to maximize the performance of enterprise WLANs is a challenging task given the complex dependencies between APs and stations. Recently, in wireless networking, the use of reinforcement learning techniques has emerged as an effective solution to efficiently explore the impact of different network configurations in the system performance, identifying those that provide better performance.

In this paper, we study if Multi-Armed Bandits (MABs) are able to offer a feasible solution to the decentralized channel allocation and AP selection problems in Enterprise WLAN scenarios. To do so, we empower APs and stations with agents that, by means of implementing the Thompson sampling algorithm, explore and learn which is the best channel to use, and which is the best AP to associate, respectively. Our evaluation is performed over randomly generated scenarios, which enclose different network topologies and traffic loads. The presented results show that the proposed adaptive framework using MABs outperform the static approach (i.e., using always the initial default configuration, usually random) regardless of the network density and the traffic requirements. Moreover, we show that the use of the proposed framework reduces the performance variability between different scenarios. Results also show that we achieve the same performance (or better) than static strategies with less APs for the same number of stations. Finally, special attention is placed on how the agents interact. Even if the agents operate in a completely independent manner, their decisions have interrelated effects, as they take actions over the same set of channel resources.

1 Introduction

In the past few years, multimedia contents, such as social networks, virtual reality, on-demand video platforms and video-streamed gaming, have witnessed a remarkable growth in terms of bandwidth consumption. The widespread use of IEEE 802.11 based wireless local area networks (WLANs), commonly known as WiFi, has helped network providers to cope with the increasing demands of wireless communications [1].

Since the growing demand of data services, we can find enterprise WLANs in a wide range of private and public spaces. Enterprise WLANs are composed by several APs, also called basic service sets (BSSs), that coexist under the same extended service set (ESS), allowing different APs to keep the same service set identifier (SSID). Then, end-users, i.e., smartphones, laptops, or tablets perceive the whole set of deployed APs as a unique WLAN, allowing them to roam from one AP to another keeping the connectivity. Normally, such WLANs are deployed in places like airports, shopping malls, or university campuses, resulting in a very attractive solution to provide Internet access.

Although being a cost-effective solution, it is well-known that WiFi networks may suffer from severe performance degradation in dense deployments. First, a loss in performance can be related to the limited number of channels that are available in the Industrial Scientific and Medical (ISM) band, which make WLANs to highly suffer from co-channel and adjacent-channel interference (CCI and ACI) issues. For example, the need to avoid interference has lead to only use three of the different available channels in the 2.4 GHz band. During the years, the spectrum scarcity has been a severe problem. In order to prevent this situation, new IEEE 802.11 WLAN amendments, such as the IEEE 802.11ac [2], and the upcoming IEEE 802.11ax [3, 4, 5] and 802.11be [6], promote the use of 5 GHz and 6 GHz bands.

On the other hand, the use of a random channel access mechanism may arise an unsatisfactory user experience in dense areas. In WiFi networks, the medium access control (MAC) is performed through the Distributed Coordination Function (DCF), which implements a carrier sense multiple access with collision avoidance (CSMA/CA) mechanism. The DCF operation is simple but functional, and in general, it performs well as long as the user density or throughput requirements are kept low. However, the higher the number of stations in a given area, the higher the probability of having unsuccessful data exchanges. In addition, this effect may be boosted by the strongest signal first (SSF) AP selection mechanism, which may create an unbalanced use of the different APs. The main reason is that the SSF is based purely on a physical (PHY) metric, such as the received signal strength indicator (RSSI), without taking into consideration how much traffic is already being handled by the destination AP. Although this approach may work under low load conditions, it is unfeasible to apply when considering dense scenarios. Then, this situation evokes to think in a different AP selection strategy based on a more complex metric capable to capture the network conditions (i.e. users' positions, current channel load, nodes' configuration, etc.)

As envisioned, resource management strategies must consider the instantaneous users' activity requirements as part of network configuration procedures. To proceed with this shift, we address the channel allocation (CA) and AP selection (APS) problems in enterprise WLANs through machine learning (ML) mechanisms, and in particular with the reinforcement learning (RL) branch. We adopt the well-known multi armed bandits (MABs) framework to implement a dynamic CA (DCA), as well as a dynamic APS (DAPS). To do so, we provide stations and APs with agents that are expected to autonomously learn the best performing action based on the environment. Besides, we address the learning problem in a decentralized adversarial multi-player context, in which APs and stations' agents, compete for a common set of resources using different action sets. Setting up this kind of adversarial multi-player environment is very challenging due to the fact that the reward experienced by an agent is conditioned not only by its own actions, but also by the actions performed by others[1].

In this work we want to assess the feasibility of a decentralized solution using learning MABs to concurrently perform channel allocation and AP selection. Rather than proposing an actual solution, the relevance of our work remains in the study of MABs algorithms in adversarial multi-player environments. Hence, the goal of this paper is to provide different insights on how nodes intelligently modify their configuration, by learning from past experiences to improve future performance. Then, efficient learning and intelligent decision-making algorithms are key to achieve such objective. In this paper, the lack of information, and the high level of uncertainty among the different actions are two conditions that have been addressed through the use of the MAB framework, as the decision-making process is carried out through a learning-by-interaction approach.

The contributions of this paper are:

- We consider the use of MABs for network optimization in an adversarial multi-player setup. To do so, we take into account aspects such as user and network dynamics, asynchronous and continuous time operation of the agents, realistic reward computation, age of information, and the interaction between agents.

- We study the case of concurrent decentralized channel allocation and AP

[1]Note that we extend the definition of adversarial setting to the case where actions taken by other players change the distribution of our rewards. This situation could also be seen as a non-stationary setting, where environment conditions change along with the actions taken by the other players.

selection problems in large enterprise WLANs scenarios. We use the Thompson sampling (TS) algorithm as an action-selection strategy, showing the effectiveness of such a technique to adapt to changing conditions, as well as to adversarial scenarios where actions from one agent affect to the distribution of the system rewards observed by the other agents.

- We evaluate the system performance for different network setups (i.e., different traffic patterns, different number of stations and APs). We show that using the developed framework, the system is able to convergence to a better solution. Finally, we consider a non-stationary case beyond the default adversarial setting, in order to validate the ability of the TS algorithm to adapt to changes.

The remainder of this article is organized as follows. In Section 2 we describe the related work. Then, Section 3 describes the system model under consideration as well as all the considerations to formulate the joint DCA and DAPS. The problem statement is assessed in Section 4 jointly with the MABs framework, whereas the simulation results are presented in Section 5. Finally, we provide some conclusions in Section 6.

2 Related work

Channel allocation and AP selection in WiFi networks have been widely studied by the research community. In this section we overview some relevant works that can be found in the literature, with the aim to properly frame the contributions of this paper. Note that, since we focus in WiFi networks, we have only considered works related to this technology.

2.1 Channel Allocation

Two main approaches regarding channel allocation have been discussed over the years. We refer to *opportunistic channel allocation* [7, 8, 9], which is intended to overcome frequency holes, and promote higher frequency utilization by accessing them in a short-interval basis, and *dynamic channel allocation*, which is intended to be responsive against changing network conditions in a long-term basis.

In this paper, we are only interested in the study of DCA. In this regard, recent approaches have incorporated the use of ML to improve the channel allocation process. In [10], authors propose an approach that exploits an smart channel selection strategy by classifying the traffic pattern on primary channels, and choosing the channel with the longest idle time. It is shown that using their approach the amount of collisions can be reduced drastically, increasing the system performance. Besides, [11] propose an online CA by adopting the MABs framework. On top of the MABs framework, they implement a weighted algorithm in order to carry the action selection process, in which the probability of selecting a certain action is adjusted according to the regret observed. More recently, authors in [12] proposed a channel allocation method based on graph analysis, linear programming and regression to minimize the overlap among APs. An study about the exploration-exploitation trade-off for different learning algorithms with the objective to achieve the best pair of channel and power allocation is presented in [13]. In addition, authors compare the performance of the different considered action selection strategies, while studying the implications of applying them under an adversarial setting. In [14], authors propose a dynamic-wise, light-weight and decentralized, online primary channel selection algorithm for performing dynamic channel bonding, which considers the activity on both target primary and secondary channels in order to maximize the expected throughput.

Finally, we can find other works that are based on a centralized architecture. Nowadays, this type of architectures are taken a lot of interest as they can get a global picture of the network state. Under this approach, a dynamic channel selection for sectorized WiFi cells is implemented in [15]. Here, the network is continuously monitored in order to identify the WiFi interference sources, so can be mitigated through establishing a configuration. In addition, other common framework used is the software defined network (SDN) paradigm, which is employed in [16] to address the spectrum congestion in dense deployments. Then, the objective is to capture the network state, so an optimized channel assignment can be executed to minimize the interference. Although solutions based on centralized architectures are very powerful as the central controller has an overall picture of the network, these solutions may not be appropriate for high dynamic scenarios where the network state changes fast.

2.2 Access Point Selection

In regards of access point selection by the stations, we can identify different solutions depending on the optimization target. For instance, we can find mechanisms that try to minimize the number of stations per AP, whereas other schemes try to maximize the RSSI, or the throughput achieved. In this context, authors in [17] evaluate an association control algorithm to optimize the throughput in WLANs. Bandwidth demands of users are considered as constraints to carry the association process, which is performed through estimating the AP utilization. In [18] authors present two different association schemes. The first one is based on channel quality in both uplink and downlink, whereas the second one uses the airtime metric of each cell. In [19] the average workload of the network is used to redistribute the traffic when a new station joins the network or when the signal quality of a client deteriorates. The proposed approach however requires changes to the standard beacon frames. A similar scheme is presented in [20] where the stations are migrated to the least loaded AP in order to balance the traffic load. However, since the channel quality is not considered, this approach may significantly reduce the aggregate throughput, as no consideration regarding the performance anomaly effect is done.

As well as in the case of CA, some articles explored the station association in centralized environments. In [21], authors explore an online AP selection process for 802.11n with heterogeneous clients (802.11a/b/g/n), with the objective to evaluate the impact of legacy clients. Moreover, authors in [22] use an SDN based solution to solve an unbalanced distribution of the stations among APs. The APs that are congested due to a high number of connected stations are requested to reconfigure their transmission power in order to force a hand-off process in some stations. However, this kind of approaches may not work properly since the number of attached clients is not an accurate estimator of the load. In this context, authors in [23] propose, over a software defined WiFi network, an association scheme capable to detect situations in which the traffic is not efficiently distributed and so, reschedule to other APs the clients whose transmissions are causing performance issues.

In regards of works using ML techniques, we can already find some papers. Authors in [11] also tackled the association process. They use the same weighted algorithm to perform both channel allocation and AP selection. In [24] is proposed a decentralized approach to perform the AP selection through the MABs framework. In their solution, authors propose an extension of the epsilon greedy algorithm that includes stickiness to perform the AP selection, which results into a notable improvement in the system performance. It is important to mention that to the best of our knowledge and up to date, we have not found any other works that adopt ML for improving the association process.

3 System model

In this section, we introduce the enterprise WLAN scenario considered in this paper. We expose the main assumptions that have been done, as well as presenting the CSMA/CA abstraction used to model the WiFi operation. Finally, we introduce the performance metrics we will use in Section 5. Table 1 summarizes the notation used through this paper.

3.1 WiFi Network description

We consider an enterprise WLAN composed by a set of APs $\mathcal{A} = \{A_1, \ldots, A_n\}$, and a set of stations $\mathcal{S} = \{S_1, \ldots, S_m\}$, where n and m are the total number of APs and stations respectively. Over a given area, both types of devices are

Table 1: Notation used in the system model

Notation	Description
n	Number of APs in the network
m	Number of stations in the network
k	Number of available channels
$\mathcal{A} = \{A_1, \ldots, A_n\}$	Set of deployed APs
$\mathcal{S} = \{S_1, \ldots, S_m\}$	Set of deployed stations
$\mathcal{C} = \{c_1, \ldots, c_k\}$	Set of available radio channels
$\mathcal{S}_j$	Stations attached to AP j
$\mathcal{N}_j$	Neighboring APs detected by AP j
$\mathcal{N}_j^c$	Neighboring APs using the same channel c that AP j
$\mathcal{A}_i$	APs detected by station i above the RSSI_{th}
B_i	Bandwidth requested by station i
L_{d}	Data packet size
$r_{i,j}$	Bit rate for station i, when associated to AP j
$u_{i,j}(B_j, L_{\text{d}}, r_{i,j})$	Airtime required by station i when associated to AP j
$t_s(r_{i,j})$	Duration of successful packet transmission between station i and AP j
t_e	Duration of an empty slot
$E[\psi]$	Expected backoff duration
p_e	Packet error probability
Ψ_j^c	Average channel reward for an AP j using channel c
$\Omega_{j,i}$	Average satisfaction for a station i attached to an AP j
ℓ_j^c	Channel load experienced at AP j using channel c

randomly placed following a uniform distribution. In addition, we consider that they all implement 802.11k and 802.11r amendments [25]. The 802.11k amendment introduces new functionalities to support resource management, whereas the 802.11r amendment addresses the transition from one AP to another within the same WLAN aiming to minimize the interruption of connectivity.

In regards of APs, let us define their action set as $\mathcal{C} = \{c_1, \ldots, c_k\}$, which is composed by the different available radio channels. Thus, an AP j will select a channel $c \in \mathcal{C}$. It is worth mention that all the APs' share the same action space $\mathcal{C}$ and so, different APs may select the same channel c. The set of stations within the Clear Channel Assessment (CCA) area of an AP j is denoted as $\mathcal{S}_j$, whereas the neighboring APs is expressed as $\mathcal{N}_j$. Additionally, $\mathcal{N}_j^c$ will refer to the neighboring APs using the same radio channel c of AP j. Note that since positions are assigned randomly, each AP may have different entries for $\mathcal{S}$ and $\mathcal{N}$.

On the other hand, from the stations' perspective, we define as $\mathcal{A}_i$ the action set for a station i, which is composed by all the APs seen by station i that are above a certain RSSI_{th} threshold, which is a system parameter added to improve the AP selection process[2]. By not including APs below the $\mathrm{RSSI}_{\mathrm{th}}$ threshold, we avoid exploring APs placed too far, in which the probability of being unsatisfied is significantly high. In addition, properly configuring this threshold allows to minimize the performance anomaly [26], which is produced when one station occupies the channel for a long time due to its low transmission rate, penalizing other stations that use higher rates. Notice that the size of the stations' action set may be different for each station, as it only depends on the number of detected APs a station i detects above the mentioned threshold.

Only downlink traffic is considered. We model stations' activity using an on/off Markovian model, where both active periods (the station requires a certain downlink throughput) and inactive periods (the station is idle) are exponentially distributed with mean T_{on} and $\mathrm{T}_{\mathrm{off}}$, respectively. Every time a station activates (moves to the 'on' state), we will say a new downlink traffic flow, or simply a flow, starts.

3.2 CSMA/CA model abstraction

In order to evaluate the DCA and DAPS over large-scale WLAN Networks for large periods of time (several hours), we abstract the CSMA/CA operation. While the considered abstraction does not capture low-level details of the PHY and MAC layers operation, it maintains the essence of the CSMA/CA: the 'fair' share of the spectrum resources among contending APs and stations. Basically, the considered abstraction takes into account the aggregate channel load at each AP to calculate the airtime that can be allocated to each station.

To explain how the proposed CSMA/CA abstraction works, refer to Figure 1, where an illustrative example of an enterprise WLAN (3 APs, under the same ESS) is depicted. In this example, we consider that two APs use the same channel, whereas the remaining AP uses a different one. Some stations are deployed within the coverage area of each AP, which will act as receivers of the data flows. Depending on the throughput requirement B_i of each flow, the length of a data packet L_{d}, and the bit rate $r_{i,j}$, the total airtime that an station i will require from its serving AP j is given by

$$u_{i,j}(B_i, L_{\mathrm{d}}, r_{i,j}) = \frac{1}{(1-p_e)} \left\lceil \frac{B_i}{L_{\mathrm{d}}} \right\rceil (E[\psi]t_e + t_s(r_{i,j}, L_{\mathrm{d}})), \tag{1}$$

where p_e is the packet error probability, the term $\frac{1}{(1-p_e)}$ represents the average number of transmissions per packet, $E[\psi]$ is the average backoff duration, t_e is the

[2]If an station i only detects one AP over the CCA threshold, whose RSSI is lower than the $\mathrm{RSSI}_{\mathrm{th}}$, the action set $\mathcal{A}_i$ will be composed only by this entry, as connectivity is ensured for all stations. Therefore, stations with a unique entry in their action set will not be able to perform any learning

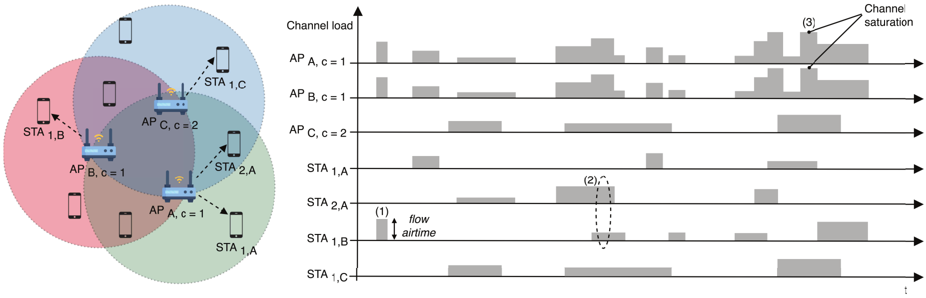


Figure 1: Example of the CSMA/CA abstraction considered in this work. Horizontal axis represents time, whereas vertical axis represents the channel load. In the APs' axes it is represented the aggregate traffic, whereas the individual traffic is represented in the stations' axes. At (1), it is represented a single flow transfer. At (2), there are represented two concurrent traffic flows, and their repercussion on the neighboring APs's channel load using the same channel. At (3), $\mathrm{AP_A}$ and $\mathrm{AP_B}$ reach saturation conditions due to the higher traffic demands.

duration of an empty slot, and $t_s(r_{i,j}, L_\mathrm{d})$ is the duration of a successful packet transmission, which is given by:

$$t_s(r_{i,j}, L_\mathrm{d}) = t_\mathrm{RTS} + 3t_\mathrm{SIFS} + t_\mathrm{CTS} + t_\mathrm{DATA}(r_{i,j}, L_\mathrm{d}) + t_\mathrm{ACK} + t_\mathrm{DIFS} + t_e, \quad (2)$$

where t_DIFS and t_SIFS are the distributed inter-frame space (DIFS) and the short inter-frame apace (SIFS), $t_\mathrm{DATA}(r_{i,j}, L_\mathrm{d})$ is the duration of a data packet at the transmission rate used between station i and AP j, and t_ACK correspond to the time that an acknowledgment (ACK) packet lasts. In A, more details about how these parameters are computed can be found, as well as the description of the 11ax path-loss model considered to obtain the transmission rates.

As Figure 1 shows, at point (1), $\mathrm{AP_A}$ initiates a flow transfer to its attached station $\mathrm{STA_{1,A}}$. Due to this event, $\mathrm{AP_B}$ starts sensing the medium busy, registering an increment of the channel load. Note that this effect is a consequence of $\mathrm{AP_B}$ using the same channel as $\mathrm{AP_A}$. Moving to point (2), there are depicted two concurrent flows from $\mathrm{AP_A}$ and $\mathrm{AP_B}$ to their corresponding stations. Here, we observe that the effective channel load experienced by both APs corresponds to the aggregate airtime for all active flow transfers, either if they are due to their own flows or from neighboring APs. It is important to highlight that this effect will not happen if the APs were not allocated to the same channel. In addition, it is interesting to note that new incoming flows either from $\mathrm{AP_A}$ or $\mathrm{AP_B}$ to their stations may be successfully served, since the aggregate airtime of the active flows does not exceed the maximum allocable time[3]. Finally, at point (3), several concurrent flows lead $\mathrm{AP_A}$ and $\mathrm{AP_B}$ to experience channel saturation, as the maximum allocable airtime is achieved. Under this conditions, the airtime of each flow will be reduced in order to remain proportional to the maximum allocable value, and so, remaining at the saturation point.

It is worth mention that $\mathrm{AP_C}$ does not perceive any change since it uses a different channel and so, its load is not affected by activity registered in the other channel. In Section 3.3.2 a numerical example is presented to further illustrate these interactions.

3.3 Performance metrics

To evaluate the performance of the system, we define two different metrics that will be used when carrying out the decision-making process. First, we define the channel reward metric for the DCA, which will be based on the channel occupancy. Then, we define the stations' satisfaction, which is related to the airtime, and it will be employed to evaluate the performance of the DAPS scheme. Both metrics will be used by the MAB framework introduced in Section 4.

3.3.1 Channel reward

We define the channel reward as a metric to characterize the occupancy level experienced when using a certain channel. Since the purpose of this metric is to evaluate how channels are performing, it is only intended for APs to use it. To compute the channel reward, APs will register new instances of this metric every time they detect a change on the channel load in the channel in which they are operating. Then, every time the DCA agent activates, it will average all the obtained values, within a time interval specified by a temporal window, in order to get a quantitative representation of the channel occupancy level (i.e., the average channel load). In further Section 4.2.1 the temporal window concept is presented. By using this metric, we consider that APs will have a detailed vision of the spectrum usage, accurately capturing all the intrinsic dynamics of a wireless environment. We define the channel reward for an AP j that uses a channel c at time t as

$$\Psi_j^c(t) = \max(0, 1 - \ell_j^c(t)) \leq 1, \tag{3}$$

[3]Since throughput is measured in Mbit/s, the maximum allocable time for transmissions is 1 second. Thus, if the effective channel occupancy experienced, and consequently the effective channel load, surpasses this value, the traffic requirements will not be met and so, we will talk about a saturated channel.

where $\ell_j^c(t)$ is the effective channel load at the AP j when it uses the channel c expressed as

$$\ell_j^c(t) = \sum_{\forall n \in \mathcal{N}_j^c} \ell_n^c(t) + \sum_{\forall i \in \mathcal{S}_j} u_{i,j}(B_i, L_{\mathrm{d}}, r_{i,j}). \tag{4}$$

Indeed, as stated in Section 3.2, $\ell_j^c(t)$ will depend on the channel load due to the own flows of AP j, and $\ell_n^c(t)$, which is the channel load registered due to flows from neighboring APs (using the same channel). Note that the effective channel load experienced can be higher than 1.

3.3.2 Station's satisfaction

We adopt the concept of station's satisfaction to assess whether an association pair between an station and its serving AP is performing well or not, and therefore, it will be only used by stations. Conceptually, we define the satisfaction as the ratio between the required airtime by an station, and the actual amount that can be allocated by its serving AP. Then, we will refer to satisfied stations if the resulting value of the metric is one (traffic requirements fulfilled), whereas we will refer to unsatisfied stations if the resulting value is lower than one (traffic requirements can not be fulfilled).

To compute the metric, stations are intended to ask for the total amount of channel load that APs have experienced. In order to use this capability, stations are considered to be compliant with the 802.11k amendment, which defines the *channel load request* messages. As defined in the standard, this type of frame is composed by a request-response sequence in which stations can ask for the amount of time in which the channel has been measured as busy (either through physical or the virtual carrier sense mechanism). As well as the channel reward, when the DAPS agent is activated, this metric averages all the tracked measures within a time interval specified by a temporal window, in order to get a quantification of the performance. Note that the satisfaction may change during the lifetime of a flow, and so we track all those changes, to average them at the end. We define the satisfaction for a station i associated to AP j operating in channel c at time t as:

$$\Omega_{i,j}(t) = \frac{\min(1, \ell_j^c(t))}{\ell_j^c(t)} \leq 1 \tag{5}$$

where $\ell_j^c(t)$ is the channel load as defined in (4). Since we consider that all resources are proportionally allocated in our CSMA/CA abstraction, the satisfaction value obtained by stations under the same AP will be the same.

To clearly understand how this metric works, refer to the point marked as (2) in Figure 1. Here, we assume that $\mathrm{STA}_{1,\mathrm{A}}$ and $\mathrm{STA}_{1,\mathrm{B}}$ require a traffic load of 40% and 30% respectively. As both APs share the same channel, the channel load perceived adds up to 70%, lower than the maximum 100%, and therefore, making stations to be satisfied as they receive the airtime allocation that they need. On the contrary, in point (3), we consider that $\mathrm{STA}_{1,\mathrm{A}}$ still requires 40%, but $\mathrm{STA}_{1,\mathrm{B}}$ increases its traffic needs to 90%. This higher requirement of $\mathrm{STA}_{1,\mathrm{B}}$ makes APs to enter in saturation, since the total channel load raises up to 130%. As a result, the satisfaction experienced by $\mathrm{STA}_{1,\mathrm{A}}$ and $\mathrm{STA}_{1,\mathrm{B}}$ scores a value of 76,9%. Essentially, this value indicates that only the 76,9% of the required airtime of both stations (i.e., 30.76% for $\mathrm{STA}_{1,\mathrm{A}}$, and 69.21% for $\mathrm{STA}_{1,\mathrm{B}}$) will be allocated. Again, note that, all active stations will receive the same satisfaction as we consider that resources are proportionally distributed.

Once we have the satisfaction, the throughput achieved by station i, associated to AP j at time t is given by

$$\Gamma_{i,j}(t) = B_i \Omega_{i,j}(t) \tag{6}$$

3.4 Problem formulation

By using airtime related metrics, we intend to make both APs and stations capable to keep track of the network congestion, and how it affects the stations' satisfaction. Here, we define the target objectives for DCA and DAPS, respectively.

First, from the APs' perspective, the strategy to ensure a good network performance is to select the less congested channel. Therefore, the optimizations problem is reduced to maximize the channel reward and so, it can be expressed as

$$c^* = \underset{\forall c \in \mathcal{C}}{\operatorname{argmax}} \ \Psi_j^c \tag{7}$$

On the contrary, from an station perspective, the strategy to enhance its own satisfaction is to minimize the congestion observed at its serving AP. Therefore, we have designed the DAPS relying on the satisfaction metric in order to decide whether an AP can be considered as a potential serving AP or not. Then, we formulate this problem as

$$a^* = \underset{\forall a \in \mathcal{A}_i}{\operatorname{argmax}} \ \Omega_{i,a} \tag{8}$$

We can observe that the optimization problem has been formulated as a maximization for both channel reward and station satisfaction. Then, using the proposed decentralized framework, APs and stations will take decisions autonomously to try to accomplish their target. Moreover, since decisions are taken asynchronously, we will see high variations when analyzing the behavior of the system as the reward observed by each agent will depend on the action selection of all other agents. However, in the long term, we expect a reduction of the variance as the system will enter in a steady state regime, in which APs will have a fair distribution of the network load, and so of the stations associated to them.

4 DCA and DAPS through the MABs framework

In this section we introduce the operation of the agent-based framework for decentralized channel allocation and access point selection, the MABs framework in which it is based, and the action-selection strategy that will be used following the performance metrics presented in Section 3.

4.1 Multi-Armed Bandits

The multi-armed bandit problem models an interaction between a learning agent, often called player, and an environment. Traditionally, the agent decides on a number of alternative arms or actions, which iteratively pulls, one at a time, during a number of rounds ($t = 1, 2, 3, ..., T$). From each action played, the learning agent receives a reward from the environment, which is used to evaluate the performance of the action, as well as to select subsequent actions. Then, the goal of the learning agent is to maximize the long-term reward to reach an optimal result. In addition, this strategy typically involves an exploration/exploitation trade-off, in which the agent must deal between learning at a faster or slower pace. To manage this trade-off, the learning rate parameter is used to balance both exploration and exploitation tasks in order to acquire enough knowledge to maximize the payoff. Note that a faster learning rate may lead to not exploring enough, ending into a suboptimal solution, whereas an slower learning rate may waste too much time on bad decisions. Therefore, tuning and selecting the appropriate learning rate is fundamental in order to achieve good results.

We can find different types of MABs depending on the characteristics of the reward. Typically, they are classified into stochastic, bayesian, contextual and adversarial bandits. For instance, in stochastic bandits actions have and independent and identical reward distribution, whereas in bayesian bandits, an arm is selected following a probability distribution that is proportional to the historic of the rewards experienced by that arm. Works such as [27], [28] [29], [30], [31] show the wide variety of applications in wireless communications. In addition, a further and extensive introduction to the MABs framework can be found in [32].

Independently of the type, the main objective of the MABs framework is to find the arm or action that maximizes the obtained reward. To do so, a common way of measuring the performance of MABs algorithms is by means of the regret function. The regret for a player i at time t, after T rounds as stated in [33], is

$$R_{i,t} = Tr^*_{i,t} - \sum_{t=1}^{T} r_{i,t}, \tag{9}$$

where $r^*_{i,t}$ is defined as the reward given by the optimal action at time t, and $r_{i,t}$ is the reward obtained by the current action selected. From the regret definition, learn is said to happen if the cumulative regret function grows sublinearly, and therefore, the algorithm is able to identify the action with the highest reward. In this case, the expected regret, $\mathbb{E}[R_{i,t}]$, will decrease over time, converging to zero.

4.1.1 Thompson sampling

We use the Thompson Sampling (TS) [34] algorithm to carry on the decision-making process. The TS algorithm is a Bayesian algorithm that selects a given action based on its past noticed performance. To do so, during the learning stage, TS observes the reward, and updates its prior belief in a way that the probability of a particular arm being optimal matches with the probability of each arm being selected. In practice, this is done by sampling each arm from its posterior distribution, and selecting the one that returns the maximum expected reward. This property will result very useful, allowing us to tackle the intrinsic non-stationarity of our environment. Hence, arms that were chosen initially because of their good rewards, can be discarded over time if they start to perform badly. Section 5.4 tackles this feature in-depth.

In our study, the prior belief on the rewards is assumed to be Gaussian distributed, as performed in [13]. Further details on the application of TS using Gaussian priors can be found in [35]. Under this model, TS takes a sample for each action (θ_x) according to a Gaussian distribution, which is provided by $\mathcal{N}(\hat{\mu}_x(t), \sigma_x^2(t))$, and so, selecting the action returning the maximum value of θ_x.

For the considered distribution, the mean and variance are calculated as

$$\hat{\mu}_i(t) = \frac{\sum_{w=1:i}^{t-1} r_i(t)}{n_i(t)+1} \quad , \quad \sigma_i^2(t) = \frac{1}{n_i(t)+1}$$

where $r_i(t)$ is the reward experienced for the action i until round t, and $n_i(t)$ is the number of times that the action i has been selected until round t. It is important to note that at the first TS iteration for each action will be given by a $\mathcal{N}(0, 1)$. The implementation of the TS algorithm can be found in Algorithm 1.

Recently, authors in [13] have proved that TS performs better than ϵ-greedy, upper confidence bound (UCB) and EXP3 in complex WiFi scenarios, as it results in faster convergence rates. In addition, empirical results from [36] demonstrate as well that TS outperforms UCB despite its simplicity.

Algorithm 1: Implementation of Thompson sampling.

Input : set of possible actions, $\mathcal{X} = \{x_1, ..., x_N\}$

Initialize: for each arm $x_i \in \mathcal{X}$, set $\hat{\mu}_i = 0$ and $n_i = 0$

while *active* **do**

  For each arm $x_i \in \mathcal{X}$, sample $\theta_i(t)$ from $\mathcal{N}(\hat{\mu}_i(t), \sigma_i^2(t))$

  Select arm $x_i = \underset{i=1,...,N}{\text{argmax}}\ \theta_i(t)$

  Observe and compute the reward experienced $r_i(t)$

  $\hat{\mu}_i(t) \leftarrow \frac{\hat{\mu}_i(t) \cdot n_i(t) + r_i(t)}{n_i(t)+2}$

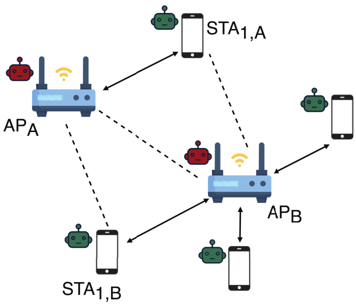
  $n_i(t) \leftarrow n_i(t) + 1$

end

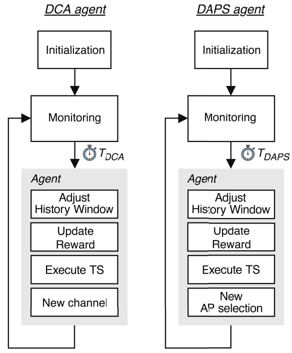
4.2 Decentralized DCA and DAPS: an adversarial MAB approach

In our study case, we consider that both DCA and DAPS problems fit into the adversarial MAB framework. The adversarial environment is developed through the fact that rewards experienced by actions depend on how the other players behave. So, from a point of view of single player, its opponents will have control over its rewards. Players are naturally classified into APs and stations, which rewards are defined accordingly to the metrics presented in Section 3.3. It is important to note that players belonging to different groups may interact during the decision-making process. Indeed, an interesting contribution of this work is to consider these interactions between two different types of players.

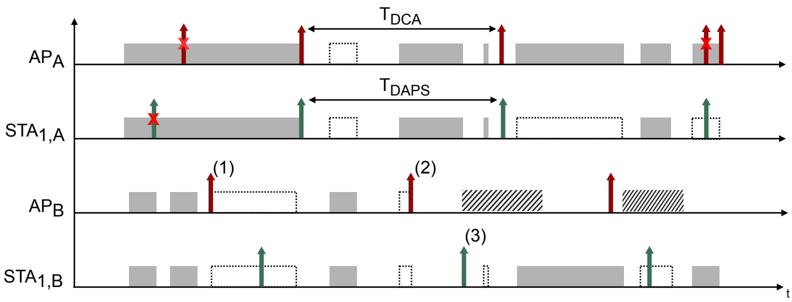
To further explain how agents perform, let us refer to Figure 2a, in which a simple WLAN scenario is represented. Here, we consider that APs have selected different channels, whereas the stations have been considered to be attached following the SSF mechanism. Besides, APs and stations are considered to have the DCA (coloured in red) and DAPS (coloured in green) agents, respectively. For instance, focusing on the AP$_B$, it may experience a bad channel reward if it suddenly changes to the channel used by its neighbor. However, this player is also conditioned by the fact that negative rewards can be produced either if its neighbor changes to its operating channel, or if all the stations select it as the serving AP. Therefore, under an adversarial setting, players' decisions are highly conditioned to the ones made by the other contestants. This can be also confirmed from a station's point of view, as it may experience low satisfaction values if it select an overloaded AP, if its serving AP changes to a congested channel, or if many other stations switch to its serving AP.



(a) Topology example



(b) Flow diagram



(c) Time line operation of DCA/DAPS agents. Colored boxes represent active flow transfers from APs to stations, whereas doted boxes represent flow transfers sensed over the medium. Finally, boxes represented with a diagonal pattern are addressed to stations not represented in the time line. Red and green arrows represent that a DCA or a DAPS agent has been triggered, respectively.

Figure 2: Representation of the DCA and DAPS agent operation

In order to learn, we consider APs and stations to be equipped with an agent, which purpose is to select the most appropriate channel and AP, respectively, based on previously gathered experience. We define the concept of agent as a learner that performs background tasks, such as data-driven decisions, on behalf of an AP or an station. In Figure 2b, it is shown the agents' operation cycle, which comprises two phases. After the node initialization, we find the network *monitoring* phase, in which the agent is intended to remain silent, observing and collecting the performance measurements for the last selected action. Although this phase does not involve the agent explicitly, it is considered part of its operation as it needs to observe the environment in order to acquire information. Then, once the timers, T_{DCA} and T_{DAPS}, are expired, the second phase starts. We call it new action selection phase, and it comprises the performance evaluation of the last selected action, which process is subdivided in four steps.

1. Adjust history window. Since agents keep track of all the data collected, they must get the past entries of the current action that fall inside the time boundaries specified by the sliding window size. By doing this procedure, agents average the performance observed during the *monitoring* phase among past records. In further Sections 4.2.1 and 5.4.2, the sliding window concept is explained and analyzed.

2. Update reward. Once agents obtain the average data, they must compute the parameters $\hat{\mu}_i(t)$ and $\sigma_i^2(t)$ for the current action, in order to update the probability distribution with the last performance data observed.

3. Execute TS. After updating the estimated parameters of the distribution, the TS algorithm is executed by drawing a sample $\theta_i(t)$ from $\mathcal{N}(\hat{\mu}_i(t), \sigma_i^2(t))$.

4. New channel/AP selection. The action returning the largest value of $\theta_i(t)$ will be selected as the new one. Then, agents will inform to APs and stations to update their configuration accordingly.

Finally, a representative illustration of the agents' time line operation is depicted in Figure 2c. For the sake of visualization we have only represented two of the different stations represented in Figure 2a. In the point marked as (1), we find that after a proper monitorization of the environment, AP_B's DCA agent is triggered, motivating a channel switch for AP_B. Immediately after this event, AP_B finds the medium busy as an ongoing data transfer is being carried by AP_A. It is interesting to note that during that transfer, the DCA agent of AP_A, as well as the DAPS agent of $STA_{1,A}$, should have been triggered. However, since the ongoing transfer was still active, their timers have been postponed until its ending. In addition, asynchronous operations of the agents are naturally supported, which implies that during the *monitoring* period (i.e., two consecutive inter DCA/DAPs agent activation epochs), we may observe other agents changing the configuration of their respective APs or stations. This effect is reflected in (2), in which AP_B, again, changes its channel. Finally, at (3) $STA_{1,B}$'s DAPS agent decides to change to AP_A in order to meet its traffic demands, decision that was caused by the unbalanced situation between APs.

4.2.1 Activation Timers and Sliding Window

Before getting into the performance analysis of the system under evaluation, we first assess the implications of the parameters used by the agents. So, we need to decide the rate at which agents will become active and choose a new action. From the point of view of learning agents, the more frequent the learning algorithm is executed, the faster will the learning process be. Thus, we need to choose the time interval between two consecutive agent activation epochs as short as possible to reduce the convergence time, while ensuring that network reconfiguration overheads can be assumed negligible. As we rely on IEEE 802.11r amendment, which reduces the roaming time between APs nearly to 50 ms, we consider that 3 minutes (180 s) is large enough and so, we set it as the time between two consecutive agent activation epochs for both DCA and DAPS.

On the other hand, the channel switch in WiFi networks is specified in the IEEE 802.11h amendment. This mechanism enables APs to announce a channel switch using channel switch announcement (CSA) frames before effectively moving to that channel [37]. When enabled, APs advertise through CSA frames the new channel, helping clients to switch to the target channel, saving scanning time. Therefore, clients, who support CSA, can perform the transition to the new channel with minimal downtime, instead of having to scan and discover the new channel in which their AP has switched. However, this process is not performed immediately as the AP sends a variable -*and normally vendor-dependant*- number of frames, which contain the CSA announcement. Thus, the delay of the channel switch depends on the number of CSA frames being broadcasted. In this work, we have considered that this process may last up to 100 ms. Therefore, it allows us to set the time between two consecutive agent activation epochs also to 3 minutes (180 s), as the impact of the channel switching frames can be considered negligible.

Another consideration when using learning algorithms is the age of information. It is important to identify the existing trade-off between still valid and outdated information. In very dynamic scenarios, keeping track of old observations can lead agents to take decisions based on information that is outdated. However, not considering enough past data will reduce the ability to select a proper new action, as agents may lose useful information. To tackle this trade-off, we use the concept of sliding window, which is intended to filter the useful information from the outdated one. We will further discuss the impact of the window size in Section 5.4.2. Unless otherwise stated, we set by default the window size to 9 minutes (540 s), which corresponds to three agent activation periods.

5 Performance evaluation

In this section, we test the DCA and DAPS under different density conditions and throughput requirements to evaluate the performance of the learning MABs. To perform the evaluation, we have implemented from scratch our own simulator in C++ using the COST simulation libraries [38], which works as presented in

Table 2: Simulation parameters

Parameter	Description	Value
f_c	Central frequency band	Depends on channel selection
W	Channel bandwidth	20 MHz
P_{tx}	Transmission power	15 dBm
G_{tx}	Antenna transmission gain	0 dB
G_{tx}	Antenna reception gain	0 dB
PL (d)	Path loss	Refer to A
CCA	CCA threshold	-80 dBm
$\mathrm{RSSI_{th}}$	RSSI threshold	-75 dBm
$\mathrm{N_{ss}}$	Number of spatial streams	2
L_{d}	Data packet size	12000 bits
$\mathrm{T_{on}}$	Avg. connection duration	1 s
$\mathrm{T_{off}}$	Avg. connection interarrival time	3 s
$\mathrm{CW_{min}}$	Min. contention window	16
p_e	Packet error rate	0.1
$\mathrm{T_{sim}}$	Simulation time	86400 s
$\mathrm{T_{DCA}}$	Time to trigger AP activation agent	180 s
$\mathrm{T_{DAPS}}$	Time to trigger station activation agent	180 s
$\mathrm{T_{sw}}$	Sliding window interval	540 s
$\mathrm{P_{th}}$	Performance threshold	85%

Section 3. The simulation platform developed is called Neko, and it can be found on GitHub[4]. The main reason that has led us to develop our own simulation tool is the need to simulate large scale networks for long periods of time.[5] To achieve that goal, Neko considers the CSMA/CA abstraction described in Section 3.2.

5.1 Simulation set-up

All considered scenarios consist of multiple IEEE 802.11ax-capable APs and stations. All of them include capabilities of the IEEE 802.11k and 802.11r amendments, as mentioned in Section 3. In all scenarios the transmission power of APs

[4]`https://github.com/wn-upf/Neko`

[5]As an example, it takes 6 minutes to simulate a medium-large scale network (100 APs and 1000 stations) in the Neko platform for a simulation time of 1 day, over an average quad-core Intel i5 3.8 GHz processor. On the contrary, in simulation platforms such as network simulator (ns) 3, it takes roughly 3 h to run a 1 minute simulation for a moderate scenario with 20 APs and 200 stations.

and stations is set to 15 dBm. In all cases, we guarantee that stations detect at least one AP (i.e., the received power is higher than the CCA threshold), as otherwise, the station is re-located. The expected duration of the downlink traffic flows, and the expected time between two consecutive flows is given by T_{on} and T_{off}, respectively. Transmission rates between stations and their serving APs are determined by the RSSI of the link. We have selected the IEEE 802.11ax path-loss model for enterprise scenarios [39], which can be found in the A. The main reason to select this path-loss model is that it takes into account the effect of walls, as the 5 GHz band is very sensitive to this parameter. Finally, we have provided APs with 3 different channels of the UNII-1 band. We have constrained the channel availability to numbers 36 (5.18 GHz), 40 (5.20 GHz) and 44 (5.22 GHz) in order to reduce the action set for the APs, as well as to get the most of the spectrum reuse. In Table 2, there are detailed the other parameters considered to obtain the results. Finally, physical and MAC layer parameters of the IEEE 802.11ax standard are shown in A.

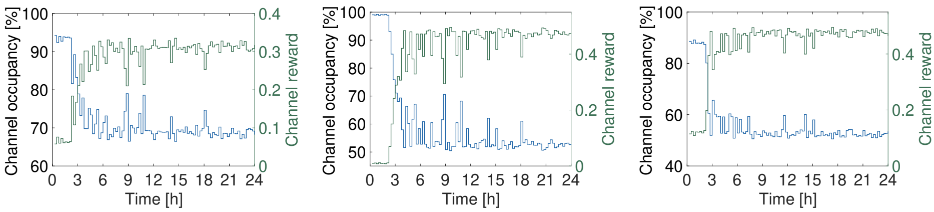
5.2 Toy scenario

For this illustrative use case, we have designed a controlled environment to study the interactions and behavior of the network when applying DCA and DAPS independently. To this purpose, we have deployed three APs in a line with partial overlapping coverage areas, and 45 stations, which have been distributed uniformly in a 3D-space with dimensions 25 x 25 x 2 m (x, y and z axes, respectively). The throughput required by each station is randomly chosen in the range [1-5] Mbps every time a new flow starts (i.e., when the station moves to the 'on' state) in order to tackle standard traffic demands. For instance, around 5 Mbps is the recommended bandwidth for high definition video quality. Regarding the action space for stations, it is important to remind that each station will construct its action set independently of the others, as it depends on the number of APs sensed over the $RSSI_{th}$. Figure 5a, shows the considered deployment and how the different APs and stations are placed. Note that the same color scheme in the APs indicate that they have been configured with the same channel, whereas different colors will indicate different channels. For this controlled use case, as the number of APs is very low, we have limited the use of channels to numbers 36 and 40. Finally, the simulation parameters correspond to the ones presented in Table 5.1.

5.2.1 DCA evaluation

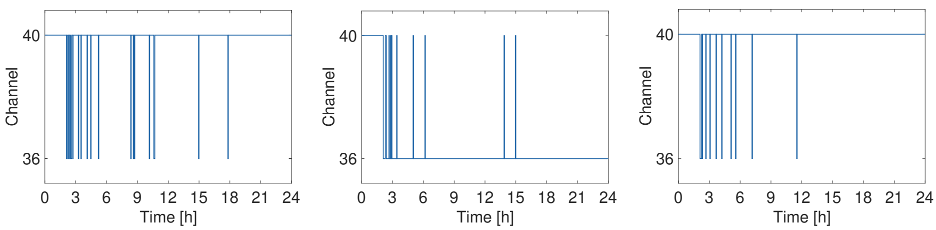
Firstly, we are going to evaluate the effect of applying the DCA mechanism in the network performance. To do so, we configure the three APs with the same radio channel, enabling the DCA agent on them. On the contrary, all the stations do not have their DAPS agent enabled, so we can analyze the implications of employing the DCA alone.

Figure 3 shows the occupancy of each AP. From $t = 0$ h and $t = 2$ h, we observe the first static stage in which APs remain with the initial configuration (i.e., all APs are configured with the same channel). Then, from $t = 2$ h and $t = 8$ h APs' DCA agents start the learning stage until they converge into a solution. From Figures 3a and 3b, for both AP_1 and AP_2, we can see two peaks at $t = 9$h that indicate a bad exploration from one of the two APs. However, this effect is not reproduced in AP_3, as seen in Figure 3c, due to AP_1 being out of range from the CCA threshold of AP_3. To analyze the consequence behind these peaks, we have represented the action evolution for all three APs. In Figure 4, it can be observed that AP_1 is the cause of the aforementioned peaks, since its channel switch from 40 to 36 downgrades not just its own reward, but AP_2's reward too. Moreover, we can see that during the learning stage, the TS algorithm constructs its probabilistic model as the time advances by exploring different actions. Consequently, at the end, exploring is less frequent and exploitation is almost exclusively performed. It is worth noticing that channel reward is inverse to the AP occupancy as stated in Section 3.3.1.



(a) AP1 (b) AP2 (c) AP3

Figure 3: Evolution of the APs' performance. In blue is represented the channel occupancy, whereas in green is represented the channel reward



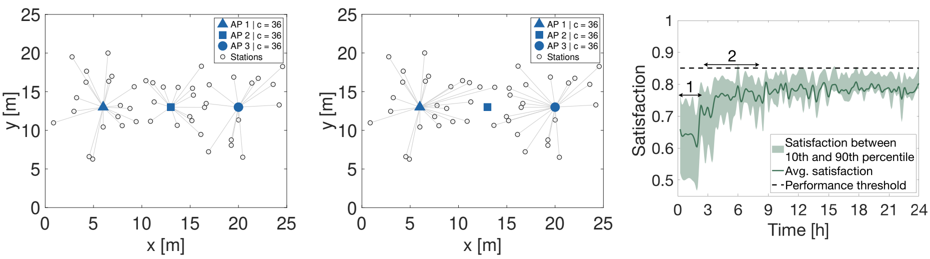
(a) AP1 (b) AP2 (c) AP3

Figure 4: Action evolution for the APs

5.2.2 DAPS evaluation

Now, we evaluate the effect of the DAPS mechanism. To do so, again, we configure the three APs with the same radio channel, but deactivating their agents. Now, all the stations have their DAPS agent activated. The initial association at the beginning of the simulation is done through the SSF mechanism. Before running the simulation, we can observe from this configuration that AP_2 will suffer from starvation, and AP_1 and AP_3 will be able to allocate most of the available airtime to their stations. Therefore, stations attached to AP_2 will be encouraged to re-associate as they will receive a poor satisfaction value.

Figure 5b shows the final association scheme of the scenario, and as expected, the stations that have perceived low satisfaction values leave their serving AP, in order to associate to a different AP that ensures a higher airtime allocation, even if that means using lower transmission rates. Figure 5c shows the average satisfaction of all the stations. It can be observed that during the beginning of the simulation (stage marked as 1) the traditional association performed badly. However, at $t = 2$ h, the DAPS agent is activated, and stations are allowed to explore different APs during the learning stage (marked as 2). After some time, stations converge into a solution, even though it is below the performance threshold set (Table 2). Therefore, we consider that in this case, using only the DAPS agent at the stations, the network is not able to reach a feasible solution. This effect is related to the fact that APs sharing the same radio channel prevent any feasible re-association option. Then, we can conclude that the performance of the DAPS mechanism is severely conditioned to the efficiency of the DCA mechanism to properly allocate orthogonal channels to overlapping BSSs. However, we find that the use of DAPS can be useful to overcome an unbalanced distribution of the stations.



(a) Initial association (b) Final association (c) Satisfaction evolution

Figure 5: Performance of DAPS in the toy scenario

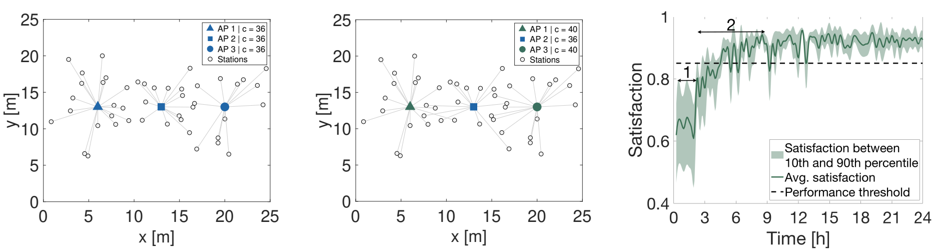
5.2.3 Concurrent DCA and DAPS

Finally, we consider the case in which both agents operate concurrently in the toy scenario. Again, the initial channel allocation is the same as for all APs, whereas stations are associated following the SSF mechanism. By evaluating the concurrent operation of DCA and DAPs, we expect to see the effects over the network of both DCA and DAPS agents, so we can observe the potential advantages of running them at the same time.

Figure 6b shows the final result. We can observe that the concurrent execution of both DCA and DAPS have accomplished the expectations. We observe that APs have been reconfigured into a feasible solution. AP_1 and AP_3 have been allocated with the same channel, as they are out of their CCA range, whereas AP_2 has been allocated with a different one. Therefore, DCA agents overcome the flow starvation effect that, initially, AP_2 was suffering. On the other hand, the effect of the DAPS agents can be observed over the new distribution of the stations. However, the relevance of the DAPS in this scenario is significantly lower, than the DCA, due to the fact that only few stations have been reallocated. Figure 6c shows the satisfaction evolution for all the stations, in which it is represented the mean value, as well as the 90th and 10th percentile of the measurements that are represented by the upper and lower bounds of the shaded area. From the figure, we observe that now the average satisfaction surpasses the performance threshold, which confirms that the efficiency of the solution remains mostly in the DCA mechanism to properly allocate orthogonal channels to overlapping BSSs.

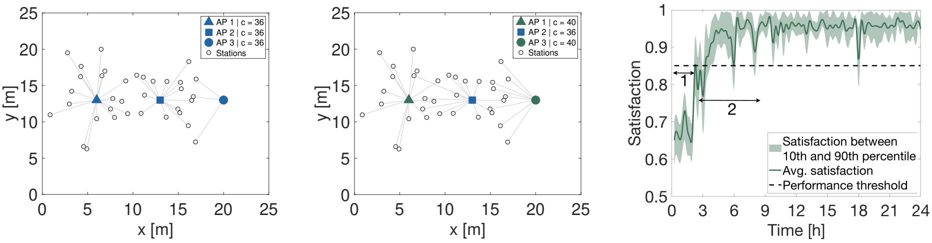
Although some random bad performances around $t = 9$ h and $t = 12$ h can be observed in Figure 6c, the network converges into a solution approximately at time $t = 5$ h. This fact is quite relevant as we can see that by applying both mechanisms we can get a considerable gain over the static approach (i.e., performance observed during the stage marked in point 1).

Finally, we assess the concurrent operation over an unbalanced scenario to



(a) Initial scheme (b) Final scheme (c) Satisfaction evolution

Figure 6: Joint performance of DCA and DAPS



(a) Initial scheme (b) Final scheme (c) Satisfaction evolution

Figure 7: Joint performance of DCA and DAPS in an unbalanced scenario

properly evaluate the effects of the DAPS in such conditions. Results are shown in both Figure 7a and Figure 7c. Again, we observe that DCA agents have accomplished their tasks. However, now, we are more interested to evaluate the performance of the DAPS agents. From the figure, we observe that the unbalanced situation has been correctly mitigated, as the station distribution is fairer among the different APs. Specifically, we can observe that AP_3 has doubled the number of stations attached, which from AP_2 to AP_3. At the same time, such effect, has caused that stations from AP_1 migrate to AP_2 to further balance the number of stations associated to every AP. Here, it is presented a clear example in which actions from a set of players may condition decisions taken by others.

5.3 Large random scenarios

Once we have seen the implementation and benefits of applying the DCA and DAPS mechanisms in a controlled scenario, the objective now is to evaluate their capabilities in multiple scenarios randomly generated. Such evaluation will prove whether the DCA and DAPs mechanisms can help to improve the overall network performance. We aim to assess how DCA and DAPS react for different throughput requirements, number of APs and number of stations.

To proceed with the evaluation, we increment the 3D area considered previously to 30 x 30 x 2 m. We have simulated 100 different scenarios, either for the static approach (i.e., using always the initial configuration), and when DCA and DAPS are applied concurrently. All simulations represent 1 day of virtual time (i.e., 86400 seconds). It is worth mention that, at the start of each simulation, APs select their initial channel in a random fashion, and stations are attached to APs based in the SSF criteria. Besides, no channel restriction is placed, so APs can select channel numbers 36, 40 and 44. Both APs and stations are agent-enabled. The rest of the simulation parameters remain the same as the presented in Section 5.1.

In terms of performance metrics, we compute the average satisfaction achieved for all the stations during the simulation time for each one of the 100 scenarios. Then, we analyze the distribution obtained from the 100 average satisfaction values obtained. The representation of the obtained average results is done through box plots, in which the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles. Whiskers extend to the most extreme data points that are not considered outliers, whereas the outliers are plotted using the 'o' symbol.

5.3.1 Throughput requirements

We first study the performance of both techniques by considering different stations' downlink throughput requirements. To do so, we fix the network density to 15 APs and 225 stations, and consider four different throughput ranges: [1-3] Mbps, [1-5] Mbps, [1-7] Mbps and [1-9] Mbps. Figure 8 shows the obtained results for the satisfaction, aggregate throughput and the throughput drop ratio (i.e., the percentage of traffic that cannot be served) metrics. Note that in Figure 8, the x-axis represents the average required throughput per station, i.e.,

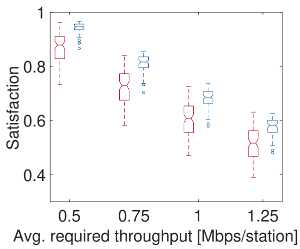
$$\overline{B} = \mathbb{E}[B]\left(\frac{T_{\text{on}}}{T_{\text{on}} + T_{\text{off}}}\right),$$

where $\mathbb{E}[B]$ is the average value of the chosen range.

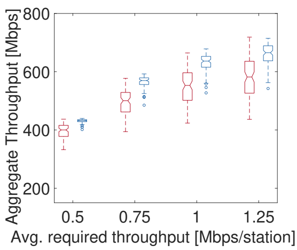
Comparing the adaptive MABs approach with the static one, we can observe that both schemes are very sensitive to the stations' required throughput. For instance, regarding the satisfaction achieved, we find that for both cases it becomes lower when the throughput range increases. However, taking the median value as a reference, we see that the DCA and DAPS perform 10% better regardless the stations' required throughput. In addition, we can see a clear tendency in the throughput values when assessing the static and traditional approach. It can be appreciated that from the 0.75 Mbps/station to the 1.25 Mbps/station cases, the whiskers of the box encompass a larger range of values, clearly indicating a high variability in the data, as results are very sensitive to the specific topology of each scenario. Comparing the DCA and DAPs performance against the static scheme, we find that this dependency on the scenario's topology, and so the variance in the results, is highly reduced using the adaptive MABs approach. For instance, we observe a variance reduction of 60% between the 25th and 75th percentile for the 1 Mbps/station case. In fact, we want to make special mention to the good performance of our mechanisms regardless the throughput demands, in which the MABs strategy outperform the static scheme reducing the variability, as all the obtained values are closer to the median value. Moreover, results show that in high traffic conditions, the only solution remains in densify the network with more APs, since the throughput drop ratio values indicate that almost half of the airtime required can not be allocated properly.

5.3.2 APs densification

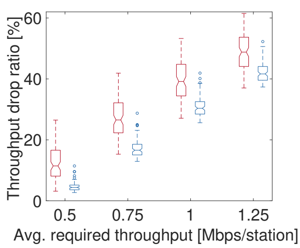
As we have seen, high traffic demands can prevent the network to deliver the required service. In this section, we investigate if adding more APs to the network may contribute to improve the network performance. To do so, we keep the same number of stations (225 stations), while traffic demands from stations are pick from the range [1-5] Mbps. Figure 9 show the obtained results. We can observe how adding more APs improve the network performance, as the uptrend evolution of the satisfaction and throughput values indicate. Also, notice in Figure 9a that the gains of adding new APs when applying the DCA and DAPS strategies get lower at each step, in special from the 20 APs case to the 25 APs, which is about 1%. This effect is an indicator that the network will barely improve even though more APs will be placed. In order to overcome this effect, APs should be strategically placed in order to be detected above the $\mathrm{RSSI}_{\mathrm{th}}$, so the DAPS mechanism can be triggered, and an effective user re-association produced. Although network densification can be a good solution to tackle network congestion, Figure 9c shows that it does not improve the performance by itself. Analyzing the throughput drop ratio in the case of 25 APs, we notice that the static approach still performs badly since the 75th percentile of the measurements surpasses a value of 20%. This effect is associated to the topology dependency mentioned before, as overlapping APs with the same channel configuration are more likely to happen. In this type of scenarios, the benefits of using the DCA mechanism are very relevant, since the network is able to better manage the spectrum re-



(a)



(b)



(c)

Figure 8: Increasing throughput requirements case. From right to left, we find the satisfaction, aggregate throughput and the throughput drop ratio, respectively. In red is shown the joint SSF and static channel allocation schemes, whereas in blue are represented the results of applying DCA and DAPS mechanisms.

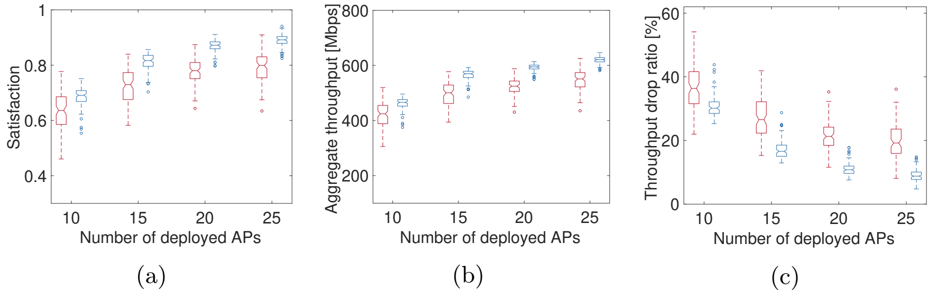


Figure 9: AP densification case. From right to left, we find the satisfaction, aggregate throughput and the throughput drop ratio, respectively. In red is shown the joint SSF and static channel allocation schemes, whereas in blue are represented the results of applying DCA and DAPS mechanisms.

sources, achieving a better performance, and very low variance between different scenarios.

5.3.3 Station density

In the following, we study the performance of the DAC and DAPS with different number of stations. We have evaluated scenarios with 75, 150, 225 and 300 stations, keeping the station/AP ratio to 15. Thus, for each density we will have 5, 10, 15 and 20 APs, respectively. Traffic demands per station remain in the range [1-5] Mbps. Figure 10 shows the results obtained. As it can be seen, the learning approach remains as the best solution in terms of performance. Despite

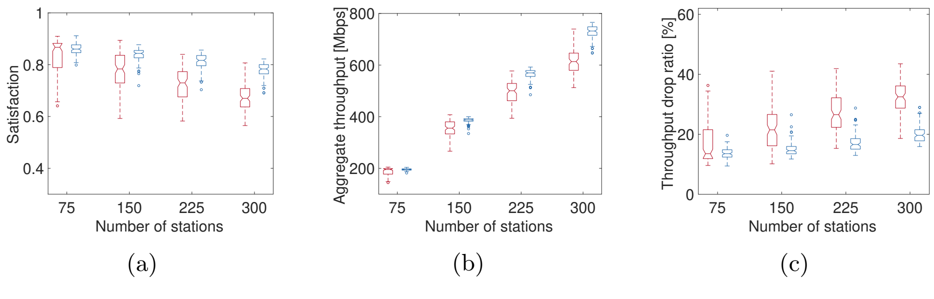


Figure 10: Variable AP and user density case. From right to left, we find the satisfaction, aggregate throughput and the throughput drop ratio, respectively. In red is shown the joint SSF and static channel allocation schemes, whereas in blue are represented the results of applying DCA and DAPS mechanisms.

the downtrend of the boxes, the difference between the static and the adaptive MABs approach gets higher, as the network gets denser. For instance, for 150 stations the gain is around 7%, whereas for the 225 and 300 stations the gain gets up to 11% and 16%, respectively. This effect shows us that the adaptive MABs approach is capable to support a larger number of users before downgrading significantly the performance. Again, the low variance results observed in all figures show that in a wide diversity of scenarios the output values are very constant in front of the static approach, which presents a high variability even in scenarios with few APs as indicated by the larger whiskers range. Finally, it can be observed that, when applying the DCA and DAPS, agents are able to successfully learn even in complex and challenging scenarios, such as the case of having 20 APs and 300 stations.

Finally, Figure 11 shows the empirical cumulative distribution function (CDF) of the convergence time for the different considered scenarios. We define the convergence time in each individual scenario as the instant of time in which the median value of the satisfaction gets above the performance threshold, P_{th}. As it is shown, the results display that there exist a temporal dependence between convergence and network density. Thus, for 75, 150, 225 and 300 stations, considering the time at which the 80 % of the individual scenarios have converged, we find out that their convergence times are 1, 7, 9, and 13 hours, respectively. Therefore, the denser the network, the more time the network will need to reach a solution. As we consider scenarios in which mobility is not significant, such as office buildings, where stations remain quiet during long periods of time, we consider that the obtained times are low enough to be acceptable in practice.

Nevertheless, it is unrealistic to think that a network will remain static for very long periods of time (i.e., more than 10h). In order to overcome such a requirement, by identifying the periods of maximum traffic load (i.e., the busy hour), and enabling the learning agents only during these periods of time through multiple days, we could expect to obtain similar results as shown here.

5.3.4 Agent Action-Selection timers

We finally asses the implications of varying both T_{DCA} and T_{DAPS} timers in the network response, in order to study the behavior of the system when the agents are executed more and less frequently. First, we tackle the AP selection case alone, fixing T_{DCA} to 180 s, and testing different values for T_{DAPS}. Again, simulations comprise 100 random scenarios, in which the network consists of 15 APs and 225 stations. The traffic load of each flow are kept in the range of [1-5] Mbps. As in previous sections, the simulation time is set to one day (84600 s). The agents are activated at $t = 3\ h$. From the results obtained, although not shown here, we observe that the DAPS timer does not have a significant impact on the performance of the stations as we only experience differences of 2 % between

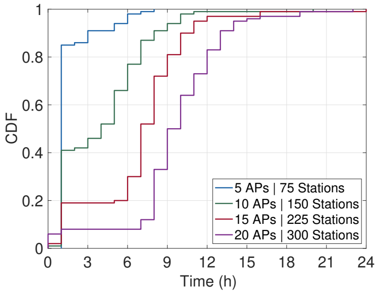


Figure 11: Empirical CDF for different network station density cases

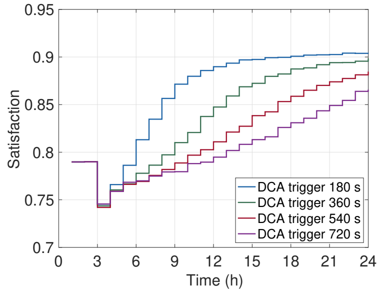


Figure 12: Evolution of the satisfaction for different DCA timer values

different timer configurations.

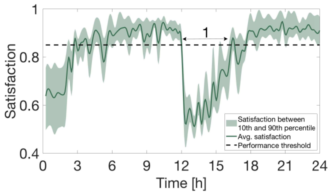
In the same conditions as in previous case, now, we vary the T_{DCA}, while keeping T_{DAPS} to 180 s. In this case, the results are shown in Figure 12, where the average satisfaction evolution is presented for different T_{DAPS} values. We tested values of 360 s, 540 s and 720 s, which correspond to 2, 3 and 4 times the default T_{DAPS} value of 180 s. We observe that for small values of T_{DAPS} convergence is reached in much less time that if timers are set to higher values. Therefore, we observe that there is no gain by increasing the value of the timers. Then, we suggest to set those timers to the minimum possible value given the reconfiguration overheads are kept negligible.

5.4 Dynamic environment

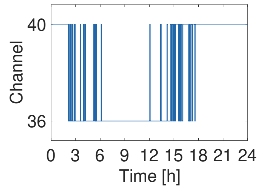
5.4.1 A sudden channel change

So far, network conditions remained static, meaning that no variation or anomaly was introduced in the network. However, as a final use case, we want to analyze the behavior of the learning agents in a changing environment. For the sake of practicality, we evaluated this scenario in the controlled deployment of the toy scenario used in Section 5.2. In this case, we consider that the whole set of stations are agent-enabled, whereas AP_1 and AP_3 are agent-enabled but AP_2 is not. As in the toy scenario, we use only channels 36 and 40.

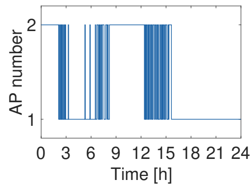
In order to assess changing conditions, we initially configure the three APs in the same radio channel, expecting that AP_1 and AP_3 will learn to reconfigure themselves as the simulation goes by. At $t = 12$ h, we observe that AP_1 and AP_3 have chosen the channel not used by AP_2. Then, we trigger a channel reconfiguration on AP_2, resulting in AP_1 and AP_3 to start exploring again the other channel



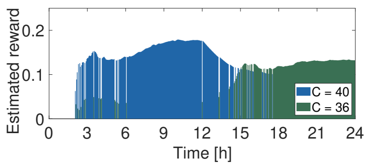
(a) Stations satisfaction



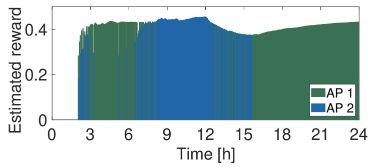
(b) AP_3 channel selection



(c) STA_2 AP selection



(d) AP_3 estimated reward



(e) STA_2 estimated reward

Figure 13: Satisfaction, rewards, and channel and AP selected by AP 3 and station 2, respectively, when assessing changing conditions.

as the reward of the currently selected channel decreases. Figure 13a shows the average satisfaction experienced by the stations. At $t = 2$ h the learning procedure starts. We can observe how rapidly the network converges to a feasible solution. However, at $t = 12$ h, we trigger the AP_2 reconfiguration from channel 40 to channel 36, remaining in this channel until the end of the simulation. As a result, both agent-enabled APs, AP_1 and AP_3, and stations start to perceive low rewards, and so they begin to explore the different available actions again. Therefore, channel 40, which was discarded at the early stages of the simulation, is chosen as the preferable one now. We can see that in the stage marked as (1) the second learning phase happens, and the decision-making process evolves satisfactory. After a while, convergence is achieved and the satisfaction is stabilized over the performance threshold.

For a more specific overview of the process, refer to Figures 13d and 13e, where we have represented the estimated reward evolution experienced by AP_3 and STA_2, respectively. We have selected these two nodes as they can give us better insights on how the learning process is performed once the AP_2 changes its channel. Considering first Figure 13d, we observe that AP_3 starts to explore the available channels, reaching a solution by selecting channel 36. However, at the half of simulation, due to the overlap caused by AP_2 being reconfigured to channel 36 too, AP_3's reward starts to decay. Consequently, AP_3 explores again, and the action discarded at the beginning starts now to receive higher rewards, and so become more likely. Therefore, AP_3 switches to channel 40, moving to

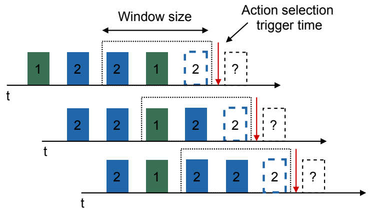


Figure 14: Sliding window time line representation.

the new expected solution.

On the other hand, Figure 13e depicts the behavior of STA_2. We observe that, when the learning stage finishes, the AP selected through the SSF criteria (AP_2) is discarded in order to be attached to AP_1. However, we can observe that even before $t = 12$ h, this station starts a learning phase again. This particular effect is caused due to other stations lasting more time to finish their learning process, which lastly may cause AP_1 to be saturated. As a result, STA_2 learns and attaches itself again to AP_2, since it provides better rewards due to the fact that at $t = 9$ h it is configured with a different channel than its neighbors AP_1 and AP_3. Finally, at $t = 12$ h, and since the AP_2 is forced to change the channel, STA_2 explores again, selecting AP_1 as final AP.

5.4.2 Age of data and sliding window

Through this paper, we have considered a reinforcement learning approach in which agents' decisions are based on the previous selected actions and their performance. In wireless environments, such a procedure requires to carefully keep track of past records as network conditions may change at a fast pace. Differentiate between valid and obsolete data is necessary in order take actions according to the current state of the network. So, old observations can lead to perform bad decisions, but considering only recent data may result in losing useful information. Then, the age of data becomes a variable that must be included into the decision making process.

To tackle this trade-off, we applied the concept of sliding window, which is intended to filter useful information from the outdated one. Particularly, the sliding window operation consists in a time interval that moves along with the simulation time, so performance records outside the window are not considered when evaluating the performance of an action. Using it, we allow agents to filter data, as well as to control and prevent bad decisions due to sporadic actions from other nodes. Figure 14 shows the sliding window feature. Once the trigger time (i.e., T_{DCA} or T_{DAPS}) is finished, agents will update, for the last action taken (depicted with a dashed line in the figure), the parameters $\hat{\mu}_x(t), \sigma_x^2(t)$, so the TS updates the distribution of the action. Only past records inside the window boundaries are used to update such parameters. A new action will be selected by drawing new values from $\mathcal{N}(\hat{\mu}_x(t), \sigma_x^2(t))$ and selecting the arm returning the higher value.

We have studied the implications of the sliding window size in the decision making process. We carried out this study under the changing environment conditions described in Section 5.4.1, as it includes both stationary and non-stationary changes in the environment. Figure 15 shows the estimated reward evolution of AP_1 for different sliding window sizes.

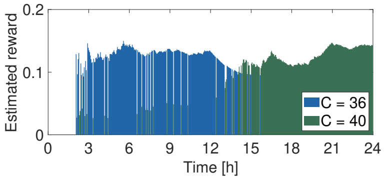
From the figures presented, we can see different behaviors as the sliding window interval increases. First, notice the sharper shape caused by having small size windows. This effect is related to the fact that the agents are reacting fast to changes from other players, since the reward evaluation is only averaged over a small set of reward entries. Then, it makes agents to become more vulnerable to others' decisions, and any random exploration by an agent may lead to an action change in all the others. Indeed, this issue can be corroborated on Figure 15a as frequent explorations are performed even after reaching a convergence state at $t = 6$ h. On the contrary, larger windows help to control and minimize the impact of the other agents on its own behavior, as it can be seen in Figure 15d, in which smoother transitions reveal that agents are more robust against sporadic changes.

Although having a larger window size helps out agents to overcome the case of intermittent bad performances, it costs agent reaction time. We refer to the agents' ability to detect and avoid an action that has been repeatedly performing bad. Then, setting a conservative approach in order to provide robustness to agents may lead to unfeasible large reconfiguration times. This issue is shown in Figure 15c and Figure 15d, in which agents require much more time to change to a better action. For instance, from $t = 12$ h, we find that for the 900 s and 3600 s cases, changing an action requires up to 6 h and 12 h, respectively. Therefore, we observe that there exists a clear trade-off between robustness and effective learning.

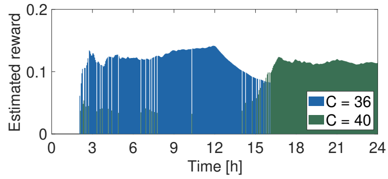
From all the simulations done, we observe that a window size of 540 s works well in the scenarios considered in this paper. As it can be seen in Figure 15b, exploration stages are barely performed after reaching convergence, and a quick reaction time (i.e., 3 h after $t = 12$ h) is registered.

6 Conclusions

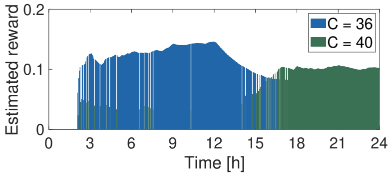
In this work, we have evaluated the implications of introducing learning algorithms for dynamic network adaptation. By means of the multi-armed bandits



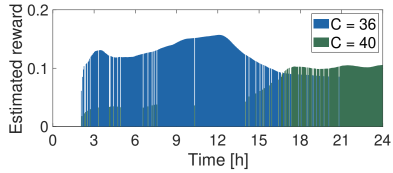
(a) Sliding window interval 60 s



(b) Sliding window interval 540 s



(c) Sliding window interval 900 s



(d) Sliding window interval 3600 s

Figure 15: Estimated reward evolution for AP_1 when considering different time windows.

framework, we asses the concurrent decentralized channel and AP selection by enabling agent-empowered APs and stations, which learn by interacting with the environment. Through simulations, we provide insights on how MABs perform in high dense deployments, and their potential use in large enterprise WLANs. In this context, we have seen that DCA and DAPS improve the channel utilization and fairness. In addition, we have found that the DCA has a high but coarse impact on the network performance, whereas the DAPS mechanisms allow stations to eradicate unbalanced situations due to SSF association criteria, so fine tuning the network performance. Then, we can state that an effective AP selection scheme must consider and evaluate, not only link quality, but load metrics in order to ensure a good network performance. In addition, since channel interference is a determinant issue, a proper channel allocation mechanism must be executed along with the AP selection, as the impact in the performance highly depends on it. Additionally, obtained results have shown that the application of learning agents highly improves and minimizes the topology dependency that static strategies suffer, as the dynamic agents adapt the network configuration to the observed conditions.

We conclude that MAB-enabled agents work well under the presented conditions, showing the potential of ML mechanisms to significantly improve the network operation. Moreover, we have shown that an agent-enabled network using ML is capable of solving anomalies and adapting itself in base of the metrics presented. As future work, we will asses the potential implications of machine learning for non-static scenarios, as well as the use of learning algorithms in a central network controller, where a global view of the network state can be used to further enhance the system performance. In this regard, we are interested in the study of the trade-off between centralized and decentralized operations, as well as on the limitations of learning algorithms for optimization purposes.

Appendix A

We assume that all stations and APs operate using the IEEE 802.11ax amendment. The PHY and MAC parameters considered during the simulations are presented in Table 3. We compute the duration of both data and control packet transmissions as detailed below. It is worth mention that we do not consider packet aggregation, and only one spatial stream per station is employed.

Table 3: Simulation parameters

Parameter	**Description**	**Value**
t_e	Empty slot duration	9 μs
t_{SIFS}	SIFS duration	16 μs
t_{DIFS}	DIFS duration	34 μs
t_{PHY}	Legacy preamble	20 μs
$t_{\text{PHY}-\text{HE}_{\text{su}}}$	HE single-station preamble	164 μs
σ_{leg}	Legacy OFDM symbol duration	4 μs
σ	OFDM symbol duration	16 μs
L_{SF}	Service field length	16 bits
L_{RTS}	RTS packet length	160 bits
L_{CTS}	CTS packet length	112 bits
L_{D}	Data packet size	12000 bits
L_{MH}	MAC Header length	320 bits
L_{ACK}	ACK packet length	112 bits
L_{TB}	Tail bits length	18 bits

$$t_{\text{RTS}} = t_{\text{PHY}} + \left\lceil \frac{L_{\text{SF}} + L_{\text{RTS}} + L_{\text{TB}}}{L_{\text{DBPS}}(\gamma_{i,j})} \right\rceil \sigma_{\text{leg}}$$

$$t_{\text{CTS}} = t_{\text{PHY}} + \left\lceil \frac{L_{\text{SF}} + L_{\text{CTS}} + L_{\text{TB}}}{L_{\text{DBPS}}(\gamma_{i,j})} \right\rceil \sigma_{\text{leg}}$$

$$t_{\text{DATA}} = t_{\text{PHY}-\text{HE}_{\text{su}}} + \left\lceil \frac{L_{\text{SF}} + L_{\text{MH}} + L_{\text{d}} + L_{\text{TB}}}{L_{\text{DBPS}}(\gamma_{i,j})} \right\rceil \sigma$$

$$t_{\text{ACK}} = t_{\text{PHY}} + \left\lceil \frac{L_{\text{SF}} + L_{\text{ACK}} + L_{\text{TB}}}{L_{\text{DBPS}}(\gamma_{i,j})} \right\rceil \sigma_{\text{leg}}$$

where $L_{\text{DBPS}}(\gamma_{i,j})$ is the number of bits in each OFDM symbol, which in fact depends on the MCS accordingly selected to the RSSI value received ($\gamma_{i,j}$) for the link pair of station i and AP j. Thus, $L_{\text{DBPS}}(\gamma_{i,j}) = N_{sc}N_mN_cN_{ss}$ where N_{sc} is the number of data sub-carriers, N_m is the number of bits per modulation symbol, N_c is the coding rate and N_{ss} is the number of spatial streams. In addition, note that control frames are transmitted in legacy mode using the lowest rate at MCS 0, and therefore being $L_{\text{DBPS}} = 24$ bits.

Regarding the path loss, we have selected the IEEE 802.11ax enterprise model described in [39], since we are considering a single floor environment. The path loss between a station i and AP j is given by

$$\begin{aligned}\mathrm{PL}(d_{i,j}) = 40.05 + 20\log_{10}\left(\frac{f_c}{2.4}\right) + 20\log_{10}(\min(d_{i,j}, d_{bp}))\\ + (d_{i,j} > d_{bp}) \cdot 35\log_{10}\left(\frac{d_{i,j}}{d_{bp}}\right) + 7W_{i,j}\end{aligned}$$

where f_c is the AP's central frequency in GHz, $d_{i,j}$ is the distance between station i and AP j in meters, d_{bp} is the breaking point distance in meters, and $W_{i,j}$ are the number of traversed walls. We set the breaking point distance, d_{bp}, to 5 m and the number of traversed walls, $W_{i,j}$, to 4. Note that the resulting propagation losses are given in dB.

Acknowledgments

This work has been partially supported by the Spanish Ministry of Economy and Competitiveness under the Maria de Maeztu Units of Excellence Programme (MDM-2015-0502), by the Spanish Government under grant WINDMAL PGC2018-099959-B-I00 (MCIU/AEI/FEDER,UE), by the Catalan Government under grant 2017-SGR-1188, and by a Gift from the Cisco University Research Program (CG#890107, Towards Deterministic Channel Access in High-Density WLANs) Fund, a corporate advised fund of Silicon Valley Community Foundation.

References

[1] VNI Cisco. Cisco visual networking index: Forecast and trends, 2017–2022. *White Paper*, 2018.

[2] Eng Hwee Ong, Jarkko Kneckt, Olli Alanen, Zheng Chang, Toni Huovinen, and Timo Nihtilä. Ieee 802.11 ac: Enhancements for very high throughput wlans. In *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 849–853. IEEE, 2011.

[3] Boris Bellalta. Ieee 802.11 ax: High-efficiency wlans. *IEEE Wireless Communications*, 23(1):38–46, 2016.

[4] M Shahwaiz Afaqui, Eduard Garcia-Villegas, and Elena Lopez-Aguilera. Ieee 802.11 ax: Challenges and requirements for future high efficiency wifi. *IEEE Wireless Communications*, 24(3):130–137, 2016.

[5] 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.

[6] David López-Pérez, Adrian Garcia-Rodriguez, Lorenzo Galati-Giordano, Mika Kasslin, and Klaus Doppler. Ieee 802.11 be extremely high throughput: The next generation of wi-fi technology beyond 802.11 ax. *IEEE Communications Magazine*, 57(9):113–119, 2019.

[7] Dusit Niyato and Ekram Hossain. Cognitive radio for next-generation wireless networks: An approach to opportunistic channel selection in ieee 802.11-based wireless mesh. *IEEE Wireless Communications*, 16(1):46–54, 2009.

[8] Yuhua Xu, Alagan Anpalagan, Qihui Wu, Liang Shen, Zhan Gao, and Jinglong Wang. Decision-theoretic distributed channel selection for opportunistic spectrum access: Strategies, challenges and solutions. *IEEE Communications Surveys & Tutorials*, 15(4):1689–1713, 2013.

[9] Arunesh Mishra, Vivek Shrivastava, Dheeraj Agrawal, Suman Banerjee, and Samrat Ganguly. Distributed channel management in uncoordinated wireless environments. In *Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 170–181. ACM, 2006.

[10] Marko Hoyhtya, Sofie Pollin, and Aarne Mammela. Classification-based predictive channel selection for cognitive radios. In *2010 IEEE International Conference on Communications*, pages 1–6. IEEE, 2010.

[11] Wang Lingzhi, Hua Cunqing, Zheng Rong, and Ni Rui. Online channel selection and user association in high-density wifi networks. In *2015 IEEE International Conference on Communications (ICC)*, pages 1571–1576. IEEE, 2015.

[12] Olivier Jeunen, Patrick Bosch, Michiel Van Herwegen, Karel Van Doorselaer, Nick Godman, and Steven Latré. A machine learning approach for ieee 802.11 channel allocation. In *2018 14th International Conference on Network and Service Management (CNSM)*, pages 28–36. IEEE, 2018.

[13] Francesc Wilhelmi, Cristina Cano, Gergely Neu, Boris Bellalta, Anders Jonsson, and Sergio Barrachina-Muñoz. Collaborative spatial reuse in wireless networks via selfish multi-armed bandits. *Ad Hoc Networks*, 88:129–141, 2019.

[14] Sergio Barrachina-Muñoz, Francesc Wilhelmi, and Boris Bellalta. Online primary channel selection for dynamic channel bonding in high-density wlans. *IEEE Wireless Communications Letters*, 2019.

[15] Jeremy Mack, Saeed Gazor, Amir Ghasemi, and John Sydor. Dynamic channel selection in cognitive radio wifi networks: An experimental evaluation. In *2014 IEEE International Conference on Communications Workshops (ICC)*, pages 261–267. IEEE, 2014.

[16] Mirghiasaldin Seyedebrahimi, Faycal Bouhafs, Alessandro Raschellà, Michael Mackay, and Qi Shi. Sdn-based channel assignment algorithm for interference management in dense wi-fi networks. In *2016 European conference on networks and communications (EuCNC)*, pages 128–132. IEEE, 2016.

[17] Heyi Tang, Liu Yang, Jiang Dong, Zhonghong Ou, Yong Cui, and Jianping Wu. Throughput optimization via association control in wireless lans. *Mobile Networks and Applications*, 21(3):453–466, 2016.

[18] George Athanasiou, Thanasis Korakis, Ozgur Ercetin, and Leandros Tassiulas. A cross-layer framework for association control in wireless mesh networks. *IEEE Transactions on Mobile Computing*, 8(1):65–80, 2008.

[19] Issam Jabri, Nicolas Krommenacker, Thierry Divoux, and Adel Soudani. Ieee 802.11 load balancing: an approach for qos enhancement. *International Journal of Wireless Information Networks*, 15(1):16–30, 2008.

[20] Yigal Bejerano, Seung-Jae Han, and Li Li. Fairness and load balancing in wireless lans using association control. In *Proceedings of the 10th annual international conference on Mobile computing and networking*, pages 315–329, 2004.

[21] Dawei Gong and Yuanyuan Yang. On-line ap association algorithms for 802.11 n wlans with heterogeneous clients. *IEEE Transactions on Computers*, 63(11):2772–2786, 2013.

[22] Chia-Ying Lin, Wan-Ping Tsai, Meng-Hsun Tsai, and Yun-Zhan Cai. Adaptive load-balancing scheme through wireless sdn-based association control. In *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)*, pages 546–553. IEEE, 2017.

[23] Estefania Coronado, Jose Villalon, and Antonio Garrido. Wi–balance: Sdn–based load–balancing in enterprise wlans. In *2017 IEEE Conference on Network Softwarization (NetSoft)*, pages 1–2. ieee, 2017.

[24] Marc Carrascosa and Boris Bellalta. Multi-armed bandits for decentralized ap selection in enterprise wlans. *Computer Communications*, 159:108–123, 2020.

[25] M Isabel Sanchez and Azzedine Boukerche. On ieee 802.11 k/r/v amendments: Do they have a real impact? *IEEE Wireless Communications*, 23(1):48–55, 2016.

[26] Martin Heusse, Franck Rousseau, Gilles Berger-Sabbatel, and Andrzej Duda. Performance anomaly of 802.11 b. In *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, volume 2, pages 836–843. IEEE, 2003.

[27] Bingtao Yang, Xue Wang, and Zhihong Qian. A multi-armed bandit model-based vertical handoff algorithm for heterogeneous wireless networks. *IEEE Communications Letters*, 22(10):2116–2119, 2018.

[28] MGS Sriyananda, Imtiaz Parvez, Ismail Güvene, Mehdi Bennis, and Arif I Sarwat. Multi-armed bandit for lte-u and wifi coexistence in unlicensed bands. In *2016 IEEE Wireless Communications and Networking Conference*, pages 1–6. IEEE, 2016.

[29] Doris Xin, Ahmed El-Kishky, De Liao, Brandon Norick, and Jiawei Han. Active learning on heterogeneous information networks: A multi-armed bandit approach. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 1350–1355. IEEE, 2018.

[30] Henrique D Moura, Daniel Fernandes Macedo, and Marcos AM Vieira. Automatic quality of experience management for wlan networks using multi-armed bandit. In *2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, pages 279–288. IEEE, 2019.

[31] ABM Alim Al Islam, SM Iftekharul Alam, Vijay Raghunathan, and Saurabh Bagchi. Multi-armed bandit congestion control in multi-hop infrastructure wireless mesh networks. In *2012 IEEE 20th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, pages 31–40. IEEE, 2012.

[32] Aleksandrs Slivkins et al. Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning*, 12(1-2):1–286, 2019.

[33] Joannes Vermorel and Mehryar Mohri. Multi-armed bandit algorithms and empirical evaluation. In *European conference on machine learning*, pages 437–448. Springer, 2005.

[34] William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.

[35] Shipra Agrawal and Navin Goyal. Further optimal regret bounds for thompson sampling. In *Artificial intelligence and statistics*, pages 99–107, 2013.

[36] Olivier Chapelle and Lihong Li. An empirical evaluation of thompson sampling. In *Advances in neural information processing systems*, pages 2249–2257, 2011.

[37] IEEE. Ieee standard for information technology– local and metropolitan area networks– specific requirements– part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications - spectrum and transmit power management extensions in the 5 ghz band in europe. *IEEE Std 802.11h-2003 (Amendment to IEEE Std 802.11, 1999 Edn. (Reaff 2003))*, pages 1–75, Oct 2003.

[38] Gilbert Chen and Boleslaw Szymanski. Cost: a component-oriented discrete event simulator. 1:776–782 vol.1, 12 2002.

[39] Simone Merlin, G Barriac, H Sampath, L Cariou, and T Derham. Tgax simulation scenarios. *IEEE802*, pages 11–14, 2015.

IEEE 802.11be Multi-Link Operation: When the Best Could Be to Use Only a Single Interface

Álvaro López-Raventós and Boris Bellalta

Abstract

The multi-link operation (MLO) is a new feature proposed to be part of the IEEE 802.11be Extremely High Throughput (EHT) amendment. Through MLO, access points and stations will be provided with the capabilities to transmit and receive data from the same traffic flow over multiple radio interfaces. However, the question on how traffic flows should be distributed over the different interfaces to maximize the WLAN performance is still unresolved. To that end, we evaluate in this article different traffic allocation policies, under a wide variety of scenarios and traffic loads, in order to shed some light on that question. The obtained results confirm that congestion-aware policies outperform static ones. However, and more importantly, the results also reveal that traffic flows become highly vulnerable to the activity of neighboring networks when they are distributed across multiple links. As a result, the best performance is obtained when a new arriving flow is simply assigned entirely to the emptiest interface.

1 Introduction

Since its appearance back in the late '90s, Wi-Fi has been continuously innovating to adapt its performance to the new user demands. Nowadays, the appearance of modern applications are pushing wireless local area networks (WLANs) to their performance limits again. For instance, remote office, cloud gaming and virtual and augmented reality (VR&AR) are straight-forward examples that not only demand high-throughput but reliable and low-latency communications [1]. To that end, in may 2019, the Institute of Electric and Electronic Engineers (IEEE) established a Task Group (TG) to address and design a new physical (PHY) and medium access control (MAC) amendment, known as IEEE 802.11be Extremely High Throughput (EHT).

The IEEE 802.11be EHT seeks to further increase the throughput performance, reduce the end-to-end latency and increase the reliability of communications [2]. To do so, different technical features have been suggested in both PHY and MAC layers [3]. Regarding PHY layer, TGbe propose the complementary adoption of the 6 GHz band, empowering the use of wider bandwidth channels up to 320 MHz, additionally to a new 4096 QAM high-order modulation (i.e., 12 bits per symbol), as immediate approaches to increase Wi-Fi peak-throughput. Besides, the multiple-input multiple-output (MIMO) capabilities are also being revised to upgrade them to support up to 16 spatial streams, introducing also an implicit channel sounding procedure. Despite those enhancements, it is not in the PHY, but in the MAC layer where we can find the most disruptive updates. We refer to the adoption of multi-link communications, which represents a paradigm shift towards concurrent transmissions. Although under the multi-link label we find the multi-AP coordination and the multi-band/multi-channel operation features, this article is focused on the analysis of the latter one.

Upon its current version, the IEEE 802.11 standard already defines two MAC architectures for supporting the multi-band/multi-channel operation. However, both designs present a common limitation: MAC service data units (MSDUs) belonging to the same traffic flow can not be transmitted across different bands [4]. As a result, stations are tied to a single band, preventing a dynamic and seamless inter-band operation. That is, one link is selected between transmitter and receiver to carry out data exchanges, remaining the other unused [5–8]. To benefit from the fact that modern APs and stations incorporate dual, or even, tri-band capabilities, TGbe is working on developing the multi-link operation[1] (MLO) framework to allow concurrent data transmission and reception in different frequency channels/bands.

MLO allows APs and stations to exploit the fact of having multiple radio interfaces to transmit and receive data. However, how to use them to maximize the WLAN performance, i.e., how to properly distribute the traffic across the multiple interfaces, is still an open question. Therefore, in this paper, we tackle such question by considering different traffic balancing policies. The obtained results show the advantages of coordinating the available interfaces through the MLO framework, as well as provide some insights on how to distribute the traffic across them. It is worthy to anticipate our conclusion that distributing the traffic over multiple interfaces may not be always the best solution.

2 Multi-link operation: an overview

The introduction of MLO is a breaking point for Wi-Fi, as its adoption represents a paradigm shift towards multi-link communications. However, this implementa-

[1]Throughout this paper, we will refer to the multi-band/multi-channel operation feature as the MLO, following the notation of the TGbe.

tion involves new challenges in terms of APs' and stations' architecture design, transmission modes, channel access, and management.

To enable a concurrent operation across multiple interfaces, the existing multi-band MAC architecture have been exhaustively revised. In this context, TGbe introduces the concept of multi-link capable device (MLD), which consists of a single device with multiple wireless PHY interfaces that provides a unique MAC instance to the upper layers. Such implementation is achieved by dividing the MAC sub-layer in two parts [9]. First, we find the unified upper MAC (U-MAC), which is the common part of the MAC sub-layer for all the interfaces. Apart from performing link agnostic operations such as A-MSDU aggregation/de-aggregation and sequence number assignment, the U-MAC implements a queue that buffers the traffic received from the upper-layers. That is, traffic awaits in the U-MAC before it is assigned to a specific interface to be transmitted. Also, the U-MAC provides some management functions such as multi-link setup, association and authentication. On the other hand, there is the independent low MAC (L-MAC), an individual part of the MAC sub-layer for each interface that performs link specific functionalities. There, we find the link specific EDCA queues (one for each access category) that hold the traffic until its transmission. Also, procedures such as MAC header and cyclic redundancy check (CRC) creation/validation, in addition to management (e.g., beacons) and control (e.g., RTS/CTS and ACKs) frame generation are implemented at the L-MAC [10].

This two-tier MAC implementation enables frames to be simultaneously transmitted over multiple links, as the U-MAC performs the allocation and consolidation of MPDUs when acting as transmitter and receiver, respectively. Also, it enables seamless transitions between links minimizing the access latency, and addressing an efficient load balancing.

At a link level (i.e., in the L-MAC), channel access takes place. In current proposals, TGbe defines different channel access methods accordingly to two different transmission modes: asynchronous and synchronous. First, there is the asynchronous transmission mode. Under this operational mode, a MLD can transmit frames asynchronously on multiple links, while keeping for each one its own channel access parameters (e.g., contention window (CW), arbitration inter-frame spacing number (AIFSN), etc.). That is, each link has its own primary channel, independent of the others. Also, this transmission mode allows the simultaneous transmission and reception (STR) capability over multiple links, enabling concurrent uplink and downlink communications. Such implementation has been proved to minimize latency [11], while maximizing the throughput [12]. Ideally, the asynchronous mode should be selected as operational scheme, however, a MLD may not be STR-capable due to the in-device coexistence (IDC) interference. This issue is caused by an excessive power leakage between interfaces that do not have sufficient frequency separation in their operating channel (e.g., two channels in the 5 GHz band). As a result, IDC prevents frame reception on one interface, during an ongoing transmission on the other interface. To avoid the IDC issue, TGbe proposes the synchronous mode, which relies on synchronized frame transmissions across the available interfaces. With that, APs or stations are prevented to perform an STR operation, avoiding IDC problems but at the cost of a lower network throughput [13, 14]. Under this operation mode, the channel access can be performed either following the single primary channel (SPC) or the multiple primary channel (MPC) methodology. Basically, the SPC performs contention on a unique channel, whereas in the MPC contention is performed on all channels. Either using the SPC or the MPC method, if one channel wins contention the others are checked during a PCF inter-frame space (PIFS) time to see if they can be aggregated. As a result, channels in idle state are aggregated and a synchronous transmission starts [15]. It is worth mention that AP MLDS may change its transmission modes (e.g., asynchronous to synchronous, and vice versa) at any time. Devices operating in a synchronous mode are referred to as constrained MLDs, or non-STR MLDs, since they are not allowed to transmit through an idle interface at the same time they are receiving through another.

Regarding MLO feature management, the TGbe proposes the multi-link setup process. Instead of designing new management frames, TGbe agreed to reuse the current association request/response frames by adding an extra multi-link element or field. That is, the setup (i.e., MLO capability exchange) is performed jointly with the association mechanism. Then, through the multi-link element, AP MLDs and non-AP MLD negotiate and establish their subsequent operation scheme by exchanging their capabilities. For instance, they exchange information regarding the number of supported links, their ability to perform STR, their transmission operation (asynchronous or synchronous), and other per-link information (e.g., frequency band, supported bandwidth, number of spatial streams, etc.) [16, 17]. To reduce overhead, the described multi-link setup process is proposed to be performed only on a single link, which, indeed, will be the lowest in frequency due to propagation constraints.

The interface usage negotiation is performed by measuring link qualities at all interfaces. That is, those receiving a quality value above the clear channel assessment (CCA) threshold are set as enabled, while disabled otherwise. Hence, for each enabled interface, we will have an enabled link. It is worth mention that, as users may keep themselves mobile, any link listed as disabled may be added afterwards to the set of enabled links, by requesting a re-setup.

3 System model

3.1 Node placement

We consider a set of IEEE 802.11be MLO-capable WLANs with N APs and M stations. Over a given area, APs are placed uniformly at random. However, we only accept the generated scenario as valid if the inter-AP distance is higher than 5 m to avoid unrealistic overlaps. Otherwise, we discard it and generate a new one. Then, we decide the number of stations that will be associated to each AP, placing them around it at a distance d within the interval [1-8] m and an angle θ between $[0\text{-}2\pi]$, both selected uniformly at random.

Regardless of its type, APs and stations are configured with 3 wireless interfaces (i.e., i_1, i_2 and i_3), each one using 2 SU-MIMO spatial streams. Additionally, we consider that each interface operates at a different frequency band. That is, for each AP MLD interface, one channel c is selected uniformly at random from the set of channels of each band, where $\mathcal{C}_b$ is the set of available channels of band b. Thus, $c_{i_1} \in \mathcal{C}_{2.4}$, $c_{i_2} \in \mathcal{C}_5$ and $c_{i_3} \in \mathcal{C}_6$. The different set of channels are detailed in Table 1. Note that depending on the band, different channel widths are considered.

3.2 MLO capabilities

The stations' set-up, as well as the establishment of the enabled interfaces, is performed through the 2.4 GHz link. Link qualities are exchanged, and therefore, interfaces that meet the quality criteria (i.e., above CCA threshold) are set as enabled, while disabled otherwise. In essence, this setup relies on a simplified version of the MLO setup process described in Section 2. Additionally, we assume that all nodes perform an asynchronous channel access for each of their enabled interfaces, while their default policy is set as *Multi Link Same Load to All interfaces* (MLSA) (see Section 3.5), which will be our baseline evaluation policy unless stated differently.

3.3 Channel model and data rate selection

Path loss is characterized following the IEEE 802.11ax enterprise model for a single floor environment. The path loss between an station i and AP j is given by

$$\begin{aligned}\mathrm{PL}(d_{i,j}) = 40.05 + 20\log_{10}\left(\frac{f_c}{2.4}\right) + 20\log_{10}(\min(d_{i,j}, d_{bp})) \\ + (d_{i,j} > d_{bp}) \cdot 35\log_{10}\left(\frac{d_{i,j}}{d_{bp}}\right) + 7W_{i,j}\end{aligned}$$

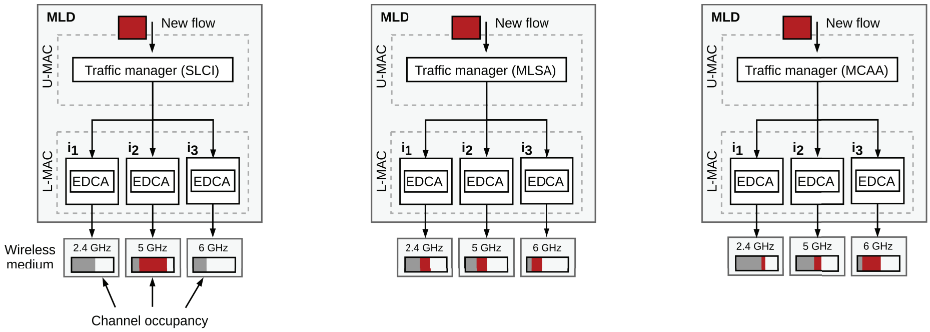


Figure 1: Policy implementation. From left to right: SLCI, MLSA and MCAA representation. Under the wireless medium representation, the gray area corresponds to the channel occupancy already seen by the MLD at the given interface, which can be from its neighbor APs, as well as its own ongoing flows.

where f_c is the AP's carrier frequency in GHz, $d_{i,j}$ is the distance between station i and AP j in meters, d_{bp} is the breaking point distance in meters, and $W_{i,j}$ are the number of traversed walls. We set the breaking point distance to 5 m and the number of traversed walls to 4. Note that the resulting propagation losses are given in dB.

The Modulation and Coding Scheme (MCS) used by each interface is selected according to the signal-to-noise ratio (SNR) between the AP and the stations. For instance, an station's interface i, with 2 spatial streams and a SNR of 11 dB, will achieve a data rate of 243.8 Mbps, using modulation 1024-QAM with a coding rate of 5/6 and guard interval of 3.2 μs in a 20 MHz channel.

3.4 Traffic generation and CSMA operation

Only downlink traffic is considered. The traffic directed to each station is modeled as an on/off Markovian model. In the on period, the AP receives a Constant Bit Ratio (CBR) traffic flow of ℓ Mbps, whereas zero during the off period. Both on and off periods are exponentially distributed with mean duration $\mathrm{T_{on}}$ and $\mathrm{T_{off}}$, respectively. Either $\mathrm{T_{on}}$ and $\mathrm{T_{off}}$ have been set to resemble a mixture of different types of traffic as in real scenarios. The value of the two parameters is presented in Table 1. We refer to the traffic generated during the on period as a traffic flow, and to ℓ as the required bandwidth.

Regarding the CSMA/CA operation, it follows the abstraction presented in [18], which considers the aggregate channel load observed by each AP to calculate the airtime that can be allocated to each flow. Although the considered abstraction does not capture low-level details of the PHY and MAC layers oper-

Table 1: Evaluation setup

Parameter	**Description**
Carrier frequency	Depends on channel selection
$\mathcal{C}_{2.4}$ channel set	1 (20 MHz), 6 (20 MHz), 11 (20 MHz)
$\mathcal{C}_5$ channel set	38 (40 MHz), 46 (40 MHz), 58 (80 MHz)
$\mathcal{C}_6$ channel set	55 (80 MHz), 71 (80 MHz), 15 (160 MHz)
System bandwidth	60 MHz/160 MHz/320 MHz
AP/STA TX power	20/15 dBm
Antenna TX/RX gain	0 dB
CCA threshold	-82 dBm
AP/STA noise figure	7 dB
Single user spatial streams	2
MPDU payload size	1500 bytes
Path loss	Same as [18]
Avg. flow duration	$T_{\text{on}} = 1$ s
Avg. flow interarrival time	$T_{\text{off}} = 3$ s
Min. contention window	15
Packet error rate	10%
Simulation time	120 s (1 per simulation)
Number of simulations	100 (per evaluation point)

ation, it maintains the essence of the CSMA/CA: the 'fair' share of the spectrum resources among contending APs and stations. In addition, this presented abstraction allows us to simulate larger networks, and larger periods of time, and so study the network performance at the flow-level.

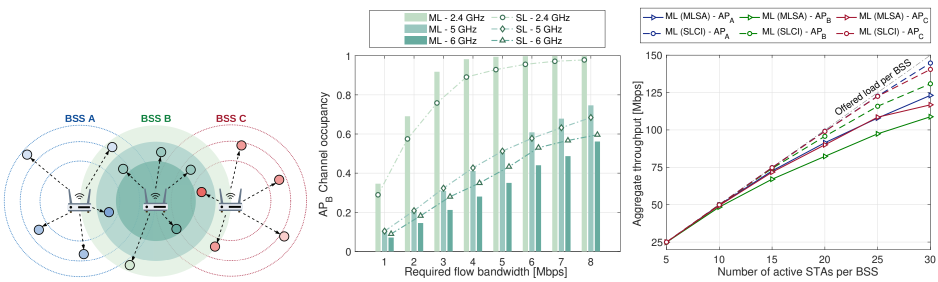
3.5 Traffic allocation policies

When a new traffic flow arrives at the AP, the traffic manager determines how it is distributed over the different enabled interfaces of each station. The three traffic allocation policies considered in this work are:

- **Single Link Less Congested Interface (SLCI):** upon a new flow arrival, pick the less congested interface and assign it to the incoming flow.
- **Multi Link Same Load to All interfaces (MLSA):** upon a new flow arrival, distribute the incoming traffic flow equally between all the enabled interfaces of the receiving station. That is, let ℓ be the required bandwidth of the incoming flow, and N_{int} the number of enabled interfaces in the destination station. This is, traffic allocation per interface is given by: $\ell_i = \ell/N_{int}$, with ℓ_i the bandwidth allocated to interface i.
- **Multi Link Congestion-aware Load balancing at flow arrivals (MCAA):** upon a new flow arrival, distribute the incoming traffic flow

accordingly to the channel occupancy observed at the AP considering the enabled interfaces at the receiving station. That is, let ρ_i the percentage of available (free) channel airtime at interface i. Then, the required bandwidth allocated to interface i is given by $\ell_{i \in \mathcal{J}} = \ell \frac{\rho_i}{\sum_{\forall j \in \mathcal{J}} \rho_j}$, with $\mathcal{J}$ the set of enabled interfaces at the target station.

While MLSA allocates traffic regardless of the congestion level of each interface, the other policies take it into account. When a new flow arrives, SLCI evaluates the current congestion level of each interface to identify the emptiest one, and so, it assigns the whole traffic flow to it. Then, we find MCAA, which tries to maximize the spectrum used by balancing the traffic allocated to the different interfaces taking into account their current occupancy. Under the MCAA, for instance, a flow with a bandwidth requirement of 10 Mbps, and capable to be distributed across three interfaces with $\rho_1 = 0.2$, $\rho_2 = 0.6$, and $\rho_3 = 0.5$, will be split as follows: 1.54 Mbps allocated to the interface 1, 4.61 Mbps to the interface 2, and 3.85 Mbps to interface 3. Although this approach seems very convenient, traffic may become more vulnerable to contention with external networks, due to the fact of being spread across multiple interfaces. Figure 1 shows the different policies presented.



(a) (b) (c)

Figure 2: Controlled scenario evaluation. From left to right: a) scenario representation, b) average channel occupancy for AP B, and c) aggregate throughput performance. The high, medium and low shaded areas represent the operation range for the 6 GHz, 5 GHz and 2.4 GHz bands, respectively.

4 Performance evaluation

To assess the performance evaluation of the MLO, we use the Neko[2] simulation platform.

4.1 A controlled scenario

First, we focus our evaluation on the comparison between a MLO-capable deployment against a traditional multi-band single link (SL). To do so, we consider a controlled scenario to pinpoint the potential benefits about the introduction of MLO, while identifying some pitfalls without the complex interactions that random deployments may originate. Figure 2a depicts a typical controlled WLAN deployment, in which three basic service sets (BSSs) are placed inline. Under such scenario, we assess the network behavior under two different cases: i) high traffic demands, and ii) high station density deployments. For each evaluation point considered in each case, we simulate 100 scenarios in which stations' positions are newly generated following the stated in Section 3.1.

To begin with, we characterize the network performance for different traffic loads. In this regard, we fixed the average number of associated stations per AP to 20, while increasing the required bandwidth for each flow from 1 Mbps to 8 Mbps. That is, 800 total simulated scenarios, as we evaluate 8 traffic load ranges for 100 different scenarios. Besides, the whole deployment is considered to be either multi-band SL (i.e., 3 SL APs), or MLO-capable (i.e., 3 AP MLDs). In both scenarios, the channel configuration is the same for the 3 APs, whereas the other simulation parameters are kept as presented in Table 1. It is worth mention that when using the multi-band SL approach, stations select the interface to be attached to in a random fashion from the set of enabled interfaces. For MLO, the configuration of each interface is set as stated in Section 3.1, and both APs and stations use baseline MLSA policy.

Figure 2b shows the experienced channel occupancy by $\mathrm{AP_B}$. We observe that $\mathrm{AP_B}$ suffers from the well-known flow-in-the-middle effect in its 2.4 GHz interface, due to $\mathrm{AP_A}$ and $\mathrm{AP_C}$ keeping the wireless medium in busy state for most of their time. Consequently, traffic bandwidth requirements in this interface are only met for low level demands in both approaches, since $\mathrm{AP_B}$ will start to drop data traffic as soon as traffic load increases. On the contrary, such effect is alleviated in 5 GHz and 6 GHz links as a consequence of nodes' spatial distribution, which provides all APs with two downlink contention-free links (i.e., only the 2.4GHz link satisfies the energy detection threshold). Although such condition may benefit the MLO capability implementation, in this case it is wasted by not using a congestion-

[2]The Neko simulation platform can be found in GitHub at: `https://github.com/wn-upf/Neko`

aware traffic allocation policy to leverage both contention-free links. Such issue has a negative impact over AP_B performance, as its experienced occupation is nearly the same either when using both multi-band SL or MLO approaches. That is, there is no difference between spreading in the same proportion the traffic over the different links through MLO, than balancing and/or steering clients across links, like current multi-band SL deployments do. Thus, it is easy to conclude that without a proper traffic allocation policy, MLO underperforms.

Similarly, Figure 2c shows the aggregate throughput of each AP when the number of stations per BSS increases. Differently from the previous case, we set now that the required bandwidth for each generated flow ℓ is selected uniformly from the 2 Mbps to 8 Mbps interval (i.e., $\ell \sim U[2,8]$). The other simulation parameters are kept as previously. Now, we introduce the SLCI congestion-aware policy presented in Section 3.5 to provide a comparison between a non-congestion and a congestion aware policy. From the figure, we observe how the MLSA policy struggles to cope even with 15 stations per BSS. For instance, AP_B already registers throughput losses around 12% for 15 stations. On the contrary, the SLCI assisted MLO operation supports a higher number of stations, while AP_B only experiences a 5% throughput losses when considering 20 stations. Additionally, we observe that the flow-in-the-middle effect is reduced by the implementation of a congestion-aware policy. For instance, we observe that AP_B's aggregate throughput is reduced 10% in comparison to its neighbors when considering MLSA for 20 stations, whereas only a 4% when considering SLCI.

4.2 Random deployments

In order to get more insights with respect to the performance of the different policies presented in Section 3.5, we conduct a set of simulations in random generated scenarios. Similarly to the previous case, we address two different cases: i) high user demands, and ii) high AP density deployments. Both cases are conducted over an area of 45x45m^2, with nodes being placed and configured as stated in Section 3. Besides, we consider that both APs and stations within the evaluation area implement the same policy. Other simulation conditions are kept as stated in Table 1. For each evaluation point considered in each case, we simulate 100 different scenarios in which APs and stations' positions are generated as stated in Section 3.1.

First, we conduct the evaluation with respect to stations traffic requirements. For this purpose, we locate 10 AP MLDs within the considered area, as well as a random number of stations (i.e., $M \sim U[15,25]$) for each AP MLD. Then, we carry out simulations by gradually increasing the required bandwidth ℓ of each incoming flow from 1 Mbps to 8 Mbps. To evaluate the performance of the network we use the satisfaction s metric. We define the satisfaction as the

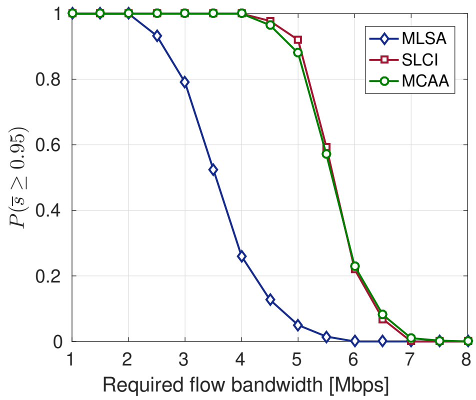


Figure 3: Prob. of average satisfaction

ratio between the required airtime to achieve the required flow bandwidth, and the actual amount that can be allocated. This metric is calculated for each AP, and then, $\overline{s}$ indicates the average satisfaction of the network computed as the sum of the average satisfaction experienced by each AP divided by the total number of APs in the WLAN. Figure 3 shows the probability that the whole WLAN achieves an average satisfaction $\overline{s}$ greater or equal than 95%. As expected, the baseline MLSA operation is outperformed by the rest of policies, due to the fact that traffic allocation is no longer allocated blindly, but following a congestion-aware approach. In this regard, we find that both SLCI and MCAA congestion-aware policies are capable to deal with higher traffic loads without losing performance. For instance, for $\ell = 5$ Mbps, SLCI and MCAA are able to provide excellent satisfaction values in 92% and 88% of the scenarios evaluated, respectively, whereas MLSA only manages to get good results in 5% of them. Finally, notice that SLCI works better for medium load ranges, while MCAA works better for high loads as its ability to distribute traffic across the different interfaces results in slightly higher satisfaction values.

Table 2: Traffic allocation efficiency

bandwidth req. per flow	Policy		
	MLSA	SLCI	MCAA
2 Mbps	0.996	1.00	1.00
4 Mbps	0.925	0.989	0.985
6 Mbps	0.830	0.931	0.930
8 Mbps	0.750	0.833	0.842

To further examine the obtained results, we show in Table 2 the traffic allocation efficiency. We define this efficiency as the ratio between the required bandwidth of each flow, and the actual throughput, for each of the policies. Once more, we can see that congestion-aware policies are able to outperform the baseline operation, as all of them can effectively allocate more than the 90% of the required bandwidth per flow for values between 2 Mbps and 6 Mbps. Also, notice how the performance of the two congestion-aware policies (i.e., SLCI and MCAA) is highly similar.

Finally, we assess the performance of the different policies under high AP density deployments. To do so, we gradually increase the number of APs, to generate four different density deployments: low ($N = 5$), medium ($N = 10$), med-high ($N = 20$) and high ($N = 40$). For each AP, we placed a random number of stations (i.e., $M \sim U[15, 25]$). Figure 4 shows the empirical cumulative distribution function (CDF) of the average throughput drop ratio experienced for each density. The throughput drop ratio represents the percentage of traffic that cannot be served. Then, it is computed as one minus the ratio between the achieved throughput and the required one. As expected, the probability of having higher drop ratio values increases upon network densification. However, the different congestion-aware policies are able to reduce significantly those values. For instance, for medium density deployments, we find that drop ratio is decreased 2.25x by means of SLCI implementation for the 75th percentile compared to MLSA. Additionally, we find that SLCI further improves MCAA as the density increases. This phenomena occurs as MCAA policy allocates traffic to more interfaces, which in fact increases the exposure to the dynamics of the other APs. That is, in high density deployments, APs tend to have more neighboring APs, and so, the use of a high number of interfaces to transmit makes the traffic much more vulnerable as the probability of having a channel overlap increases.

5 Conclusions & Future Work

In this paper, we assessed and evaluated the adoption of multi-link communications by the upcoming IEEE 802.11be amendment. Through a wide variety of scenarios, our results show that the implementation of MLO can help the next generation Wi-Fi networks to satisfy the highly demanding requirements of modern applications. However, we show that the performance of MLO depends mainly on the implemented traffic allocation policy. From that side, as expected, congestion-aware allocation policies able to adapt to the instantaneous state of the network are the best performing ones. Additionally, we have observed that allocating an incoming flow to the emptiest interface is almost as good, if not better, than proportionally distributing the flow over multiple interfaces. Such conclusion relies on the fact that a single interface policy not only reduces traffic

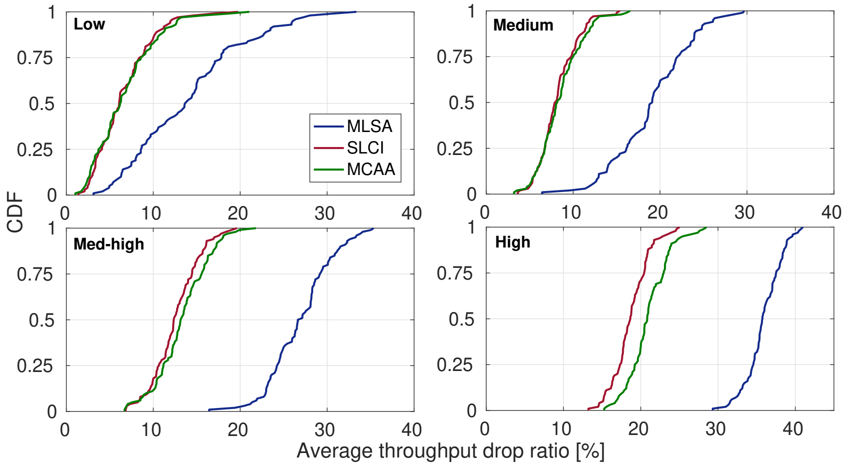


Figure 4: Drop ratio CDF for different network densities

exposure, but minimizes the complexity of the traffic allocation procedure.

Regarding future work, we consider three potential extensions to further improve MLO operation. First, extend MCAA to re-allocate the traffic over the different interfaces periodically. Indeed, such implementation may minimize the negative impact of the activity of neighboring networks in terms of performance, but at the cost of entailing a much complex policy structure. Second, throughout this paper, we only considered IEEE 802.11be MLO capable networks, whereas in real scenarios they may coexist with legacy single link networks. Therefore, it may be relevant to study the different MLO traffic allocation policies under those conditions. Third, and last, to integrate the MLO operation in a Software Defined Networking framework [19] to orchestrate the different APs that belong to the same WLAN, as well as to design and evaluate centralized multi-AP MLO policies.

Acknowledgments

This work has been partially supported by the Spanish Government under grant WINDMAL PGC2018-099959-B-I00 (MCIU/AEI/FEDER,UE), by the Catalan Government under grant 2017-SGR-1188, and by a Gift from Cisco (2020).

References

[1] M. Carrascosa and B. Bellalta, "Cloud-gaming: Analysis of Google Stadia traffic," *arXiv preprint arXiv:2009.09786*, 2020.

[2] T. Adame, M. Carrascosa, and B. Bellalta, "Time-sensitive networking in IEEE 802.11 be: On the way to low-latency WiFi 7," *arXiv preprint arXiv:1912.06086*, 2019.

[3] D. López-Pérez, A. Garcia-Rodriguez, L. Galati-Giordano, M. Kasslin, and K. Doppler, "Ieee 802.11 be extremely high throughput: The next generation of wi-fi technology beyond 802.11 ax," *IEEE Communications Magazine*, vol. 57, no. 9, pp. 113–119, 2019.

[4] Y. Fang, B. Sun, H. Zhiqiang, and N. Li, "Multi-link architecture and requirement discussion." IEEE 802.11 Documents. 2019. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1095&is_group=00be`.

[5] Z. M. Fadlullah, Y. Kawamoto, H. Nishiyama, N. Kato, N. Egashira, K. Yano, and T. Kumagai, "Multi-hop wireless transmission in multi-band wlan systems: Proposal and future perspective," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 108–113, 2017.

[6] P. Kyasanur and N. H. Vaidya, "Routing and interface assignment in multi-channel multi-interface wireless networks," in *IEEE Wireless Communications and Networking Conference, 2005*, vol. 4, pp. 2051–2056, IEEE, 2005.

[7] H. Singh, J. Hsu, L. Verma, S. S. Lee, and C. Ngo, "Green operation of multi-band wireless lan in 60 ghz and 2.4/5 ghz," in *2011 IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 787–792, IEEE, 2011.

[8] N. Choi, Y. Seok, and Y. Choi, "Multi-channel mac protocol for mobile ad hoc networks," in *2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No. 03CH37484)*, vol. 2, pp. 1379–1382, IEEE, 2003.

[9] Y. Fang, B. Sun, H. Zhiqiang, and L. Lu, "Multi-link reference model discussion." IEEE 802.11 Documents. 2020. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1131&is_group=00be&is_year=2020`.

[10] J. Levy and X. Wang, "802.11be architecture/association discussion." IEEE 802.11 Documents. 2020. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1122&is_group=00be`.

[11] G. Naik, D. Ogbe, and J.-M. J. Park, "Can wi-fi 7 support real-time applications? on the impact of multi link aggregation on latency,"

[12] S. Naribole, W. B. Lee, S. Kandala, and A. Ranganath, "Simultaneous transmit-receive multi-channel operation in next generation wlans," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–8, IEEE, 2020.

[13] W. Murti, S. Kim, J.-H. Yun, J. Son, G. Ko, and K. Jin Sam, "Performance and fairness of multi-link operations." IEEE 802.11 Documents. 2019. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1633&is_group=00be`.

[14] A. Patil, G. Cherian, A. Asterjadhi, and D. Ho, "Multi-link aggregation - gain analysis." IEEE 802.11 Documents. 2019. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=764&is_group=00be&is_year=2019`.

[15] I. Jang, J. Choi, J. Kim, S. Kim, S. Park, and T. Song, "Channel access for multi-link operation." IEEE 802.11 Documents. 2019. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1144&is_group=00be`.

[16] I. Jang, J. Choi, J. Kim, S. Kim, S. Park, and T. Song, "Discussion on multi-link setup." IEEE 802.11 Documents. 2019. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1509&is_group=00be`.

[17] I. Jang, J. Choi, J. Kim, T. Song, and N. Kim, "Indication of multi-link information: Follow-up." IEEE 802.11 Documents. 2020. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=741&is_group=00be`.

[18] Á. López-Raventós and B. Bellalta, "Concurrent decentralized channel allocation and access point selection using multi-armed bandits in multi bss wlans," *Computer Networks*, vol. 180, p. 107381, 2020.

[19] Á. López-Raventós, F. Wilhelmi, S. Barrachina-Muñoz, and B. Bellalta, "Combining software defined networks and machine learning to enable self organizing WLANs," in *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 1–8, IEEE, 2019.

Multi-link Operation in IEEE 802.11be WLANs

Álvaro López-Raventós and Boris Bellalta

Abstract

The multi-link operation (MLO) is a new feature proposed to be part of the IEEE 802.11be Extremely High Throughput (EHT) amendment. Such feature represents a paradigm shift towards multi-link communications, as nodes will be allowed to transmit and receive data over multiple radio interfaces concurrently. To make it possible, the 802.11be Task Group has proposed different modifications in regards to nodes' architecture, transmission operation, and management functionalities. This article reviews such changes and tackles the question of how traffic should be distributed over multiple links, as it is still unresolved. To that end, we evaluate different load balancing strategies over the active links. Results show that in high load, dense and complex scenarios, implementing congestion aware load balancing policies to significantly enhance next-generation WLAN performance using MLO is a must.

1 Introduction

Commonly known as WiFi, the IEEE 802.11 standard was released back in the late 90s, with the aim to provide a low complex, and cost efficient, wireless connectivity solution. Currently in its 6th generation, the proliferation of WiFi has been driven by the constant revision of the standard, since periodic amendments have made possible to face the increasing requirements of newer use-cases. Wireless data services will continue to grow, with upcoming applications, such as virtual/augmented reality, video/game streaming and cloud based services, requesting vasts amounts of data with the most demanding throughput, latency, and reliability requirements. To address such expectations, the 802.11be Task Group (TGbe) was created in May 2019 to address the development of new specifications to fuel the upcoming WiFi 7.

Referred to as IEEE 802.11be Extremely High Throughput (EHT) [1], this amendment aims to increase the WiFi throughput, while reducing the end-to-end latency and improving the reliability of communications [2]. For such purpose, the Multi-link Operation (MLO[1]) is considered a main candidate feature, as it promotes the use of multiple wireless interfaces to allow concurrent data transmission and reception in access points (APs) and stations (STAs) with dual- or tri-band capabilities.

Indeed, the interest in the use of the MLO framework is rapidly increasing. Latency in real-time applications has been already studied in [3–5], showing that MLO is able to significantly reduce worst-case latency. Besides, authors in [4] extended their analysis to evaluate the reliability over multiple links, showing a high delivery rate when having multiple uncorrelated links. Also, an end time alignment mechanism to allow the use of parallel downlink transmissions through different links to stations without simultaneous transmit and receive (STR) capability is presented in [6]. Such approach is intended to maximize the spectrum efficiency. Analogously, an opportunistic backoff mechanism is proposed in [7] to allow non-STR stations to resume their backoff timers, if an ongoing transmission is identified to not cause a collision. Authors in [8] suggest that the use MLO per se may not be sufficient enough without coordination between APs, proposing a coordination framework to achieve high throughput requirements in high density areas.

The integration of a framework capable to operate at the same time over multiple wireless interfaces brings up new challenges and research opportunities. In this context, we find that MLO compliant devices will have the ability to transmit and receive packets with different quality-of-service (QoS) requirements over multiple links. Such functionality, which was not allowed in past amendments, is called traffic identifier (TID) to link mapping, and opens up to conceive new traffic management mechanisms. For instance, we may find all TIDs to be assigned to all links, allowing a full adaptive load balancing strategy, as traffic may be moved partially or fully between multiple links. On the contrary, other approaches may rely on having a dedicated link assigned to an specific QoS traffic, which implies a more rigid and less flexible load balancing solution, but ensuring that only traffic with the same QoS requirements share the same set of resources.

In this article, we assess different allocation strategies that follow either an adaptive or link-dedicated implementation, with the aim to provide some insights on how to distribute the traffic across multiple interfaces, and how it may affect to the network performance. The main contributions are:

- We provide a comprehensive overview on how the MLO framework is being devised by the TGbe, pointing out the different modifications in regards the nodes' architectural changes, transmission modes and management functions.

[1]Throughout this paper, we will refer to the multi-band/multi-channel operation feature as the MLO, following the notation of the TGbe.

- We discuss the potential benefits, and challenging issues related to the overviewed modifications. Also, we point out some open issues and research directions faced by the MLO framework.

- We assess different policy-based strategies in order to tackle the traffic allocation problem. Also, we adopt the TID-to-link mapping functionality to showcase its implementation, benefits and drawbacks.

- We evaluate the presented strategies under different traffic requirements, showing that a link-dedicated approach may not be suitable neither in high dense, nor high load use-case scenarios.

2 Multi-link operation

In the following, we explore the different, and the most relevant, proposals that are likely to be included in the IEEE 802.11be amendment regarding the MLO feature.

2.1 Architecture

The first architectural change is found in the redefinition of classical APs or STAs into the so-called multi-link capable devices (MLDs). Either AP MLDs or STA MLDs, refer to single devices with multiple wireless interfaces[2]. The most relevant aspect about that remains on the fact that MLDs will provide a unique MAC instance to the upper layers, without losing the independent parameters of each interface. To achieve that, TGbe proposed to divide the MAC sub-layer functionalities in two different levels [9]. Figure 1 depicts the MLD architecture, representing both MAC sub-layer levels.

First, there is the upper MAC (U-MAC), which is a common part of the MAC sub-layer for all the interfaces. In the U-MAC, we find that link agnostic operations take place. We refer, for instance, to sequence number assignation, and MAC service data units (MSDUs) aggregation/de-aggregation. In this context, it is important to point out that the sequence number assignation must be performed at the U-MAC, since packets belonging to the same traffic flow can be fragmented and transmitted over different links. Such approach, then, eases the packet reordering at the receiver side. Additionally, common management functions for all links, such as setup, association and authentication take placed in this layer.

[2]Instead of interfaces, the TGbe defines them as affiliated AP/STAs. However, for sake of simplicity, and comprehensive purposes, we will keep referring to them as interfaces.

Below the U-MAC, we find the low MAC (L-MAC). This lower level, which is independent for each interface, is in charge of link specific functionalities like the channel access. In this context, we find that having individual L-MAC instances allow interfaces to keep their own channel parameters if needed. Inherently, this implementation also grants each interface to keep track of their own enhanced distributed channel access (EDCA) queues (one for each access category) to hold the traffic until its transmission. Other functionalities in the L-MAC layer are the management and control frame generation, as well as the MAC header creation and validation, when transmitting and receiving respectively [10].

The motivation behind this two-tier architecture is to permit MLO-capable devices to move traffic from one link to another, being totally transparent to upper-layers. Hence, load balancing techniques may be useful to minimize the spectrum usage inefficiency of current standardized multi-band approaches, in which per client transmissions are only performed either in one band or another, by leveraging the use of all the available resources. However, such architecture entails a more complex design, requiring not only to design new methods to perform traffic to link allocation, if not also to rethink low-level aspects regarding how channel contention and packet transmissions are done in presence of multiple links.

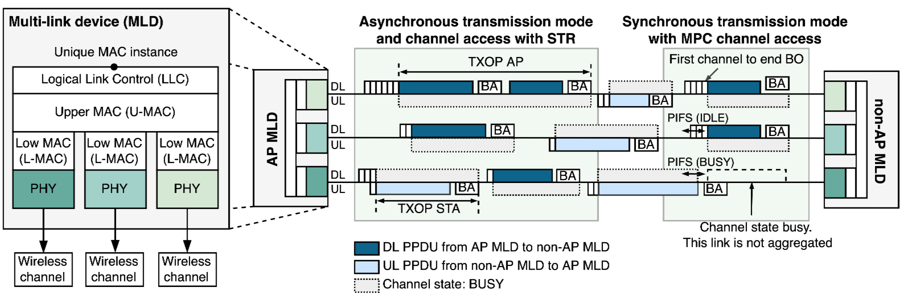


Figure 1: Multi-link architecture and transmission modes representation. Each PHY color represents a different band/channel for each one of the different interfaces.

2.2 Transmission modes

The TGbe defines two different transmission modes for MLDs. First, the asynchronous transmission mode allows a MLD to transmit frames asynchronously on multiple links. Under this mode, each interface keeps its own channel access parameters with an independent behavior respect the others. Also, it allows the

STR capability, enabling concurrent uplink (UL) and downlink (DL) communications, as depicted Figure 1. Ideally, it is suggested that the asynchronous mode should be selected as the default operational scheme by all 802.11be compliant nodes, since it provides a higher throughput performance [11]. However, such operation must be followed by a power save mechanism, specially for handheld devices, as the power consumption may be significantly high due to having multiple asynchronous interfaces operating at the same time. Additionally, the asynchronous operation is constrained to the in-device coexistence (IDC) interference. That is, the power leakage between interfaces may prevent a frame reception on one interface, during and ongoing transmission on the other interface, as a result of not having enough separation between operating bands/channels (e.g., two channels in the 5 GHz band).

To avoid the IDC issue, the TGbe defines the synchronous mode, which relies on synchronized frame transmissions across the available links. Devices operating under a synchronous mode are referred to as constrained MLDs, or non-STR MLDs, since they are not allowed to transmit through an idle interface at the same time they are receiving through another. To perform synchronization, the end-time alignment or the defer transmission mechanisms [1] may be implemented. While the former relies on ending transmissions on different channels at the same time, the latter defers the transmission of a link that has finished its backoff, until the end of the same counter in other links. With that, APs or stations are prevented to perform an STR operation, avoiding IDC problems, but at the cost of a lower throughput, if compared to the asynchronous scheme. In regards of channel access, it can be performed either following a single primary channel (SPC) or a multiple primary channel (MPC) methodology. Basically, the SPC performs contention on a unique channel, whereas in the MPC contention is performed on all channels. While applying a MPC scheme offers nodes higher chances to win contention and transmit frames, SPC allows to reduce power consumption as non-primary channel interfaces' may remain under a doze state. Figure 1 shows the synchronization scheme with MPC channel access, considering the end-time alignment mechanism. Either using the SPC or the MPC method, if an interface wins contention in its channel, the others are checked during a PCF inter-frame space (PIFS) time to see if they can be aggregated, performing a transmission opportunity (TXOP) aggregation. At last, it is worth mention that MLD-capable APs may change its transmission modes (e.g., asynchronous to synchronous, and vice versa) at any time, as depicted in Figure 1.

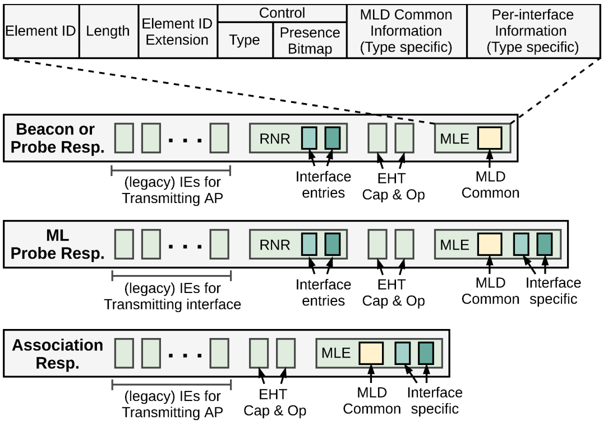


Figure 2: Multi-link element and management frames

2.3 Management

2.3.1 The Multi-Link element

The information elements (IEs) included in the different management frames allow devices to exchange their capabilities and operational parameters. With such purpose, the 802.11be defines the multi-link element (MLE). As shown in Figure 2, the MLE has been designed as a common element to the different management actions (e.g., discovery and setup). To achieve such implementation, the MLE introduces a type sub-field within the control field, that maps each operation to an specific value [12]. Hence, this information field is type-dependent, with its attributes announced by a presence bitmap. Such distinctive functionality provides a flexible structure to carry type specific information, while avoiding frame bloating and minimizing its overhead.

The current 802.11be revision defines two MLE types. First, there is the basic type, which is intended to be used for beacon frames. In such type, the MLE carries only the information that is common to all interfaces. We refer, for instance, to the MLD MAC address, the set of enabled links, or the STR capability. Second, there is the multi-link request/response type, which is expected to be used during the multi-link setup. In this type, the MLE includes, apart from the common information, the complete information of those interfaces different from the advertising one, through an individual and independent field. Any parameter not advertised in the field of a given interface is considered to have the same value as the advertising one. For instance, some advertised parameters are the channel allocation (e.g., the primary channel and bandwidth), and the number of available spatial streams. Although there are currently only two defined types, further extensions of the MLE types may be aggregated. For instance, proposals are exploring to announce buffered traffic information by means of reporting a traffic indication map (TIM), or indicating changes in regards to the mapping between TID values and links.

2.3.2 Discovery and Setup

The 802.11be discovery mechanism reuses the same principles already defined in the 802.11 standard. That is, stations can gather information of nearby APs by performing the discovery process based on either a passive or active scanning. However, the introduction of MLDs makes necessary to make some updates. As explained in Section 2.3.1, beacons and probing frames only carry partial information at the multi-link level (i.e., U-MAC related-information). Such implementation, however, may take stations more time to perform the discovery process, as they should scan all the interfaces of the MLD before doing the multi-link setup. To avoid such a situation, 802.11be reuses the already defined Reduced Neighbor Report (RNR) element to announce some basic information about the different interfaces of the same AP MLD. Note that such information will belong only to the interfaces not sending the beacon frame. With that, stations can directly probe an AP MLD requesting its complete set of capabilities, parameters and operation elements of their other interfaces. To perform such probing, they must use the multi-link request/response MLE type. Although this approach may seem inefficient, since devices need to send an extra multi-link request/response, it turns out to be the opposite as it saves energy by not requiring the non-AP MLD to enable multiple radios (i.e., scan other bands/channels of the AP MLD). Also, this approach allows to reduce the air-time occupancy of management frames, as well as, the time required by the station to pass from the discovery process to the multi-link setup. Figure 2 shows the frames described.

In regards of the setup process, 802.11be will reuse the current association request/response frames by adding the extra MLE. Then, through the MLE, AP MLDs and STA MLDs will negotiate and establish their subsequent operation scheme by exchanging their capabilities. Besides, the multi-link setup process is proposed to be performed only on a single link in order to reduce overhead. It is worth mentioning that, the set of enabled links for each STA MLD is determined by measuring link qualities at all interfaces. That is, those receiving a quality value above the clear channel assessment (CCA) threshold are set as enabled, while disabled otherwise. As users may keep themselves mobile, any link listed as disabled may be added afterwards by requesting a re-setup. Analogously, the re-setup process reuses the already defined re-association request/response frames. Figure 2 shows the association response frame with the MLE.

2.3.3 Link management

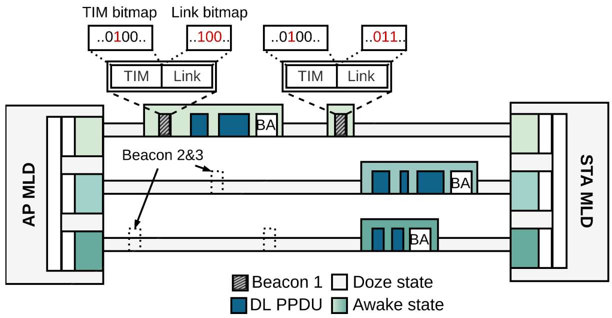
On current multi-band APs exist the main limitation that MSDUs belonging to different TIDs are not able to be sent over multiple links. It is important to recall that TIDs were created as traffic identifiers to classify different traffic types according to their QoS needs, which establish different user priorities. Past amendments used these user priorities to provide differentiation and prioritization through EDCA, by classifying each data packet into an access category, and so, associating each one to a specific MAC transmission queue with its own MAC parameters. However, even with the use of both 2.4 GHz and 5 GHz bands in the 802.11ax amendment, all the TIDs were still tied to a single link operation. With the introduction of MLO, the IEEE 802.11be promotes such a change. By default, it is suggested for AP MLDs to map all the TIDs to all links, implying that stations would be able to retrieve any type of traffic through any link. However, such condition may not be necessarily static, as APs may perform a dynamic TID transfer, allowing them to seamlessly move a TID from one link to another. In this context, a link management technique can be performed as a load balancing mechanism to avoid excessive levels of link congestion without performing any client steering. Indeed, this feature opens up new research opportunities in the area of load balancing.

2.4 Power save

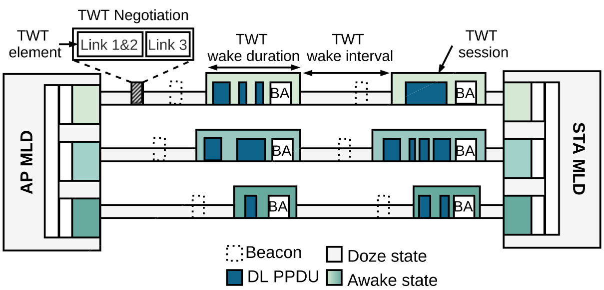
Since the Internet connectivity nowadays is mainly performed through handheld devices, power-related consumption issues must be carefully considered. To address such issue, TGbe have suggested to adopt and adapt the use of the traffic indication map (TIM) and the target wake time (TWT).

2.4.1 Traffic indication Map

The TIM mechanism is used to notify stations that its serving AP has buffered data ready to be delivered to them. Thus, the AP includes the TIM element into beacons to broadcast periodically this information. In the TIM element, APs include a bitmap formed by 2007 bits, each one of them corresponding to a unique associated station. If the bit in the bitmap corresponding to a given station is 0 the station remains in a doze state, whereas if the bit is equal to 1, it goes to an awake state, being ready to retrieve the data from the AP. Although the TIM mechanism worked for single link stations, with the introduction of MLDs it has had to be revised. In order to include the information for all the multiple links that an station may be attached to, the TGbe proposed to add a link indication field following the TIM element. Within this field, a link mapping bitmap is included, where each bit indicates a designated link. Therefore, if



(a) Multi-link TIM



(b) Multi-link TWT

Figure 3: Power save mechanisms

a STA MLD detects in the TIM element its corresponding bit set to 1, the STA MLD further checks the link mapping, finding the specific link(s) in which the buffered traffic is mapped to [1]. Figure 3a shows the described multi-link TIM indication mechanism. As shown, the multi-link TIM extends the classic TIM by providing an efficient functionality in which stations only need to awake determined interfaces on specific periods of time. Such a procedure, therefore, allows stations to minimize their power consumption, enlarging battery cycles.

2.4.2 Target Wake Time

The TWT [13] is a power save mechanism firstly included in the 802.11ah amendment, and further developed under the 802.11ax amendment. This mechanism relies on an initial negotiation, in which stations and APs agree in a common wake scheduling, namely session period (SP) or TWT session, where stations can send or receive data. To achieve this implementation, TWT requires from an initial negotiation phase to determine the SP parameters. To efficiently address a TWT operation under the MLO framework, TGbe suggests to perform TWT agreements (i.e., negotiation phase) for the different enabled links through a single link. To do so, STA MLDs include in the TWT request different TWT elements, corresponding each one to a certain link that is identified through a bitmap. Such identification is needed as the links may have different TWT parameters such as wake up time, wake interval or minimum wake duration. On the contrary, if the same parameters apply for all links, only one TWT element is needed. Figure 3b shows the described multi-link TWT mechanism. As well as in TIM, under TWT, stations move from awake and doze states when necessary, allowing to reduce their power consumption. Although the adoption of the TWT may have different performance implications, there are no works related to such issue at the time of this article being published. Indeed, their assessment is out of scope for this paper, but an interesting topic to be addressed in future works.

3 Traffic management

The existence of APs and stations with multiple interfaces makes traffic flow allocation a challenging part for the MLO. In this regard, the 802.11be U-MAC implementation should rely on a traffic manager to distribute the buffered data to a certain L-MAC, as the efficient use of the interfaces will play a critical role in terms of network performance. To that end, we introduced a set of policies in [14]. First, *Multi Link Same Load to All interfaces* (MLSA) allocates traffic equally to all interfaces. However, it was demonstrated that such operation is highly

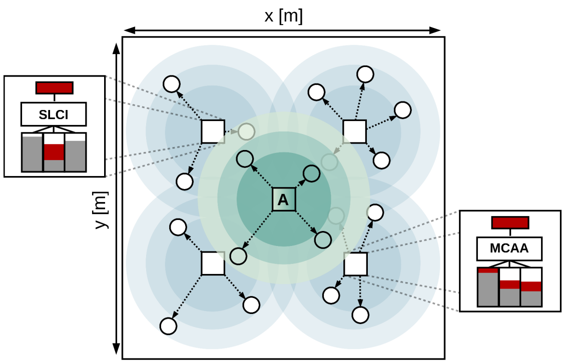


Figure 4: Scenario representation. The high, medium and low shaded areas represent the operation range for the 6 GHz, 5 GHz and 2.4 GHz bands, respectively.

inefficient, as the channel occupancy is not taken into account when driving decisions. Then, we also introduced the *Single Link Less Congested Interface* (SLCI) and the *Multi-link Congestion-Aware load balancing at flow Arrivals* (MCAA) policies. They rely on link occupancy measurements to allocate traffic either to a single or multiple bands, resulting in significant performance gains as expected. Hence, we showcased that traffic decisions must take into account the instantaneous occupancy of the channel, as well as the own traffic load. However, the non-adaptive implementation of the proposed mechanisms may not be efficient for long-lasting flows, as links may change its occupancy very rapidly. In this context, we set for further study the adoption of a dynamic strategy that not only takes into account the instantaneous channel occupancy of each interface when a flow becomes active, but tracks them continuously, so the traffic can be reallocated dynamically when changes happen.

Although the previous strategies considered all TIDs to be mapped into the multiple interfaces, the MLO opens up the possibility to perform a link-based traffic separation through the TID-to-link mapping functionality. That is, different TIDs may be mapped to different links, in order to minimize, for instance, access delays for time-sensitive traffic. Besides, such feature may be complemented by the fact that nodes' spatial distribution may create different contention-free links, specially in the 5 GHz and 6 GHz bands, as a result of favorable radio propagation conditions. Therefore, traffic with higher QoS requirements can be exclusively exchanged through those contention-free links, as long as they exist.

To showcase the benefits and drawbacks in the application of a TID-to-link mapping strategy, in this paper we introduce a new traffic allocation policy that distinguishes between traffic flows of different types. That is, traffic corresponding to data flows will be allocated to different links than the video flows. We will refer to this policy as *Video and Data Separation* (VDS), and it will allocate data flows to the 2.4 GHz or 5 GHz band, whereas video flows will be allocated at the 6 GHz band. Following the results from [14], VDS will not distribute data flows across multiple interfaces, but it will allocate the whole traffic to a single interface (i.e., either interface at 2.4 GHz or 5 GHz band, selecting always the emptiest one).

4 Performance evaluation

This section aims to conduct a flow-level performance analysis of a link-based traffic allocation strategy under the MLO framework. Simulations are done using the CSMA/CA abstraction presented in [15]. We evaluate $\mathrm{N_D} = 500$ random generated deployments, all of them with 5 BSSs as depicted in Figure 4. Each BSS consists of one AP and M stations placed around it. In every deployment, we will place the $\mathrm{BSS_A}$ at the center, and the other 4 BSSs distributed uniformly

Table 1: Evaluation setup

Parameter	**Description**
Carrier frequency	2.437 GHz/5.230 GHz/6.295 GHz
Channel bandwidth	20 MHz/40 MHz/80 MHz
AP/STA TX power	20/15 dBm
CCA threshold	-82 dBm
AP/STA noise figure	7 dB
Single user spatial streams	2
MPDU payload size	1500 bytes
Path loss	Same as [14]
Avg. data duration	$\mathrm{T_{ON}} = 3$ s
Avg. data interarrival time	$\mathrm{T_{OFF}} = 1$ s
Min. contention window	15
Packet error rate	10%
Simulation time (1 simulation)	120 s
Number of deployments	$\mathrm{N_D} = 500$

at random over a 20x20 m^2 area.

Unless stated otherwise, we consider that all MLD AP/STAs are configured with 3 wireless interfaces that operate at a different frequency band (i.e., 2.4 GHz, 5 GHz and 6 GHz). All stations are inside the coverage area of its AP for at least the 2.4 GHz interface, as shown in Figure 4. For evaluation purposes, APs' interfaces corresponding to the same frequency band are configured with the same radio channel. Except for $\mathrm{AP_A}$, which will be set either with the SLCI, MCAA, VDS or MSLA, the rest of the APs will implement either the SLCI or MCAA policies, selected with the same probability.

Only DL traffic is considered. Upon creation, stations request either a video or a data traffic flow. The two type of flows are alive during the entire simulation time, but their activity follows an ON/OFF Markovian process. The ON and OFF periods are exponentially distributed with mean duration $\mathrm{T_{ON}}$ and $\mathrm{T_{OFF}}$. For each individual video (data) flow in the ON period, the corresponding AP has to deliver ℓ_S (ℓ_E) Mbps. Table 1 details the complete set of parameters used.

Figure 5 shows different percentiles of the average throughput losses suffered by video and data flows for each policy through the different $\mathrm{N_D}$. As observed, the VDS policy is able to keep the average throughput losses under a 5% value for both video and data flows only in the 50% of the evaluated scenarios. In fact, it is noticeable that for video flows, the 5% worst case raises up to throughput losses nearly 40%, performing even worse than the MSLA policy. Such results reveal a critical drawback of the VDS policy: the traffic separation in VDS may

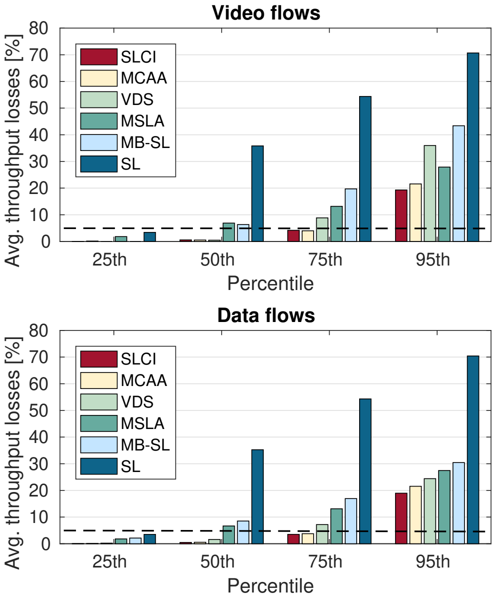


Figure 5: Avg. throughput losses for video and data flows. Black dashed line corresponds to the 5% losses threshold.

suffer from severe performance problems in conditions with high neighboring BSSs overlap, or high traffic scenarios. On the other hand, the SLCI and MCAA congestion-aware policies are able to overcome such negative issues in 75% of the scenarios due to their ability to balance the traffic load between the active links.

At last, also in Figure 5, we provide a comparison between MLO-based, and both legacy multi-band single link (MB-SL) and legacy single link (SL) deployments. Through the MB-SL, we clearly observe the advantage of adding more bands to the system, as the stations can be spread across them, reducing also their congestion levels. Although the MB-SL performance is better when compared to SL, it barely keeps the throughput losses below an acceptable 5% value for both flow types in only 25% of the scenarios. Compared to the legacy approaches, MLO is shown to be able to perform better in all the evaluated scenarios independently. In fact, we observe that with SL and MB-SL only the 25 % of the considered scenarios achieve average throughput losses below 5 %, which is increased up to the 75 % with either SLCI and MCAA. Those results, prove that the MLO framework will be a relevant new feature to WiFi, enabling currently unsuitable scenarios with SL and MB-SL solutions.

5 Open issues and challenges

Although the MLO represents a promising functionality to be implemented in next generation WLANs, the concurrent use of multiple interfaces brings new challenges to face off. In this context, we point out some open issues that require further research:

- *non-STR and legacy blindness.* This issue relates to the fact that non-STR and legacy STAs may cause different collision scenarios, as a consequence of their constrained operation. First, non-STR STAs may be unable to detect an intra-BSS transmission (either DL or UL) in one of its available links, because of performing a transmission on another. Therefore, a collision may occur if non-STR STAs attempt to transmit over that link already in use. This issue has been already tackled in [7] by allowing AP MLDs to inform non-STR stations about the channel state in other links in use by the AP MLD to prevent such a situation. On the other hand, similarly, legacy devices may not know if a transmission is taking place in others links, since they only operate in a single one. Hence, some indication, as the proposed in the non-STR case, is needed to inform legacy nodes of the activities happening in the other links.

- *Spectrum inefficiency.* Conservative approaches to avoid the IDC interference or collisions can lead to an inefficient use of the spectrum, because of suspending the backoff procedure in one link, if medium access is granted in another one. In this regard, an opportunistic backoff mechanism to maximize the spectrum utilization of non-STR nodes is proposed in [7], so transmission attempts can be resumed only when the channel state guarantees a collision with not happen.

- *Channel access fairness.* Since MLO allows to perform TXOP aggregation over different links, nodes with single link availability may experience starvation due to their higher difficulties to access the channel. Therefore, in presence of legacy stations the usage of link aggregation techniques should be limited or restricted, in order to minimize unfair situations.

- *Load balancing.* Although this has been the main topic of this paper, further research is required to fully understand which is the best strategy to balance the traffic in MLO WLANs. For instance, it is important to consider also how MLO can be used for uplink traffic, as it may require a completely different approach than its downlink counterpart. In this aspect, load balancing strategies can benefit from the use of machine learning solutions to predict future traffic and network dynamics.

6 Concluding remarks

This article has overviewed the EHT MLO framework with the objective to provide a clear and concise understanding of this upcoming disruptive WiFi functionality. The MLO framework will allow next generation of APs and stations to perform concurrent transmissions by using their multiple wireless interfaces in a coordinated way, and therefore, opening the door to both improve the network performance and achieve a more efficient use of the spectrum resources. However, further research need to be done to fully understand all new features enabled by MLO. Apart from the traffic allocation, we identified other open issues that must be tackled such as the spectrum inefficiency and the channel access fairness.

References

[1] IEEE P802.11be/D1.0 Draft Standard for Information technology— Telecommunications and information exchange between systems Local and metropolitan area networks— Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 8: Enhancements for extremely high throughput (EHT), May 2021.

[2] David López-Pérez, Adrian Garcia-Rodriguez, Lorenzo Galati-Giordano, Mika Kasslin, and Klaus Doppler. IEEE 802.11 be Extremely High Throughput: The Next Generation of Wi-Fi Technology beyond 802.11 ax. *IEEE Communications Magazine*, 57(9):113–119, 2019.

[3] Gaurang Naik, Dennis Ogbe, and Jung-Min Jerry Park. Can Wi-Fi 7 Support Real-Time Applications? On the Impact of Multi Link Aggregation on Latency. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE, 2021.

[4] Guillermo Lacalle, Iñaki Val, Oscar Seijo, Mikel Mendicute, Dave Cavalcanti, and Javier Perez-Ramirez. Analysis of Latency and Reliability Improvement with Multi-Link Operation over 802.11. In *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, pages 1–7. IEEE, 2021.

[5] Marc Carrascosa, Giovanni Geraci, Edward Knightly, and Boris Bellalta. An Experimental Study of Latency for IEEE 802.11 be Multi-link Operation. *arXiv preprint arXiv:2111.09281*, 2021.

[6] Sharan Naribole, Srinivas Kandala, Wook Bong Lee, and Ashok Ranganath. Simultaneous Multi-Channel Downlink Operation in Next Gener-

ation WLANs. In *GLOBECOM 2020-2020 IEEE Global Communications Conference*, pages 1–7. IEEE, 2020.

[7] Sharan Naribole, Wook Bong Lee, Srinivas Kandala, and Ashok Ranganath. Simultaneous Transmit-Receive Multi-Channel Operation in Next Generation WLANs. In *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–8. IEEE, 2020.

[8] Mao Yang, Bo Li, Zhongjiang Yan, and Yuan Yan. AP Coordination and Full-duplex enabled Multi-band Operation for the Next Generation WLAN: IEEE 802.11 be (EHT). In *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–7. IEEE, 2019.

[9] Abhishek Patil, George Cherian, Alfred Asterjadhi, and Duncan Ho. Multi-Link Operation: Design Discussion. IEEE 802.11 Documents. 2019. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=823&is_group=00be`.

[10] Joseph Levy and Xiaofei Wang. 802.11be Architecture/Association Discussion. IEEE 802.11 Documents. 2020. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1122&is_group=00be`.

[11] Abhishek Patil, George Cherian, Alfred Asterjadhi, and Duncan Ho. Multi-Link Aggregation - Gain Analysis. IEEE 802.11 Documents. 2019. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=764&is_group=00be&is_year=2019`.

[12] Rojan Chitrakar, Rajat Pushkarna, Yanyi Ding, and Yoshio Urabe. Multi-link element format. IEEE 802.11 Documents. 2020. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=772&is_group=00be`.

[13] Maddalena Nurchis and Boris Bellalta. Target wake time: Scheduled access in IEEE 802.11 ax WLANs. *IEEE Wireless Communications*, 26(2):142–150, 2019.

[14] Álvaro López-Raventós and Boris Bellalta. IEEE 802.11 be Multi-Link Operation: When the Best Could Be to Use Only a Single Interface. In *2021 19th Mediterranean Communication and Computer Networking Conference (MedComNet)*. IEEE, 2021.

[15] Álvaro López-Raventós and Boris Bellalta. Concurrent Decentralized Channel Allocation and Access Point Selection using Multi-Armed Bandits in multi BSS WLANs. *Computer Networks*, 180, 2020.

Dynamic Traffic Allocation in IEEE 802.11be Multi-link WLANs

Álvaro López-Raventós and Boris Bellalta

Abstract

The multi-link operation (MLO) is a key feature of the next IEEE 802.11be Extremely High Throughput amendment. Through its adoption, it is expected to enhance users' experience by improving throughput rates and latency. However, potential MLO gains are tied to how traffic is distributed across the multiple radio interfaces. In this paper, we introduce a traffic manager atop MLO, and evaluate different high-level traffic-to-link allocation policies to distribute incoming traffic over the set of enabled interfaces. Following a flow-level approach, we compare both non-dynamic and dynamic traffic balancing policy types. The results show that the use of a dynamic policy, along with MLO, allows to significantly reduce the congestion suffered by traffic flows, enhancing the traffic delivery in all the evaluated scenarios, and in particular improving the quality of service received by video flows. Moreover, we show that the adoption of MLO in future Wi-Fi networks improves also the coexistence with non-MLO networks, which results in performance gains for both MLO and non-MLO networks.

1 Introduction

The multi-link operation (MLO) is a revolutionary feature that is planed to be part of the IEEE 802.11be Extremely High Throughput (EHT) amendment [1]. By the use of multiple radio interfaces, MLO-capable devices will be able to send and receive traffic over different wireless links, allowing devices to experience higher throughput rates, as well as lower end-to-end latency delays. To support such implementation, the Task Group "be" (TGbe) has proposed several modifications to the standard, being the nodes' architecture one of the most significant. In this regard, it is suggested to split common and link-specific medium access control (MAC) functionalities into two different levels [2].

With such approach, the TGbe aims to provide nodes with a dynamic, flexible, and seamless inter-band operation. To that end, a unique MAC instance is presented to the upper-layers, while each interface is able to maintain an independent set of channel access parameters [3]. However, proper traffic balancing over the different interfaces is required to make the most out of the MLO. To implement such a load balancing, we rely on the existence of a traffic manager on top of the MLO framework, in order to apply different traffic management policies to allocate new incoming flows/packets across the enabled interfaces[1]. This approach allows to control the allocation process, ensuring a more balanced usage of the network resources.

Although MLO is gaining relevance at a very fast pace, none of the existing works have tackled how traffic allocations may be performed. For instance, existing MLO works relate to feature improvements, as the work in [4], in which the authors prove that MLO can reduce latency by means of minimizing the congestion. Similarly, [5] shows experimentally that MLO is able to reduce Wi-Fi latency in one order of magnitude in certain conditions by just using two radio interfaces. Additionally, authors in [6] suggest that the use of MLO per-se may not be sufficient enough to provide the prospected gains without a coordination between access points (AP) in high density areas. Hence, they propose a coordination framework to achieve high throughput in those circumstances. On the other hand, works in [7, 8] focus on maximizing the medium utilization, while the interference suffered by constrained nodes is minimized. As shown, none have tackled neither the implementation of a traffic manager atop MLO, nor considered the performance gains from a flow-level perspective.

A first evaluation of the capabilities of the proposed traffic manager was presented in [9]. There, it was shown —as expected— that congestion-aware policies outperform a blindfolded scheme. Additionally, and more important, it was shown that allocating the whole traffic of an incoming flow to the emptiest interface was almost as good, as proportionally distributing the flow over multiple interfaces. Such finding relies on the fact that using more interfaces, a traffic flow becomes more vulnerable to suffer a congestion episode due to the changing spectrum occupancy conditions caused by the neighboring wireless local area networks (WLANs).

In this letter, we introduce and evaluate a dynamic traffic balancing policy for the traffic manager, which periodically modifies the traffic-to-link allocation accordingly to the instantaneous channel occupancy conditions. Thus, we expect to minimize the negative impact of neighboring WLANs over the traffic flows by reacting to changes in the spectrum occupancy. The presented results show that the application of a dynamic policy has a significant impact on the spectrum

[1]We refer to those interfaces that can be effectively used, as the power received from the serving AP is above the clear channel assessment (CCA) threshold.

usage efficiency, while improving the service received by the flows. For instance, we observe that video flows are able to keep up to 95% their performance in most of the scenarios, when the dynamic policy is applied. Additionally, we showcase that the adoption of MLO in future Wi-Fi networks eases coexistence issues with non-MLO networks, which performance is improved up to 40% when surrounded by MLO BSSs.

2 Policy-based traffic management for MLO-capable WLANs

The multi-interface availability allows to naturally think of a manager in order to distribute traffic. Following the proposals of the TGbe, this logical entity should be placed at the upper MAC level, since the interface assignation is performed once traffic goes through it [10]. Once a connection[2] is established between an AP-STA pair, and traffic streams start to flow, the traffic manager is in charge to allocate the traffic to the corresponding interfaces. Such approach allows to not only achieve an efficient use of the network resources, but better control the capabilities of multi-link devices (MLDs) supporting, for instance, advanced traffic differentiation, beyond the default MLO's TID-to-link mapping functionality [2]. Figure 1 shows an schematic of a MLD architecture, with a traffic manager representation.

To perform the allocation process, the transmitting MLD gathers the instantaneous channel occupancy at each interface according to the set of enabled interfaces at the receiving node. Then, the traffic manager is able to ensure that the transmitting MLD will not allocate traffic to congested interfaces, distributing it over all of them proportionally to their occupancy. At the following, we present the different policies, which can be classified into non-dynamic and dynamic, depending on their behavior.

2.1 Non-dynamic congestion-aware policies

Under a non-dynamic strategy, each flow maintains the same traffic-to-link allocation during its lifetime. That is, upon a flow arrival, the channel occupancy is gathered, and the traffic is distributed either proportionally over multiple interfaces according to their congestion, or fully into the less congested one. We define two different non-dynamic policies:

- **Single Link Less Congested Interface (SLCI).** Upon a flow arrival, pick the less congested interface, and allocate the new incoming flow to it.

[2] A connection is defined to be a logical link, in which traffic of a certain application is exchanged between two end hosts.

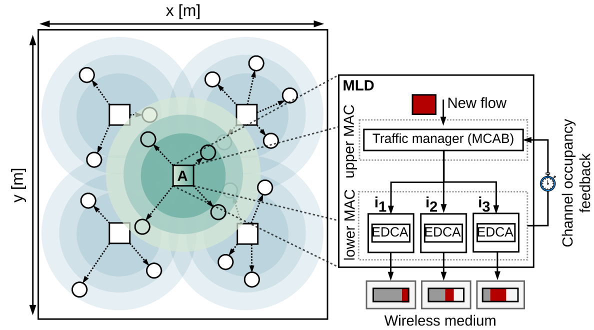


Figure 1: Scenario and architecture representation. The high, medium and low shaded areas represent the operation range for the 6 GHz, 5 GHz and 2.4 GHz bands, respectively. Colored in red is represented the traffic allocated to each interface, whereas in gray is represented the channel occupancy.

- **Multi Link Congestion-aware Load balancing at flow arrivals (MCAA).** Upon a flow arrival, distribute the new incoming flow's traffic accordingly to the observed channel occupancy at the AP, considering the enabled interfaces of the receiving station. Namely, let ρ_i the percentage of available (free) channel airtime at interface i. Then, the fraction of the flow's traffic allocated to interface i is given by $\ell_{i\in\mathcal{J}} = \ell\frac{\rho_i}{\sum_{\forall j\in\mathcal{J}}\rho_j}$, with ℓ being the traffic load, and $\mathcal{J}$ the set of enabled interfaces at the target station. If there are any other active flows at the AP, their traffic allocation remain the same as it was.

Due to their straightforward approach, the application of non-dynamic policies are well-suited for scenarios where the interfaces' congestion levels remains almost stationary. Their computational cost is low, as only few calculations are done at flow arrivals.

2.2 Dynamic congestion-aware policies

A dynamic strategy is able to periodically adjust the traffic-to-link allocation in order to follow channel occupancy changes, and so, taking the most out of the different enabled interfaces. In this regard, a traffic (re)allocation may be triggered by two different events: a new flow arrival or a periodic timer, which wakes up every δ units of time. Under both events, the channel occupancy is gathered to proportionally (re)distribute the traffic load of all active flows to any of the enabled interfaces. It is worth mention that, the dynamic reallocation of traffic is performed by adjusting the interfaces' traffic weights (i.e., traffic percentage associated to each one), which are tracked by the traffic manager at the upper MAC level. Besides, we consider such reallocation to be instantaneous. We define the following dynamic policy:

- **Multi Link Congestion-aware Load balancing (MCAB).** Upon a flow arrival or at every δ units of time, collect the channel occupancy values and sort all flows (including the incoming one) in ascending order, considering the number of enabled interfaces at the destination station (i.e., first the flows with less enabled interfaces). In case two or more flows have the same number of enabled interfaces in the destination station, they are ordered by arrival time. After, start (re)allocating the flows' traffic accordingly to the same procedure as in MCAA.

Through its dynamic implementation, the MCAB minimizes the effect of neighboring BSSs actions, as they usually result in abrupt changes in the observed congestion at each link. Therefore, such policy scheme is able to adjust the traffic allocated to each link, exploiting the different traffic activity patterns while maximizing the traffic delivery. However, it is noticeable that the MCAB gain is conditioned to perform multiple operations in shorts amounts of time, which may be impractical in high density areas, as the computational requirements to (re)distribute all flows grows with the number of active users.

3 System model

3.1 Scenario

To assess the performance of the different policies, we consider an scenario with N BSSs, each composed by an AP and M stations as depicted in Figure 1. In every scenario, we place the $\mathrm{BSS_A}$ at the center, and the other $N-1$ BSSs are distributed uniformly at random over the area of interest. To consider a random generated scenario as valid, the inter-AP distance must be equal or higher than 3 m. Otherwise, the scenario is discarded and a new one is generated. For each BSS, stations are placed within a distance $d \in [1-5]$ m and an angle $\theta \in [0-2\pi]$ from its serving AP, both selected uniformly at random.

3.2 Node operation

All APs and stations have three wireless interfaces, each one of them configured in a different frequency band (i.e., 2.4 GHz, 5 GHz and 6 GHz). For each AP-station pair, a set of enabled interfaces is established. The modulation and coding scheme (MCS) used for each enabled interface is selected accordingly to the signal-to-noise ratio (SNR). All stations are inside the coverage area of its serving AP for at least the 2.4 GHz band. For evaluation purposes, all APs' interfaces corresponding to the same band are configured with the same radio channel.

Unless otherwise stated, all APs and stations will be considered MLO-capable, using an asynchronous transmission mode [2]. It is worth mentioning that, even though we consider an asynchronous operation mode, i.e., enabling the simultaneous transmit and receive (STR) capability, the in-device coexistence interference (IDC) is not a problem, as the spectral separation between the interfaces of each node is high enough to avoid such an issue (see Table 1).

Except for $\mathrm{AP_A}$, which will be set either with the SLCI, MCAA or MCAB, the rest of the APs will implement either the SLCI or MCAA policy schemes, which will be selected with the same probability. Regarding the MCAB policy, we set the time between two adaptation periods to be δ s. In this paper, δ is set to 1 s. The MCAB dependency in regards of δ is kept out of this article due to space limitations.

3.3 Traffic considerations

Only downlink traffic is considered. The deployed stations are defined as data or video depending on the traffic that they will request. Also, only one connection is considered per station, which is set to be alive during the whole simulation time. Video traffic is modeled as a single Constant Bit Ratio (CBR) traffic flow of ℓ_S Mbps, whereas data traffic behaves following an ON/OFF Markovian model,

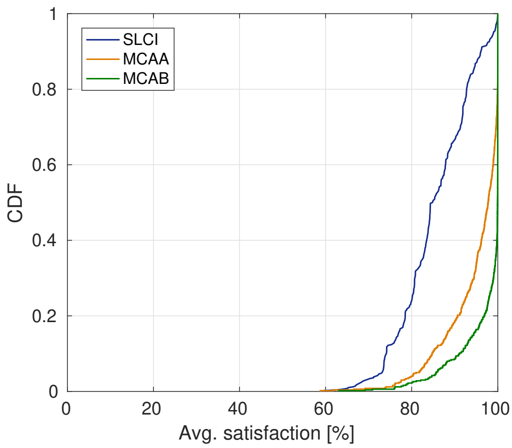


Figure 2: $\mathrm{AP_A}$'s avg. satisfaction per policy.

Table 1: Evaluation setup

Parameter	**Description**
Carrier frequency	2.437 GHz/5.230 GHz/6.295 GHz
Channel bandwidth	20 MHz/40 MHz/80 MHz
AP/STA TX power	20/15 dBm
Antenna TX/RX gain	0 dB
CCA threshold	-82 dBm
AP/STA noise figure	7 dB
Single user spatial streams	2
MPDU payload size	1500 bytes
Path loss	Same as [11]
Avg. data flow duration	$\mathrm{T_{on}} = 3$ s
Avg. data flow interarrival time	$\mathrm{T_{off}} = 1$ s
MCAB adaptation period	$\delta = 1$ s
Packet error rate	10%
Simulation time	120 s (1 simulation)
Number of simulations	$\mathrm{N_s}$ (variable)

where each ON period is treated as a new flow. Therefore, for data flows, their traffic load is ℓ_E Mbps during the ON period, and zero otherwise. Both ON and OFF periods are exponentially distributed with mean duration $\mathrm{T_{ON}}$ and $\mathrm{T_{OFF}}$, respectively.

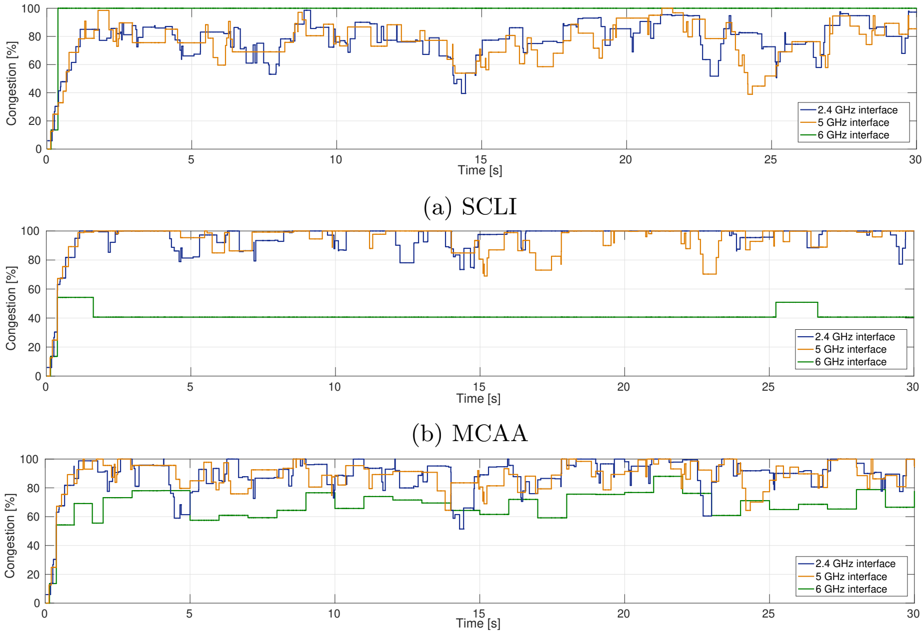
4 Performance evaluation

Flow-level simulations are performed using the Neko[3] simulation platform, which implements the CSMA/CA abstraction presented in [11]. This abstraction relies on the channel occupancy observed by each AP to calculate the allocable airtime for each flow, preserving the inherent Wi-Fi 'fair' share of the spectrum resources. Table 1 describes the complete set of parameters.

4.1 Long-lasting flows

Here, we analyze the effects between the dynamic and non-dynamic traffic-to-link allocation policies in regards of video flows (i.e., flows with constant traffic requirements, and long lifetimes). To do so, we generate $\mathrm{N_s} = 500$ scenarios, placing $N = 5$ BSSs over a 20x20 m^2 area. At the central BSS (i.e., $\mathrm{BSS_A}$), we configure

[3] The Neko simulation platform can be found in GitHub at: `https://github.com/wn-upf/Neko`



(a) SCLI

(b) MCAA

(c) MCAB

Figure 3: AP_A's congestion distribution per interface, and per policy application.

a unique video station with $\ell_S \sim U[20, 25]$ Mbps, whereas the remaining BSSs will have $M \sim U[5, 15]$ stations requesting data traffic with $\ell_E \sim U[1, 3]$ Mbps.

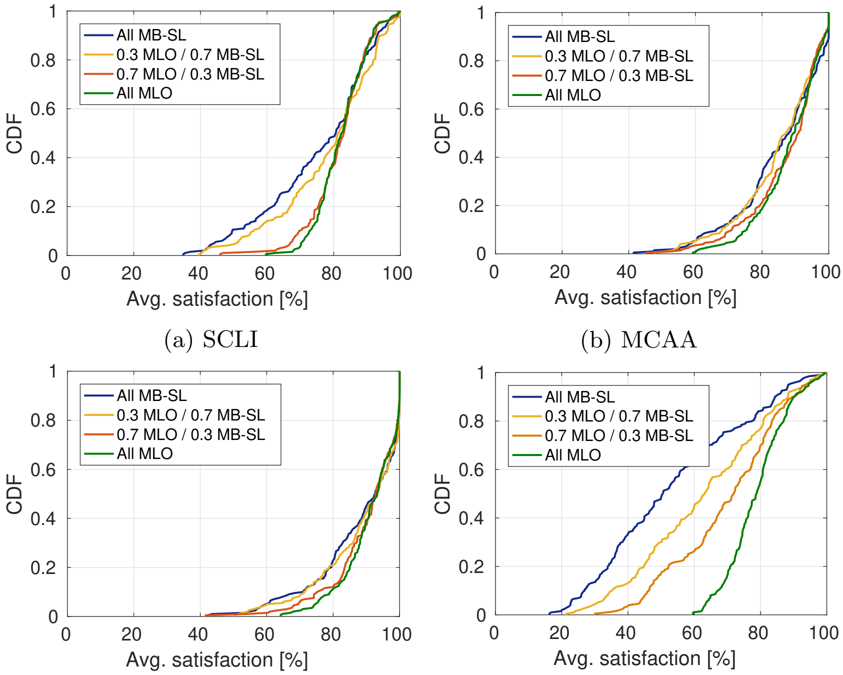
Figure 2 plots the cumulative distribution function (CDF) of the average satisfaction ($\overline{s}$) experienced by the traffic flow served by AP_A, per policy type. We define $\overline{s}$ as the sum of the satisfaction of each station divided by the total number of stations in the BSS. Also, we refer to the satisfaction of a flow as the ratio between the allocated airtime by the AP during the flow lifetime, and the total amount of airtime required. As expected, the MCAB outperforms both non-dynamic policies. For instance, it is able to increase by 17% and 6% the $\overline{s}$ in regards of the MCAA and SLCI, respectively, for the 5% worst case scenarios. Besides, we observe that the MCAB provides satisfaction values up to 95% in more than the 90% of the scenarios. This performance gains are provided by the periodic evaluation of the channel occupancy, which allows to leverage the emptiest interfaces, and so, making a better use of the available resources.

Further details are presented in Figure 3. There, we observe in detail the congestion evolution for each AP_A's interface, during the first 30 s of a single simulation. Figure 3a and Figure 3b expose the main drawbacks of SLCI and MCAA, respectively, as the temporal evolution of the congestion reveals how unbalanced the interfaces are. First, the SLCI overloads the 6 GHz link by placing the whole video flow in it, while there is still room for some traffic in the other interfaces. On the contrary, the MCAA does not leverage the fact of having empty space at the 6 GHz interface, which makes the proportional parts of the flow allocated to the 2.4 GHz and 5 GHz links to suffer from congestion. Such inefficient operation from the non-dynamic policies is shown in Figure 3c to be overcomed by the MCAB, as it reveals a more balanced use of the interfaces. However, we also observe that most of the time the congestion values for the 6 GHz interface are lower than for the other two. Such effect is related to the unequal number of neighboring nodes detected at each band. As a result, even if most of the traffic is allocated to this interface, it still manages to provide traffic with fewer congestion episodes.

4.2 Coexistence with legacy networks

Wi-Fi's constant evolution makes newer devices, which implement up-to-date specifications, to coexist with others with less capabilities. As a result, last generation devices may decay in performance due to its coexistence with legacy ones. To assess if Multi-Band Single Link (MB-SL) BSSs affects the performance of MLO ones, we analyze four different cases in which we increment the fraction of MLO BSSs around the central one from 0, to 0.3, 0.7, and 1. To do so, we generate $\mathrm{N_s} = 200$ scenarios, placing $N = 11$ BSSs. At the central BSS (i.e.,$\mathrm{BSS_A}$), we configure a single video station with $\ell_S \sim U[20, 25]$ Mbps, whereas the remaining BSSs will have $M \sim U[5, 15]$ stations requesting background data traffic of $\ell_E \sim U[1, 3]$ Mbps. It is worth mention that, MB-SL APs are equipped with 3 interfaces, considering the associated stations are distributed across all three bands uniformly at random.

Figures 4a, 4b, and 4c show the CDF of the $\overline{s}$ for each policy. Regardless of the policy used, the central $\mathrm{BSS_A}$ experiences a negative trend when it is surrounded by more legacy BSSs, as the results show lower satisfaction values when so. Although the MCAA and MCAB experience low gains when increasing the number of MLO BSSs, the SLCI presents a 17% improvement for the 25th percentile, when comparing the performance results between the best and the worst (i.e., all MLO and all MB-SL, respectively) cases. Such an improvement is caused by the higher link availability from the neighboring BSSs to allocate traffic, which also avoid to overload the interfaces by the use of congestion-aware policies. On the other hand, comparing policies, we find that the MCAB outperforms the other ones. Specially, we observe that the MCAB tends to perform better in the cases with more MB-SL neighboring BSSs. In those situations, the $\overline{s}$ when using



(a) SCLI (b) MCAA

(c) MCAB (d) MB-SL

Figure 4: Coexistence performance per policy type, and MB-SL.

MCAB is above 94% in half of the scenarios, whereas below 85% when using the SLCI and MCAA. Although the optimal solution will be to avoid coexistence issues by not having any legacy BSSs, the periodic channel evaluation of the MCAB adds the required flexibility to minimize those negative effects.

At last, Figure 4d shows the avg. satisfaction when $\mathrm{BSS_A}$ is set as a legacy MB-SL with the aim to observe if the presence of MLO devices will benefit legacy ones. As previously, we incremented the fraction of MLO BSSs from 0, to 0.3, 0.7, and 1. Figure 4a reveals that legacy MB-SL BSSs can benefit from the fact of having MLO BSSs around them, as the improvement is highly noticeable. In fact, we observe that between the best and worst cases the satisfaction increases by a 40% for half of the scenarios evaluated. Then, from the perspective of a legacy BSS, the adoption of the MLO by other BSSs represents also a performance improvement.

5 Conclusions and future work

In this letter, we assessed the implementation of a traffic manager to perform traffic allocation on top of MLO-capable BSSs. We evaluated three policy schemes under different conditions to shed some light on the potential performance gains of dynamic policies in comparison to non-dynamic ones. Under a wide variety of scenarios, our results shown that dynamic policies should be applied in presence of long-lasting flows, since their frequent adaptation to the instantaneous congestion conditions allows to minimize the effect of the neighboring AP MLDs' actions. By the nature of video flows, it has been found also that the MCAB is able to maximize the traffic delivery by keeping a satisfaction ratio of 95% for most of the evaluated scenarios. Under coexistence conditions, we observe that an excessive number of legacy BSSs may harm the performance of MLO ones. However, we found that the MCAB is able to reduce the negative impact of legacy BSSs by almost 10% compared to MCAA, as it is able to react to changes in the channel occupancy of the different interfaces.

Regarding future research, we plan to extend current traffic management policies to also support link aggregation at channel access. Regarding improving QoS provisioning in next generation Wi-Fi networks, traffic differentiation policies should be further investigated in presence of heterogeneous stations, providing solutions that go beyond the default TID-to-link mapping functionality. Finally, we also consider the redesign of the traffic management module as part of an end-to-end Software Defined Networking solution, closely working with an external controller in charge of multiple APs to properly allocate traffic flows to interfaces.

References

[1] IEEE P802.11be/D1.0 Draft Standard for Information Technology. Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks. Part 11. Amendment 8: Enhancements for Extremely High Throughput (EHT), May 2021.

[2] Álvaro López-Raventós and Boris Bellalta. Multi-link Operation in IEEE 802.11 be WLANs. *arXiv preprint arXiv:2201.07499*, 2022.

[3] Joseph Levy and Xiaofei Wang. 802.11be Architecture/Association Discussion. IEEE 802.11 Documents. 2020. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1122&is_group=00be`.

[4] Gaurang Naik, Dennis Ogbe, and Jung-Min Jerry Park. Can Wi-Fi 7 Support Real-Time Applications? On the Impact of Multi Link Aggregation on

Latency. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE, 2021.

[5] Marc Carrascosa, Giovanni Geraci, Edward Knightly, and Boris Bellalta. An experimental study of latency for ieee 802.11 be multi-link operation. *arXiv preprint arXiv:2111.09281*, 2021.

[6] Mao Yang, Bo Li, Zhongjiang Yan, and Yuan Yan. AP Coordination and Full-duplex Enabled Multi-band Operation for the Next Generation WLAN: IEEE 802.11 be (eht). In *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–7.

[7] Sharan Naribole, Wook Bong Lee, Srinivas Kandala, and Ashok Ranganath. Simultaneous Transmit-Receive Multi-Channel Operation in Next Generation WLANs. In *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–8. IEEE.

[8] Sharan Naribole, Srinivas Kandala, Wook Bong Lee, and Ashok Ranganath. Simultaneous Multi-Channel Downlink Operation in Next Generation WLANs. In *GLOBECOM 2020-2020 IEEE Global Communications Conference*, pages 1–7. IEEE, 2020.

[9] Álvaro López-Raventós and Boris Bellalta. IEEE 802.11be Multi-Link Operation: When the Best Could Be to Use Only a Single Interface. In *2021 19th Mediterranean Communication and Computer Networking Conference (MedComNet)*, pages 1–7.

[10] Mark Hamilton and Joseph Levy. Mld architecture part 2. IEEE 802.11 Documents. 2021. [Online]. Available: `https://mentor.ieee.org/802.11/documents?is_dcn=1111&is_group=00be`.

[11] Álvaro López-Raventós and Boris Bellalta. Concurrent Decentralized Channel Allocation and Access Point Selection using Multi-Armed Bandits in multi BSS WLANs. *Computer Networks*, 180, 2020.