THE UNIVERSITY
of EDINBURGH

Distributed Algorithms for Optimized Resource Management of LTE in Unlicensed Spectrum and UAV-enabled Wireless Networks

Ursula Challita

Doctor of Philosophy
Institute of Computing Systems Architecture
School of Informatics
University of Edinburgh
2018

Abstract

Next-generation wireless cellular networks are morphing into a massive Internet of Things (IoT) environment that integrates a heterogeneous mix of wireless-enabled devices such as unmanned aerial vehicles (UAVs) and connected vehicles. This unprecedented transformation will not only drive an exponential growth in wireless traffic, but it will also lead to the emergence of new wireless service applications that substantially differ from conventional multimedia services. To realize the fifth generation (5G) mobile networks vision, a new wireless radio technology paradigm shift is required in order to meet the quality of service requirements of these new emerging use cases. In this respect, one of the major components of 5G is self-organized networks. In essence, future cellular networks will have to rely on an autonomous and self-organized behavior in order to manage the large scale of wireless-enabled devices. Such an autonomous capability can be realized by integrating fundamental notions of *artificial intelligence* (AI) across various network devices.

In this regard, the main objective of this thesis is to propose novel self-organizing and AI-inspired algorithms for optimizing the available radio resources in next-generation wireless cellular networks. First, heterogeneous networks that encompass licensed and unlicensed spectrum are studied. In this context, a deep reinforcement learning (RL) framework based on long short-term memory cells is introduced. The proposed scheme aims at proactively allocating the licensed assisted access LTE (LTE-LAA) radio resources over the unlicensed spectrum while ensuring an efficient coexistence with WiFi. The proposed deep learning algorithm is shown to reach a mixed-strategy Nash equilibrium, when it converges. Simulation results using real data traces show that the proposed scheme can yield up to 28% and 11% gains over a conventional reactive approach and a proportional fair coexistence mechanism, respectively. In terms of priority fairness, results show that an efficient utilization of the unlicensed spectrum is guaranteed when both technologies, LTE-LAA and WiFi, are given equal weighted priorities for transmission on the unlicensed spectrum. Furthermore, an optimization formulation for LTE-LAA holistic traffic balancing across the licensed and the unlicensed bands is proposed. A closed form solution for the aforementioned optimization problem is derived. An attractive aspect of the derived solution is that it can be applied *online* by each LTE-LAA small base station (SBS), adapting its trans-

mission behavior in each of the bands, and without explicit communication with WiFi nodes. Simulation results show that the proposed traffic balancing scheme provides a better tradeoff between maximizing the total network throughput and achieving fairness among all network flows compared to alternative approaches from the literature. Second, UAV-enabled wireless networks are investigated. In particular, the problems of interference management for cellular-connected UAVs and the use of UAVs for providing backhaul connectivity to SBSs are studied. Specifically, a deep RL framework based on echo state network cells is proposed for optimizing the trajectories of multiple cellular-connected UAVs while minimizing the interference level caused on the ground network. The proposed algorithm is shown to reach a subgame perfect Nash equilibrium upon convergence. Moreover, an upper and lower bound for the altitude of the UAVs is derived thus reducing the computational complexity of the proposed algorithm. Simulation results show that the proposed path planning scheme allows each UAV to achieve a tradeoff between minimizing energy efficiency, wireless latency, and the interference level caused on the ground network along its path. Moreover, in the context of UAV-enabled wireless networks, a UAV-based on-demand aerial backhaul network is proposed. For this framework, a network formation algorithm, which is guaranteed to reach a pairwise stable network upon convergence, is presented. Simulation results show that the proposed scheme achieves substantial performance gains in terms of both rate and delay reaching, respectively, up to 3.8 and 4-fold increase compared to the formation of direct communication links with the gateway node. Overall, the results of the different proposed schemes show that these schemes yield significant improvements in the total network performance as compared to current existing literature. In essence, the proposed algorithms can also provide self-organizing solutions for several resource management problems in the context of new emerging use cases in 5G networks, such as connected autonomous vehicles and virtual reality headsets.

Lay Summary

The combination of billions of connected devices, petaflops of computing resources, and advanced communication capabilities that enable real-time interactions is leading to the creation of systems on a scale and complexity level that is beyond the ability of humans to manage and control. Management and operation of these systems require an extremely high degree of intelligent automation, which is one of the main components of 5G networks. In this regard, the main scope of this thesis is to leverage the use of optimization and artificial intelligence techniques for managing the wireless resources in next-generation cellular networks.

First, we focus on spectrum management in 5G networks. 5G networks will support more spectrum bands in order to adapt to the vast increase in the number of connected devices. In particular, we focus on the operation of LTE over the unlicensed spectrum. We propose a proactive resource management scheme for the coexistence of LTE and WiFi in the unlicensed spectrum. The proposed proactive mechanism can be combined with traditional reactive schemes thus guaranteeing the promised quality-of-service for different types of applications. Furthermore, we propose a traffic balancing framework for enabling small base stations to steer their traffic between the licensed and the unlicensed bands based on the congestion level on each band, respectively.

Second, we investigate UAV-enabled wireless networks. In particular, we propose an interference-aware path planning scheme for cellular-connected UAVs. The main intention of the proposed scheme is to optimize the trajectories of multiple cellular-connected UAVs while minimizing the interference level caused on the ground network. Moreover, we propose a UAV-based on-demand aerial backhaul scheme for providing backhaul connectivity to small base stations for scenarios in which high-speed ground backhaul links are either unavailable or limited in capacity.

Acknowledgements

First and foremost, I am greatly thankful to my supervisors, Murray Cole and Walid Saad, for their persistent support to ensure the success of this thesis. In particular, I would like to thank Murray Cole for his constant support and trust during my PhD studies, without which this thesis would have not been completed. My sincere gratitude for Walid Saad for offering me the opportunity to join his lab as a visiting researcher at Virginia Tech. His scientific intuition and insightful discussions have helped exceptionally inspire and enrich my growth as a researcher. Without his guidance, broad technical expertise, and constant motivation, this thesis would not have been possible.

I would like to thank also my thesis committee members, Mehdi Bennis and David Laurenson, for accepting to evaluate this thesis and for their valuable comments that have improved the presentation of this work.

I am very grateful for the opportunity to work with Li Dong, who helped me through a new research topic and was always available to offer support. I was also fortunate to work with Christian Bettstetter who gave me some insightful feedback. I am also thankful to Mingzhe Chen, and other colleagues and friends at the University of Edinburgh and Virginia Tech for I was fortunate to have met up, discussed, and worked with during my PhD journey.

My deepest appreciation to Joy whose dedication and persistent confidence in me have provided an invaluable support and inspiration during the stressful moments of this thesis. Finally, I would like to acknowledge the commitment, sacrifice and support of my parents and family, who have always motivated me.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Ursula Challita*)

To my late grandparents, Wadih and Hanne.

Table of Contents

List of Figures

List of Tables

Abbreviations

3D	3-dimensional
3GPP	Third generation partnership project
4G	Fourth generation
5G	Fifth generation
A2A	Air-to-air
A2G	Air-toground
ABS	Almost blank subframe
ACK	Acknowledgement
AI	Artificial intelligence
ANN	Artificial neural network
ASA	Authorized shared access
BPTT	Backpropagation through time
BR	Best response
BS	Base station
BW	Bandwidth
CA	Carrier aggregation
CAPEX	Capital expenditure
CCA	Clear channel assessment
CFP	Contention-free periods
CSMA/CA	Carrier sense multiple access with collision avoidance
CTS	Clear-to-send
CW	Contention window
DCF	Distributed coordination function
DFS	Dynamic frequency selection
DIFS	DCF Interframe Space
DL	Downlink
DNN	Deep neural network
DTX	Discontinuous transmission

eICIC	Enhanced ICIC
ESN	Echo state network
ESP	Echo state property
FCC	Federal communications commission
FDD	Frequency division duplex
FNN	Feed-forward neural network
GNEP	Generalized Nash equilibrium problem
GRU	Gated recurrent unit
HE	Homo egualis
HetNet	Heterogeneous network
ICIC	Intercell interference
IoT	Internet of Things
ISM	Industrial, scientic and medical
KKT	Karush-Kuhn-Tucker
LBT	Listen-before-talk
LMS	Least mean squares
LoS	Line-of-sight
LSA	Licensed shared access
LSTM	Long-short term memory
LTE	Long term evolution
LTE-LAA	Licensed assisted access LTE
LTE-U	LTE-unlicensed
LWA	LTE-WiFi link aggregation
MAC	Medium access control
MBS	Macro base station
MIMO	Multiple-input-multiple-output
MINLP	Mixed integer nonlinear optimization problem
MLP	Multi-layer perceptron
mmWave	Millimeter wave
MSE	Mean square error
MUE	Macro cell user equipment
Multi-Rat	Multiple radio access technology
NE	Nash equilibrium
NLoS	Non-line-of-sight
NZ	Non-zero

OFDM	Orthogonal frequency division multiplexing
OFDMA	Orthogonal frequency division multiple access
OPEX	Operational expenditure
PF	Proportional fairness
PHY	Physical
QoS	Quality-of-service
RB	Resource block
RC	Reservoir computing
ReLU	Rectied linear units
ResNet	Residual network
RF	Radio frequency
RL	Reinforcement learning
RL-LSTM	Reinforcement learning based on LSTM
RLS	Recursive least squares
RNN	Recurrent neural network
RTRL	Real-time recurrent learning
SBS	Small base station
SC-FDMA	Single-carrier frequency-division multiple access
SDL	Supplemental downlink
SDN	Software-defined networking
SGD	Stochastic gradient descent
SIFS	Short interframe space
SINR	Signal-to-interference-plus-noise ratio
SNR	Signal-to-noise ratio
SPNE	Subgame perfect Nash equilibrium
STA	Station
SUE	Small cell user equipment
TDD	Time division duplex
TNT	Total network throughput
TPC	Transmit power control
TVWS	TV white space
UAV	Unmanned aerial vehicle
UAV-UE	Unmanned aerial vehicle user equipment
UE	User equipment
UL	Uplink

U-NII	Unlicensed national information infrastructure
WAP	WiFi access point
WLAN	Wireless local area network

Chapter 1

Introduction

1.1 Motivation

Wireless cellular networks are witnessing a radical change that is mainly driven by new wireless use cases that significantly vary from the conventional voice-based or multimedia services. These emerging use cases, such as connected autonomous vehicles, remote controlled robots, haptic feedback-enabled drones and fixed wireless access – rivalling fiber capacity – for residential homes, will ultimately impact both consumers and industries. In essence, the emergence of these new cases will result in the proliferation of bandwidth-intensive wireless applications such as high definition video streaming and multimedia services which can drastically strain the capacity of current wireless cellular networks. Alongside the dramatic growth in mobile data traffic, future mobile networks must be able to deliver ultra-reliable, low-latency communication [10–13], that is adaptive in real-time to a rich and dynamic Internet of Things (IoT) environment. For instance, a real-time and low-latency communication system is essential for the autonomous behavior of future wireless devices such as drones and connected vehicles. Therefore, these emerging new applications and use cases have created a radically different networking environment whose quality-of-service (QoS) requirements mandate a fundamental change in the way in which wireless networks are modelled, analyzed, designed, and optimized. Consequently, the goal of the fifth generation (5G) mobile networks is to expand the broadband capability of mobile networks, and to provide the promised QoS for consumers and for various industries and society at large, hence unleashing the potential of the IoT.

To realize this, current wireless cellular networks will require major changes

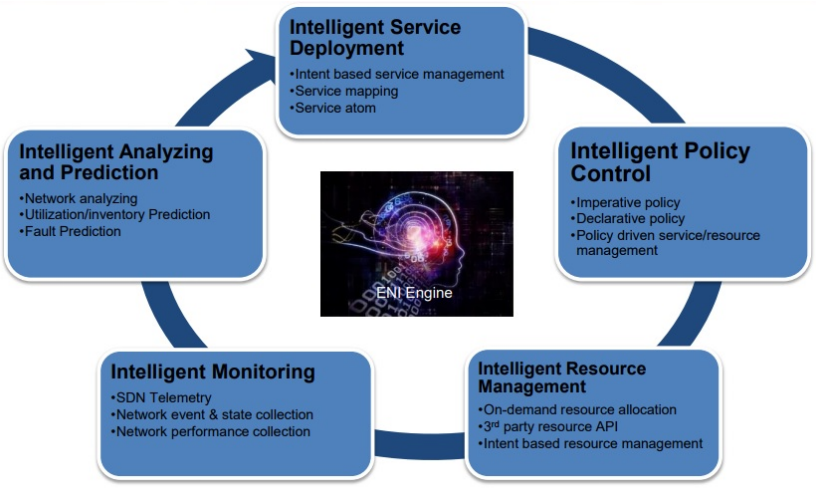

Figure 1.1. Experiential networked intelligence for future wireless cellular networks [2].

in the implementation and deployment of the networking infrastructure as well as the management of the available radio resources. In particular, 5G networks will be based on software-defined networking (SDN), network functions virtualization, dense small cell deployments, heterogeneous spectrum, millimeter wave (mmWave) communications, and device-to-device communications. While the main ingredients for 5G have been identified, integrating them into a truly harmonious wireless system requires instilling *intelligent functions across both the edge and core of the network*. These intelligent functions must be able to adaptively exploit the wireless system resources and generated data, in order to optimize network operation and guarantee, in real-time, the QoS needs of emerging wireless and IoT services. Such network intelligence can be mainly realized by integrating *artificial intelligence* (AI) [14] and online optimization techniques across the wireless infrastructure and end-user devices.

In particular, next-generation wireless cellular networks will rely heavily on *experiential networked intelligence*, alongside traditional optimization techniques. Experiential learning is described as the process of learning through experience i.e., "learning through reflection on doing". Such experiential intelligence techniques will be integrated at various components of the 5G network such as intelligent service deployment, intelligent policy control, intelligent resource management, intelligent monitoring, and intelligent analysis and prediction [2], as shown in Fig. 1.1. AI will essentially allow next-generation cellular networks to dynamically adapt to changing context in real-time and thus enabling autonomous and self-adaptive operations. For instance, machine learning-based AI techniques can be used to investigate and predict network and user behavior so as to pro-

vide users' information for solving diverse wireless networking problems such as cell association, spectrum management, computational resource allocation, and cached content replacement. These techniques can also be used to extract information from the user environment thus providing a wireless network with the ability to predict the users' future behaviors and, hence, design an optimal strategy to improve the QoS and reliability. Therefore, machine learning-based AI optimization algorithms will provide inherently *self-organizing, self-healing, and self-optimizing solutions* for a broad range of problems within the context of network optimization and resource management. Such AI-driven self-organizing solutions are particularly appropriate for ultra dense wireless networks in which classical centralized and distributed optimization approaches can no longer cope with the scale and heterogeneity of the network. Consequently, machine learning solutions will enable next-generation wireless cellular networks to evolve from a network of *connected things* to a network of *connected intelligence*.

Therefore, the main goal of this thesis is to leverage machine learning and optimization techniques for optimizing the management of the available radio resources in next-generation wireless cellular networks. We focus on spectrum management for 5G networks, and in particular, the coexistence of long term evolution (LTE) and WiFi in the unlicensed spectrum. We further address some of the challenges that arise in the context of unmanned aerial vehicles (UAVs) for wireless communications. Specifically, we tackle the problem of interference management for cellular-connected UAVs and the use of UAVs for providing backhaul connectivity to small base stations (SBSs). Next, we highlight our contributions in the aforementioned areas. Then, we give an overview on the structure of this thesis.

1.2 Thesis Contributions

The main objective of this thesis is to propose novel algorithms and methodologies for optimizing the management of the available radio resources in future wireless cellular networks. The main contributions of this thesis are summarized as follows:

- First, we propose a proactive LTE-WiFi coexistence scheme that allows a better utilization of the unlicensed spectrum as compared to the existing

literature. In this regard, current existing literature [15–18] relies on a reactive approach in which data requests are first initiated and, then, radio resources are allocated based on their corresponding delay tolerance value. Nevertheless, this sense-and-avoid approach can cause an underutilization of the spectrum due to the impulsive reconfiguration of the spectrum usage that does not account for the future dynamics of the network. Meanwhile, in a proactive approach, rather than reactively responding to incoming demands and serving them when requested, an SBS can predict traffic patterns and determine future off-peak times so that incoming traffic demand can be properly allocated over a given time window. Hence, a proactive coexistence scheme allows an efficient use of the available spectrum thus resulting in a lower collision probability for WiFi nodes and more transmission opportunities for LTE. In this context, our main contribution can be summarized as follows:

- We propose a *novel deep reinforcement learning (RL) algorithm based on long short-term memory (LSTM) cells* [19] for proactively allocating LTE resources over the unlicensed spectrum. The proposed framework enables the SBSs to autonomously learn which unlicensed channels to use along with the corresponding channel access probability on each channel taking into account future environmental changes, in terms of wireless local area network (WLAN) activity on the unlicensed channels and LTE traffic loads.
- We show that our proposed deep RL algorithm reaches a mixed-strategy Nash equilibrium (NE) upon convergence. We also show that the gain of the proposed proactive resource allocation scheme and the optimal size of the prediction time window is a function of the traffic pattern of the dataset under study.
- Simulation results using real data traces show that the proposed scheme can yield up to 28% and 11% gains over a conventional reactive approach and a proportional fair (PF) coexistence mechanism, respectively. The results also show that the proposed framework prevents WiFi performance degradation for a densely deployed licensed assisted access (LTE-LAA) network.
- The proposed approach can also be combined with online machine

learning to accommodate traffic changes, by properly re-training the weights of the neural network. Moreover, the proposed scheme can be combined with the conventional reactive approach thus serving different types of applications.

- The results of this work have been published in [20] and [21].

- Second, we present a holistic approach for LTE-LAA small cell traffic balancing by jointly optimizing the use of the licensed and unlicensed bands. In this regard, current existing literature [15, 17, 22–24] either do not consider the operation of LTE-LAA SBS in the licensed band while optimizing its operation over the unlicensed bands alongside WiFi or consider both bands but do not optimize them jointly [1]. This can however lead to a suboptimal resource allocation when seen globally. For instance, it can result in an over-utilization of the unlicensed band by LTE-LAA SBS and a decrease in WLAN performance. In this context, our main contributions can be summarized as follows:

  - We present a formulation of the optimization problem for holistic traffic balancing that seeks PF coexistence of WiFi, small cell and macro cell user equipments (UEs) by adapting the transmission probability of the LTE-LAA SBS in the licensed and unlicensed bands. The rationale behind this formulation is for the LTE-LAA SBS to switch between or aggregate licensed and unlicensed bands depending on the interference/traffic level and number of active UEs in each cell.
  - We derive a closed-form solution for the aforementioned optimization problem. An attractive aspect of the proposed approach is that it can be applied online by each LTE-LAA SBS, adapting its transmission behavior in each of the bands, and without explicit communication with WiFi nodes.
  - Simulation results reveal that approaches focusing on coexistence in one band while ignoring the other cause load imbalance and a decrease in the total network throughput and/or fairness. Meanwhile, the proposed approach, owing to its holistic nature, results in improved network performance as it achieves a better tradeoff between maximizing the total network throughput and attaining fairness among all network flows while also providing better LTE-WiFi coexistence.

  - The results of this work have been published in [25].

- Third, we propose an interference-aware path planning scheme for a network of cellular-connected UAVs. Unlike current wireless UAV connectivity that rely on short-range communication range (e.g., WiFi, bluetooth, and radio waves), providing cellular connectivity to the UAVs allows beyond line-of-sight control, low latency, real time communication, and ubiquitous coverage. Such cellular-connected UAV-user equipments (UAV-UEs) will thus enable a myriad of applications ranging from real-time video streaming to surveillance. In this regard, current existing literature [26–29] is limited to studying the impact that cellular-connected UAVs have on the ground network. Indeed, the existing literature [26–29] does not provide any concrete solution for optimizing the performance of a cellular network that serves both aerial and ground UEs in order to overcome the interference challenge that arises in the context of cellular-connected UAVs. UAV trajectory optimization is essential in such scenarios. An online path planning that accounts for wireless metrics is vital and would, in essence, assist in addressing the aforementioned interference challenge along with new improvements in the design of the network, such as 3-dimensional (3D) frequency reuse. Such a path planning scheme allows the UAVs to adapt their movement based on the rate requirements of both aerial UAV-UEs and ground UEs, thus improving the overall network performance. The problem of UAV path planning has been studied mainly for non-UAV-UE applications [30–33] with [34] being the only work considering a cellular-connected UAV-UE scenario. Nevertheless, the work in [34] is limited to one UAV and does not account for the interference that cellular-connected UAVs cause on the ground network during their mission. Moreover, the work in [34] relies on offline optimization techniques that cannot adapt to the uncertainty and dynamics of a cellular network. In this context, our main contributions can be summarized as follows:

  - We propose a novel *deep RL framework based on echo state network (ESN) cells* [35] for optimizing the trajectories of multiple cellular-connected UAVs in an online manner. This framework will allow cellular-connected UAVs to minimize the interference they cause on the ground network as well as their wireless transmission latency. Two

important features of our proposed algorithm are adaptation and generalization; the UAVs can take decisions for unseen network states, based on the reward they got from previous states.

- We show that the proposed algorithm reaches a subgame perfect Nash equilibrium (SPNE) upon convergence. Moreover, we derive an upper and lower bound for the altitude of the UAVs thus reducing the computational complexity of the proposed algorithm.
- Simulation results show that the proposed approach improves the trade-off between energy efficiency, wireless latency, and the interference level caused on the ground network. The results also show that each UAV's altitude is a function of the ground network density and the UAV's objective function and is an important factor in achieving the UAV's target.
- The results of this work have been published in [36] and [37].

- Fourth, we propose a novel backhaul scheme that relies on UAVs as an on-demand flying network. In fact, a fundamental challenge for the efficient operation of a dense SBS deployment is to provide an economical and ubiquitous backhaul connectivity to the SBSs that would allow the routing of the traffic to/from the SBSs from/to the core network. In this regard, current existing solutions for SBS backhauling [38–41] do not account for scenarios in which the high-speed ground backhaul is either congested, unavailable, or limited in capacity. In such scenarios, the backhaul connectivity of SBSs can become a bottleneck thus degrading the performance of the radio access network. To address this challenge, the authors in [41–43] propose a vertical fronthaul/backhaul framework based on UAVs and free-space optics communication. In [42, 43], the authors propose an optimization formulation for the association problem of the UAVs and the SBSs but ignoring the design of the multi-hop links among the UAVs. Therefore, one challenging area which remains relatively unexplored is the dynamic formation of the aerial graph that connects the UAVs to the core network. Indeed, the existing prior art does not provide an efficient scheme, in terms of achievable rate and delay, for the formation of a multi-hop aerial network for SBS backhauling. In this context, our main contributions can be summarized as follows:

  - We propose an online framework that relies on a network of UAVs as an on-demand flying network linking the SBSs and the core network in scenarios where the ground backhaul is either unavailable or limited in capacity. The design of the aerial backhaul network is formulated as a network formation game [44] in which the players are the UAVs. The objective of the proposed game is to allow the UAVs to autonomously learn which air-to-air (A2A) and air-to-ground (A2G) links to form in order to guarantee the connectivity of the SBSs to the core network.
  - To solve this game, we propose a dynamic network formation algorithm that is guaranteed to reach a pairwise stable network upon convergence. Moreover, to ensure an efficient backhauling process between the UAVs, we incorporate the notion of *virtual force fields* [45] into our dynamic algorithm. In essence, virtual forces allow the UAVs to adjust their location dynamically based on the links they want to form.
  - Simulation results show that the proposed network formation algorithm achieves substantial performance gains in terms of both rate and delay reaching, respectively, up to 3.8 and 4-fold increase compared to the formation of direct communication links with the gateway node.
  - The results of this work have been published in [46].

This thesis yielded the following journals and conference publications:

- **U. Challita**, A. Ferdowsi, M. Chen, and W. Saad, "Artificial Intelligence for Wireless Connectivity and Security of Cellular-Connected UAVs", Submitted to *IEEE Wireless Communications*, available at arXiv:1804.05348.

- **U. Challita**, W. Saad, and C. Bettstetter, "Interference Management for Cellular-Connected UAVs: A Deep Reinforcement Learning Approach", Submitted to *IEEE Transactions on Wireless Communications (TWC)*, available at arXiv:1801.05500.

- **U. Challita**, L. Dong and W. Saad, "Proactive Resource Management for LTE in Unlicensed Spectrum: A Deep Learning Perspective", *IEEE Transactions on Wireless Communications (TWC)*, vol. 17, no. 7, July 2018.

- A. Ferdowsi, **U. Challita**, W. Saad, and N. Mandayam, "Robust Deep Reinforcement Learning for Security and Safety in Autonomous Vehicle Systems", *the 21st IEEE International Conference on Intelligent Transportation Systems (ITS)*, Nov. 2018, Maui, Hawaii, USA.

- A. Ferdowsi, **U. Challita**, and W. Saad "Deep Learning for Reliable Mobile Edge Analytics in Intelligent Transportation Systems", Submitted to *IEEE Network*, available at: arXiv:1712.04135.

- M. Chen, **U. Challita**, W. Saad, C. Yin, and M. Debbah "Machine Learning for Wireless Networks with Artificial Intelligence: A Tutorial on Neural Networks", Submitted to *Proceedings of IEEE*, available at: arXiv:1710.02913.

- **U. Challita**, W. Saad, and C. Bettstetter, "Deep Reinforcement Learning for Interference-Aware Path Planning of Cellular-Connected UAVs", *IEEE International Conference on Communications (ICC)*, May 2018, Kansas City, USA.

- **U. Challita**, and W. Saad, "Network Formation in the Sky: Unmanned Aerial Vehicles for Multi-hop Wireless Backhauling", *IEEE Global Communications Conference (Globecom)*, Dec. 2017, Singapore.

- **U. Challita**, L. Dong, and W. Saad, "Deep Learning for Proactive Resource Allocation in LTE-U Networks", *European Wireless (EW)*, May 2017, Dresden, Germany.

- **U. Challita**, and M. K. Marina "Holistic Small Cell Traffic Balancing across Licensed and Unlicensed Bands", *The 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, Nov. 2016, Malta.

- C. Hasan, M. K. Marina, and **U. Challita**, "On LTE-WiFi Coexistence and Inter-Operator Spectrum Sharing in Unlicensed Bands: Altruism, Cooperation and Fairness", *The 17th International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, July 2016, Paderborn, Germany.

In what follows, we give an overview on the structure of this thesis.

1.3 Structure

The rest of this thesis is structured as follows:

- **Chapter 2** gives a general overview on topics of direct relevance to the work presented in this thesis. The first part of this chapter presents general background on next-generation cellular networks. The second part of this chapter provides background information about the underlying techniques of the proposed solutions in this thesis such as game theory and machine learning.

- **Chapter 3** presents a proactive resource management scheme for the co-existence of LTE and WiFi in the unlicensed band. The proposed scheme allows different SBSs to autonomously update their channel selection and channel access probabilities based on the traffic of WLAN on each of the unlicensed channels.

- **Chapter 4** introduces a holistic approach for LTE-LAA small cell traffic balancing by jointly optimizing the use of the licensed and unlicensed bands. The proposed scheme enables the LTE-LAA small cell to autonomously switch between or aggregate licensed and unlicensed bands depending on the interference/traffic level and the number of active UEs in each band.

- **Chapter 5** explores an interference-aware path planning scheme for a network of cellular-connected UAVs. The proposed scheme enables each UAV to achieve a tradeoff between maximizing energy efficiency and minimizing both wireless latency and the interference level caused on the ground network along its path.

- **Chapter 6** investigates a novel UAV-based backhaul framework for linking the SBSs and the core network in scenarios where the ground backhaul is either unavailable or limited in capacity. Based on the proposed scheme, the UAVs would either serve as a bridge among the SBSs and relay the traffic to a nearby gateway node (with core network access) or act as an intermediate relay point between different backhaul transceivers.

- **Chapter 7** summarizes the work presented in this thesis and provides some future direction in the scope of this thesis.

Chapter 2

Background

In this chapter, we give an overview on topics of direct relevance to the work presented in this thesis. The first part of this chapter presents general background information on next-generation cellular networks. In particular, we overview the different types of spectrum for future cellular networks and we focus on the coexistence of LTE-LAA and WiFi in the 5 GHz unlicensed spectrum. Moreover, we give a general background on the IoT with a focus on UAV-enabled wireless networks and connected autonomous vehicles. Finally, we discuss different 5G backhaul solutions and elaborate more on different fairness notions used in wireless networks. Background information related to the underlying analytical techniques of the proposed solutions is presented in the second part of this chapter. We overview game theory, specifically, network formation games, and deep learning techniques such as LSTM and deep ESNs.

2.1 Next-Generation Cellular Networks

5G cellular networks promise to deliver enhanced mobile broadband, mission critical services, massive IoT, and vehicular communications [3, 47]. In essence, beyond the need for high data rates – which has been the main driver of the wireless network evolution in the past decade – next-generation wireless networks must be able to deliver ultra-reliable, low-latency communication [10–13], that is adaptive, in real-time to a rich and dynamic IoT environment. In particular, the 5G key elements, given in Fig. 2.1, can be summarized as follows:

- **Higher data capacity:** Area capacity 1000× fourth generation (4G) mobile networks and edge capacity 100×4G.

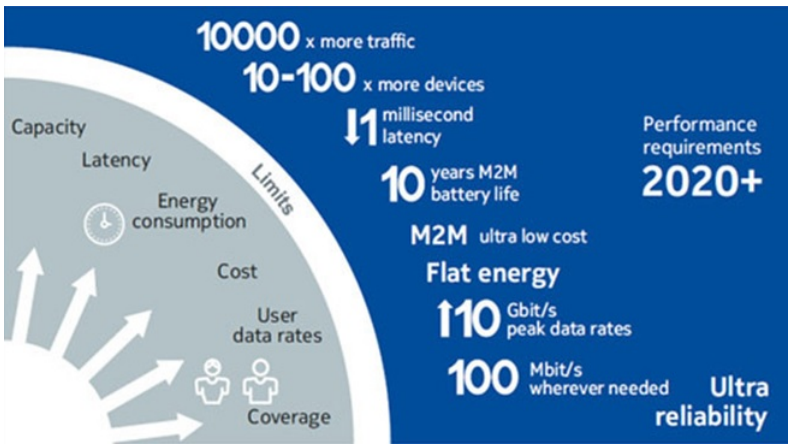

Figure 2.1. Key elements of 5G.

- **High reliability and low latency:** Roundtrip air latency of 1 msec.
- **Cost reduction:** For instance, SBSs should be $10 - 100\times$ cheaper than macro base stations (MBSs).
- **Machine type devices:** such as virtual reality headsets, connected autonomous vehicles, and UAVs.
- **Low energy consumption:** 5G will allow a huge boost to network capacity while maintaining a flat energy consumption.
- **Lower outage probability:** and thus better coverage.

Therefore, to achieve these key elements and thus cope with this ongoing and rapid evolution of wireless services, much research has been dedicated to investigate the optimal cellular network architecture within the context of the emerging fifth generation wireless networks (e.g., see [47] and references therein). For instance, the $1000\times$ data rate will be mainly achieved through combined gains in the following categories:

- **Dense SBS deployment:** 5G networks will consist of a dense deployment of SBSs with reduced cell sizes and will essentially rely on offloading in order to improve the area spectral efficiency. This in turn would result in more active nodes per unit area and Hz.
- **More efficient use of existing spectrum:** Unlike early generations of wireless cellular networks that operate *exclusively* over the licensed spectrum, 5G networks would transmit over the 5 GHz unlicensed spectrum

and other underutilized spectrum bands such as the 3.8 GHz, alongside the conventional licensed spectrum.

- **More spectrum:** 5G networks will leverage the large amount of available bandwidth (BW) in the mmWave frequencies. In fact, mmWave communications will provide high data rates by exploiting directional antennas and transmitting over a large BW.
- **Increased spectral efficiency:** This can be achieved through advances in multiple-input-multiple-output (MIMO) to support more bits/s/Hz per node.

Therefore, the drastic growth in mobile broadband data will require drastic improvements in spectrum efficiency, small cell network densification and large amounts of additional spectrum. As such, the combination of more nodes per unit area, more spectrum, and more bits/s/Hz per node, will compound higher data rates in bits/s per unit area. Next, we give a general overview on the spectrum for future 5G networks while focusing on the operation of LTE in the 5 GHz unlicensed band. This is then followed by a discussion on the internet of things and in particular, UAVs and connected autonomous vehicles. Finally, we summarize the different proposed solutions for 5G SBS backhauling.

2.1.1 Spectrum Management and Coexistence of Multiple Radio Access Technologies (Multi-RATs)

To cope with the unprecedented increase in mobile data traffic and realize the envisioned 5G services, significant enhancement of per-user throughput and overall system capacity is required [48]. Such an enhancement can be achieved through advanced physical (PHY)/medium access control (MAC)/network technologies and efficient methods of spectrum management. In fact, one of the main advancements in the network design for 5G networks relies on the integration of multiple different radio access technologies, as shown in Fig. 2.2. With multi-RAT integration, a mobile device can potentially transmit data over multiple radio interfaces such as LTE and WiFi, at the same time, thus improving its performance [49]. Moreover, a multi-RAT network allows fast handover between different RATs and, thus, providing seamless mobility experience for the UEs. Therefore, the integration of different RATs results in an improvement in the

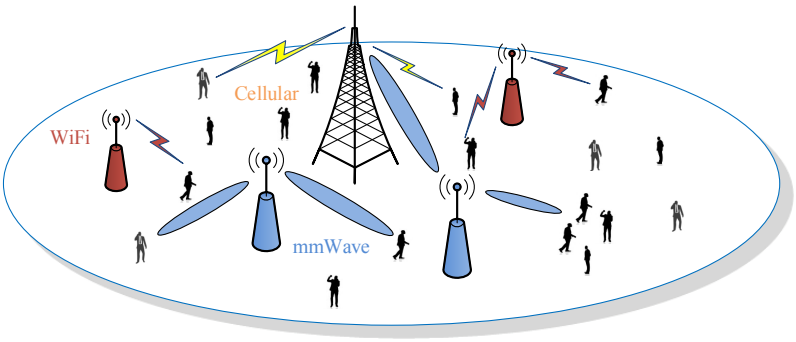

Figure 2.2. Multiple radio access networks [3].

overall utilization of the available radio resources and, thus, an increase in the system's capacity. It also guarantees a consistent service experience for different UEs irrespective of the served RAT and facilitates the management of the resulting network.

Spectrum management is also regarded as another key component of 5G networks [4, 47]. The radio spectrum is mainly divided into two types, licensed and unlicensed. The former includes frequencies that can be reserved for a specific use while the unlicensed spectrum is publicly owned and can be used without the need of a license. The unlicensed spectrum has been an attractive band for the operation of many current wireless technologies such as WiFi hotspots, cordless phones, RFID, and many applications such as medical equipment, wireless headsets and keyboards, remote car door openers and industrial systems. Meanwhile, static assignments of dedicated and exclusive licenses over many years have dominated spectrum policy for mobile services in order to assure high service quality and reliability for the subscriber. However, due to the increase in the traffic demand and the need for more spectrum, regulators are now considering *spectrum sharing* with other incumbent technologies over various frequency bands.

Therefore, unlike early generations of cellular networks that operate exclusively on the sub-6 GHz (microwave) licensed band, 5G networks are expected to transmit over the conventional licensed sub-6 GHz band, sub-6 GHz unlicensed spectrum and the high-frequency mmWave band, which can be either licensed or unlicensed [50–52]. The shared spectrum mechanism will be used as a *complement* to dedicated licensed spectrum. Future mobile networks would operate over the unlicensed (sub-6 GHz and high-frequency mmWave band) and licensed shared access (LSA) bands alongside the sub-6 GHz licensed spectrum via the carrier aggregation (CA) feature that has been defined in LTE Advanced, Rel-10. The CA feature allows a maximum of five component carriers, contiguous or

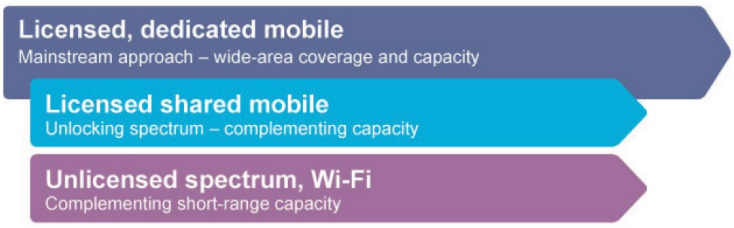

Figure 2.3. Spectrum complements for future mobile networks [4].

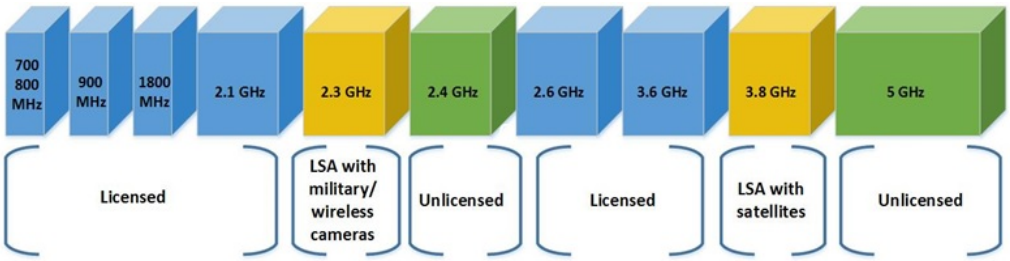

Figure 2.4. Sub-6 GHz spectrum for future mobile networks [5].

non-contiguous, to be aggregated in order to form a single LTE radio interface. It defines three different use cases for the available spectrum: licensed and licensed CA, licensed and unlicensed CA, and licensed and shared access CA as shown in Fig. 2.3.

Within this spectrum landscape, on the one hand, the classical LTE sub-6 GHz licensed band provides a reliable, albeit limited and scarce resource. On the other hand, the sub-6 GHz unlicensed bands can be used to serve best effort traffic only since the operation over this spectrum should account for the presence of other coexisting technologies. Fig. 2.4 summarizes the sub-6 GHz spectrum which is indeed valuable due to its channel characteristics that essentially provide better penetration and thus a higher coverage range as compared to high-frequency bands. Meanwhile, the high-frequency mmWave spectrum (licensed or unlicensed) has a large amount of available BW and can provide multi-gigabit communication services, however, the uncertainty and dynamic channel variations of the mmWave band due to blockage make it unreliable. Therefore, a *multi-mode BS* operating over the sub-6 GHz licensed, sub-6 GHz unlicensed, and the high-frequency mmWave bands can exploit the different characteristics and availability of the frequency bands thus providing robust and reliable communication links for the end users [52]. In what follows, we give a brief overview on LSA, sub-6 GHz unlicensed spectrum, and the high-frequency mmWave spectrum.

2.1.1.1 Licensed Shared Access

The sub-6 GHz spectrum bands are largely fragmented and challenging to harmonize. These bands have been allocated to different primary systems, and hence their exclusive use for mobile broadband would be hard to accomplish. LSA, also known as authorized shared access (ASA) is a new regulatory concept which was initially introduced by Nokia and Qualcomm in 2011 in order to support more spectrum for mobile broadband [53]. It provides a solution for bands that cannot be easily refarmed or totally vacated by their incumbent users and where spectrum usage is underutilized and infrequent. For example, spectrum for military radar may have been allocated on countrywide basis, but the radar operations may only be utilizing it at certain places such as coastline [53]. Therefore, LSA is intended to be applicable to frequency bands that are not being used by licensees at particular locations or times, extending a certain level of quality of service to all rights holders, including sub-licensees. For instance, the 2.3-2.4 GHz band has been identified in Europe as the first potential application area for spectrum sharing between a mobile network operator and incumbent spectrum users under an LSA framework. To realize such a sharing mechanism, most of the existing literature suggest the use of an ASA geo-location database by a network operator, which instructs the relevant base station (BS) to aggregate the ASA channels that are free at a particular time. Therefore, the geo-location database manages the spectrum resource allocation based on predefined policies and availability thus ensuring protection of the primary UEs.

2.1.1.2 Sub-6 GHz Unlicensed Spectrum

The unlicensed spectrum refers to the frequency bands for which no exclusive licenses are granted and on which unregistered UEs may operate wireless devices without a federal communications commission (FCC) license. This spectrum can be categorized into two main bands – the industrial, scientific and medical (ISM) and the unlicensed national information infrastructure (U-NII) bands. The ISM bands are generally confined to the 900 MHz and 2.45 GHz range, while the UNII band covers the higher 5.15-5.35 GHz and 5.725-5.825 GHz range. Numerous applications use the ISM/U-NII bands, including cordless phones, wireless garage door openers, wireless microphones, vehicle tracking and amateur radio as well as a number of access technologies such as 802.11 (WiFi), 802.15.1 (Bluetooth)

and 802.15.4 (ZigBee).

The 2.4 GHz ISM band provides 13 overlapping channels spread equally over the frequencies in addition to a 14th channel used in Japan with the center frequency 2.484 GHz. This leaves available only three non-overlapping channels in the 2.4 GHz band. On the other hand, the 5 GHz band is divided up into sub-bands named U-NII-1, U-NII-2, U-NII-2e, and U-NII-3 where U-NII-3 is not freely available worldwide. In total, there are 23 non-overlapping channels in the 5 GHz band where 4 of these have limitations based on location. The 23 available non-overlapping channels in the 5 GHz band can provide a possibility for easier planning of an interference-free and stable wireless communication as compared to three non-overlapping channels in the 2.4 GHz band. Moreover, the availability of 23 non-overlapping channels allows for increased density i.e., more wireless devices can be connected in the same radio environment. This in turn makes the 5 GHz unlicensed band an attractive spectrum for future mobile networks. Nevertheless, to operate over the unlicensed band, future mobile networks must comply with the requirements of the FCC's Part 15 Rules, which are summarized as follows [54]:

- **Dynamic frequency selection (DFS):** DFS is a mechanism that enables devices to operate over the 5 GHz unlicensed frequency bands without causing interference to existing radar systems in this frequency band. In particular, DFS requires coexisting devices to vacate a particular channel on the 5 GHz band if the level of a radar signal is detected to be above a certain threshold value on that channel.

- **Listen-before-talk (LBT):** LBT or clear channel assessment (CCA) is a technique whereby a radio transmitter is required to sense its radio environment for a period of at least 20 μsec before it starts a transmission. The LBT regulation is imposed only in markets such as Japan, India and Europe.

- **Discontinuous transmission (DTX):** DTX limits the use of a channel for a maximum transmission duration set by the regulations (1 to 10 msec). For example, the maximum transmission duration in Japan is set to 4 msec.

- **Transmit power control (TPC):** TPC limits the maximum transmission power on the unlicensed bands based on the country and the area of the

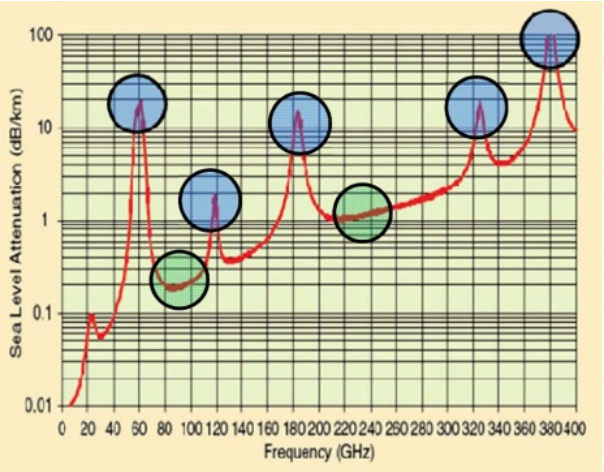

Figure 2.5. The mmWave band between 30 and 300 GHz. Atmospheric absorption across mmWave frequencies in dB/km. The attenuation caused by atmospheric absorption is 0.012 dB over 200 m at 28 GHz and 0.016 dB over 200 m at 38 GHz. Frequencies from 70 to 100 GHz and 125 to 160 GHz also have small loss. Figure from [6].

band being used. For example, FCC allocates both the 900 MHz and 2.4 GHz band with a maximum transmission power of 1 W, whereas ETSI allocates only the 2.4 GHz band with 100 mW as the maximum transmission power (900 MHz is used for GSM cell phones in Europe). For the 5 GHz band, the lower portion of the spectrum is restricted to indoor use, with a maximum transmit power of 200 mW, while the upper part allows higher transmission power, typically 1 W [55].

2.1.1.3 High-frequency mmWave Spectrum

MmWave frequencies between 30 and 300 GHz are a new frontier for cellular communications that offer the promise of orders of magnitude greater bandwidths combined with further gains via beamforming and spatial multiplexing from multielement antenna arrays [56]. In fact, the available spectrum at these high frequencies offer more than 200 times the current cellular spectrum, as depicted in Fig. 2.5. Operation over the mmWave band will essentially allow multi-Gbps wireless transmission. However, a main difference between microwave and mmWave frequencies is the sensitivity of the latter to blockages. In essence, mmWave frequency bands exhibit a path loss exponent of 2 for *line-of-sight (LoS)* propagation and 4 (plus additional power loss) for non-line-of-sight (NLoS) links [57]. Indeed, mmWave cellular research will need to incorporate sensitivity to blockages, more complex channel models, and the use of higher density infrastructure and relays [58]. Therefore, despite the potential of mmWave cellular systems, there are

a number of key challenges for realizing the vision of cellular networks in these bands, which are summarized as follows [56]:

- **Short range and directional communication:** According to Friis' transmission law, the free space omnidirectional path loss grows with the square of the frequency. Therefore, signal attenuation is high for mmWave. However, the smaller wavelength of the mmWave signals enable proportionally greater antenna gain for the same physical antenna size. Consequently, the higher frequencies of the mmWave signals do not result in any increased free space propagation loss, provided that the antenna area remains fixed and suitable directional transmissions are used. Meanwhile, the reliance on highly directional transmissions will necessitate novel design changes to current cellular systems.

- **Shadowing:** mmWave signals are extremely susceptible to shadowing. For example, materials such as brick can attenuate signals by as much as 40-80 dB [56] and the human body itself can result in a 20-35 dB loss. Meanwhile, humidity and rain fades are not an issue in cellular systems. On the contrary, the human body and other outdoor materials are very reflective for mmWave signals and thus act as scatterers for mmWave propagation.

- **Rapid channel fluctuations and intermittent connectivity:** For a given velocity, channel coherence time is linear in the carrier frequency and is therefore very small in the mmWave range. For instance, the Doppler spread at 60 km/h at 60 GHz is over 3 kHz, hence the channel will change in the order of hundreds of microseconds, which is considerably much faster than current cellular systems. Moreover, the presence of obstacles would lead to dramatic swings in path loss as a result of high levels of shadowing. Consequently, this rapid change in the path loss would result in changes for cell association and thus an intermittent connectivity for the UEs.

- **Multiuser coordination:** Current applications for mmWave transmissions are generally for point-to-point links. Nevertheless, to achieve high spatial reuse and spectral efficiency, cellular systems require simultaneous transmissions over multiple interfering links, which therefore necessitates new mechanisms for coordinating these transmissions in mmWave networks.

Here, it is important to note that, due to the limited range of the mmWave signals, most of the cellular applications for mmWave systems have focused on small-cell, outdoor deployments such as in campus- and stadium-like settings where the UEs could obtain relatively unobstructed connections to SBSs operating over the mmWave band. It is also worth noting that, due to the inherent limitations of the mmWave propagation, mmWave cellular systems cannot provide uniform, robust high capacity across a range of deployments. As such, mmWave networks will be inherently heterogeneous. For instance, mmWave cells will have to coexist with a conventional microwave cellular overlay for universal coverage due to the limitation in their coverage range [52].

Having defined the different types of bands over which 5G networks will operate, next, we elaborate more on the operation of LTE over the unlicensed spectrum, an approach known as licensed assisted access LTE.

2.1.2 Licensed Assisted Access LTE

The significant amount of unlicensed band in the 5 GHz band has recently attracted operators to deploy LTE in unlicensed spectrum bands, an approach known as licensed-assisted access using LTE or LTE-Unlicensed (LTE-U). In this section, we give a general overview on LTE-LAA along with its opportunities and challenges.

2.1.2.1 Spectrum for LTE-LAA

According to the third generation partnership project (3GPP) release 13, LTE is expected to start using the unlicensed spectrum alongside the licensed spectrum [59]. The frequency band of most interest for 3GPP is the 5 GHz band, which has up to 500 MHz of spectrum with 23 non-overlapping channels of 20 MHz BW channels available globally as compared to three non-overlapping 20 MHz channels in the 2.4 GHz frequency band. Moreover, the 5 GHz spectrum is not shared with widely used devices that utilize Bluetooth and cordless phone technology and equipment such as microwave ovens which therefore implies less interference upon transmitting over the 5 GHz band than the 2.4 GHz band.

The 5 GHz unlicensed band is divided into three main categories: U-NII-1 (5.15 - 5.25 GHz), U-NII-2 (5.25 - 5.35 GHz), extended UNII-2 (5.470 - 5.725 GHz), and U-NII-3 (5.725-5.825 GHz). These bands are categorized based on

the radio frequency (RF) requirements (e.g., the maximum transmit power, antenna gain) and other regulatory requirements that define the operation areas and license requirement. For instance, the U-NII-1 band is restricted for indoor operations with a power limit of 200 mW in order to protect co-channel mobile satellite service feeder links. Due to these restrictions, this sub-band is most suitable for U-NII devices providing communication links between indoor devices separated by short distances. On the contrary, the U-NII-2 band can be used for indoor and outdoor operations with a maximum transmit power of 1 W and is suitable for communication links within and between buildings such as the case in campus-wide local area networks. This band is shared with the federal government radio location service, earth exploration satellite service, and space research service. Finally, the U-NII-3 band (also known as U-NII/ISM due to overlap with the ISM band) is intended for outdoor and indoor transmissions with a maximum transmit power of 4 W and is suitable for communication links within and among buildings and over long distances through the use of high gain antennas. This U-NII sub-band is shared with federal government radio location, amateur, ISM, and other Part 15 devices.

Therefore, to select the most suitable band for the operation of LTE-LAA, one has to consider the availability of BW, regulation restrictions, and the inter-modulation interference with existing international mobile telecommunications bands. For example, there exists some cross-band emission issues (e.g., the inter-modulation interference between 5.47-5.725 GHz band and the 1.8 GHz band) which prohibit their aggregation with other LTE licensed bands. Moreover, the 5.15-5.25 GHz and 5.25-5.35 GHz bands are currently used by residential WLAN which therefore makes these bands undesirable for LTE-LAA operation. Consequently, the most suitable bands for LTE-LAA are 5.725-5.850 GHz. This is due to the large available BW, indoor and outdoor operation, high transmit power, relatively lower interference than other sub-bands, no inter-modulation interference with most of LTE licensed bands, and fewer regulatory requirements in most regions.

2.1.2.2 LTE-LAA Operation Modes

The goal of LTE-LAA is to leverage the unlicensed spectrum as a complement to licensed spectrum to offload best-effort traffic data through the CA framework, while critical control signalling, mobility, voice and control data will always be

transmitted over the licensed band. Therefore, the primary carrier always uses licensed spectrum in order to ensure that the connection is maintained while the secondary carrier uses unlicensed spectrum. Three modes of operation for LTE-LAA have been defined [59]: standalone, supplemental downlink (SDL), and carrier aggregation. In the *standalone mode*, both control plane and data plane traffic are carried over the unlicensed spectrum and can therefore be adopted in interference-limited scenarios only. Unlike the *supplemental downlink mode* that uses the unlicensed spectrum only for the downlink (DL), the *carrier aggregation mode* will use the unlicensed spectrum for both the DL and uplink (UL) directions. When operating in the SDL mode, LTE SBSs can perform most of the necessary operations to ensure reliable communications. For instance, the SBS would check the availability of a particular channel and should aim to select a channel that is either free or slightly loaded.

2.1.2.3 LTE-LAA Deployment Scenarios

LTE-LAA is considered as an attractive spectrum for the operation of small cells (femto, pico, and micro cells) due to the low power restriction in the unlicensed band which in turn results in a smaller coverage range. Two deployment scenarios are considered: co-located and non-co-located. In the former scenario, an SBS transmits on both bands, while in the latter an SBS transmits only on the unlicensed band. Therefore, in the non-co-located scenario, a UE receives its control and critical information from the MBS that transmits over the licensed band and the best effort data from the SBS that operates over the unlicensed band only. It is the SBS that would decide on how to split data transmission between the licensed and the unlicensed bands on the basis of an LTE subframe duration (1 msec).

LTE-LAA relies on the existing core network for the backhaul and other capabilities such as security and authentication. However, a new RF support is needed at both the UE and the SBS in order to accommodate the new frequency bands. For the CA mode, the SBS and the UE should incorporate new capabilities that would ensure proper sharing of the unlicensed frequencies such as the LBT feature and radar detection. On the other hand, these features are needed only at the SBS side for the SDL mode.

2.1.2.4 LTE-LAA Opportunities

LTE-LAA provides new opportunities for both, the UEs and the network operators. In what follows, we summarize the gains achieved upon the transmission of LTE over the unlicensed spectrum.

2.1.2.4.1 Data Rate Improvement LTE-LAA allows an increase in the transmission BW by aggregating the licensed and the unlicensed bands. This in turn allows the UEs to achieve a higher data rate and better QoS. Moreover, as compared to WiFi, LTE offers a higher data rate and hence a more efficient use of the unlicensed spectrum. For example, LTE offers a peak data rate of 1.09 Gbps at 80 MHz and 4x4 MIMO while that of WiFi is 0.75 Gbps. In [60], Qualcomm experiments have shown that, for the same number of LTE-LAA nodes, LTE-LAA provides twice the capacity as compared to WiFi. Moreover, the same capacity can be provided with fewer nodes with LTE-LAA versus WiFi.

2.1.2.4.2 Better Spectrum Efficiency As compared to WiFi and other technologies operating over the unlicensed band, LTE offers better spectrum efficiency. This is due to the fact that LTE is a scheduled system and therefore overcoming any intra-system contention. In addition, LTE offers various performance enhancing techniques (channel quality indicator scheduling, fast link adaptation, L1 hybrid automatic repeat request, coordinated multipoint, CA, intercell interference (ICIC) management, etc.) and better QoS management and control. For instance, it can operate at lower signal-to-interference-plus-noise ratio (SINR) values as compared to WiFi systems and supports eight different QoS class identifiers with different performance requirements such as guaranteed bit rates, priority, delay, and packet error loss rate.

2.1.2.4.3 Reliable and Predictable Performance As opposed to the unlicensed spectrum, the licensed spectrum offers a reliable and predictable performance. Therefore, through CA, LTE-LAA guarantees UE experience in which control and signalling information is always transmitted over the licensed band. In essence, the licensed spectrum ensures service quality in case the unlicensed spectrum becomes unusable for any reason, such as reduced coverage and interference from other systems.

2.1.2.4.4 Wider Coverage Range LTE-LAA SBSs provides a larger coverage range, as compared to WiFi access points (WAPs). For example, at 5.9 GHz and power level of 20 dBm, LTE offers a coverage range of 38 m while that of WiFi is 28 m.

2.1.2.4.5 Seamless Indoor/Outdoor Mobility What makes LTE an attractive technology over WiFi is that the former is a cellular technology while the latter is a hotspot technology that does not account for mobility. LTE offers automatic subscriber authentication, integrated backhaul, and seamless mobility to and from the MBSs. In LTE-LAA, the UE mobility is controlled via the licensed band which allows robust mobility through licensed anchor carrier. This in turn enables LTE-LAA to offer seamless indoor/outdoor mobility and to support handover and service continuity.

2.1.2.4.6 Management of a Single Network LTE-LAA can be integrated into a mobile operator's existing radio network setup which therefore allows the management of a single network. This in turn results in a unified LTE network with common authentication, security, and management thus simplifying the overall network maintenance. In essence, a unified network allows joint operation, load balancing, and interference management along with a flexible charging policy.

2.1.2.4.7 Increase in Revenue LTE-LAA is fully transparent to the LTE core network and thus does not require any upgrade for the evolved packet core elements. This in turn keeps the capital expenditure (CAPEX) of LTE-LAA deployment at a reduced level cost due to the existence of the backhaul, core network, and the SBSs that are already deployed for licensed LTE carriers. Therefore, operators would only need to upgrade the SBSs so that they can operate over the unlicensed spectrum.

From an operational perspective, LTE-LAA allows a unified operation and management between the licensed and the unlicensed spectrum, including OAM configuration, authorization, charging, and RRM management through a common RAN framework across the whole network. This in turn keeps the operational expenditure (OPEX) at a lower cost especially that the secondary cell can be activated/deactivated by the primary cell dynamically within few milliseconds.

Therefore, given the fact that the unlicensed spectrum is free to use and the low CAPEX and OPEX cost, LTE-LAA offers operators an opportunity for increasing their revenue.

2.1.2.5 LTE-LAA Challenges

To reap the benefits from the operation of LTE over the unlicensed spectrum, several challenges need to be addressed such as traffic balancing across the licensed and the unlicensed bands, the coexistence of multiple LTE-LAA operators, and the coexistence of LTE with other incumbent technologies such as WLAN. The main LTE-LAA challenges are summarized as follows.

2.1.2.5.1 Meeting the Regulatory Requirements Regulatory requirements should be met for any technology operating on the unlicensed band. This in turn necessitates an update for the MAC layer of LTE in order to support DFS, LBT, DTX, and TPC.

2.1.2.5.2 AP Discovery An LTE-LAA SBS should be aware of any other technology transmitting on the unlicensed band in order to allow a fair coexistence and coordination with these incumbent technologies. For instance, WAPs use beacon request/report pairs for AP discovery. Similarly, an LTE-LAA SBS should be able to detect nearby WAPs through new discovery mechanisms such as beacon signalling or fingerprint techniques.

2.1.2.5.3 MAC Layer Time Synchronization Inter-operator and intra-operator time synchronization is required for the coexistence of multiple operators on the unlicensed band. In fact, an LTE SBS can transmit at the beginning of a subframe only which in turn necessitates synchronization among different operators. Moreover, WAPs can only transmit during the contention-free periods (CFP) which in turn necessitates synchronization between LTE SBSs and WiFi nodes. In particular, the beginning of an LTE subframe should coincide with the beginning of a WiFi CFP.

2.1.2.5.4 Traffic Balancing An SBS would essentially operate over the licensed and unlicensed spectrum via the CA framework. As such, intelligent traffic offloading between both bands is vital. On the licensed band, an LTE-LAA

SBS causes ICIC on other neighbouring SBSs. Meanwhile, an LTE-LAA SBS can cause disruption to the operation of WiFi APs on the unlicensed band. Therefore, an LTE-LAA SBS should automatically balance its traffic in the licensed and unlicensed bands based on channel gains, interference, and traffic conditions on both bands. In this regard, new strategies are required in order to steer the traffic between the licensed and the unlicensed bands in a way that would maximize the total network performance.

2.1.2.5.5 Global Solution A global solution for the operation of LTE-LAA over the unlicensed band is required in order to address the regulatory requirements of different regions. For instance, LBT is necessary in markets such as Europe and Japan but not in the US. Therefore, a global solution is required in order to avoid market fragmentation.

2.1.2.5.6 Multi-operator Coexistence A new policy for sharing the unlicensed spectrum among different LTE-LAA operators is necessary. This is needed in order to guarantee a fair share among different operators and to avoid spectrum overlapping which in turn can result in low spectrum efficiency.

2.1.2.5.7 Multi-technology Coexistence To allow an efficient operation of LTE over the unlicensed band, LTE-LAA should coexist fairly with other incumbent technologies such as WLAN. However, the limitation for the efficient coexistence of LTE and WiFi is due to the lack of an inter-RAT coordination and mutual interference management when these two technologies share the same unlicensed spectrum. For instance, preliminary results show that WAP throughput could drop by 70% and even 100%, depending on the scenario, if mutual interference is not mitigated [61]. In fact, the LTE-WiFi coexistence challenge stems primarily from the difference in the PHY and MAC layers of these two technologies.

The PHY layer of WiFi is based on orthogonal frequency division multiplexing (OFDM) and a time division duplex (TDD) mode. During one OFDM symbol (4 μs), a device would operate over the entire BW and thus only one device can transmit during a WiFi time slot. The employed MAC mechanism of WiFi is carrier sense multiple access with collision avoidance (CSMA/CA) also known as the distributed coordination function (DCF). In this mechanism, a device is

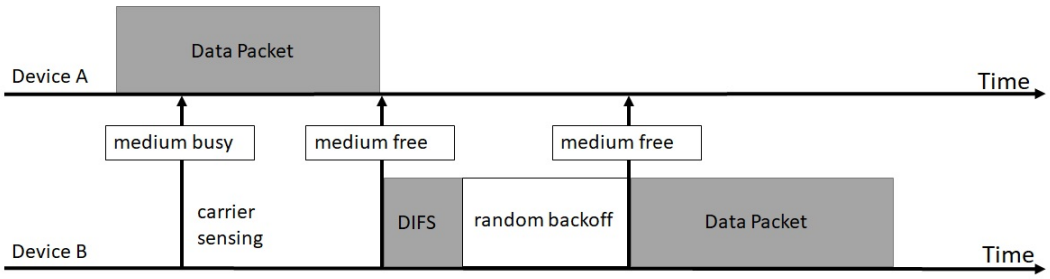

Figure 2.6. The DCF mechanism of WiFi.

required to sense the channel before attempting to transmit. If the medium is sensed to be busy, the device would wait for the channel to become idle. On the contrary, if the channel is sensed to be free for at least DCF interframe space (DIFS) (34 μs), the device will backoff for a random number of slots between [0, CW], where CW is the contention window size, and then start transmitting if the medium was still free [62]. The CW of a STA starts from $\mathrm{CW}_{\mathrm{min}}$, is reset to $\mathrm{CW}_{\mathrm{min}}$ when a packet is successfully transmitted, and doubles on each unsuccessful MAC protocol data unit transmission reaching a maximum limit of $\mathrm{CW}_{\mathrm{max}}$. When the CW is increased to $\mathrm{CW}_{\mathrm{max}}$, it remains at $\mathrm{CW}_{\mathrm{max}}$ even if there was still more collisions. If the channel becomes busy during a backoff process, the backoff is suspended. When the channel becomes idle again, and stays idle for an extra DIFS time interval, the backoff process resumes with the suspended backoff counter value. For each successful reception of a frame, the receiver acknowledges by sending an acknowledgement (ACK) frame. The ACK frame is transmitted after a short interframe space (SIFS), which is shorter than the DIFS. Other STAs resume the backoff process after the DIFS idle time. If an ACK frame is not received after the data transmission, the frame is retransmitted after another random backoff process. All of the MAC parameters including SIFS, DIFS, Slot Time, $\mathrm{CW}_{\mathrm{min}}$, and $\mathrm{CW}_{\mathrm{max}}$ are dependent on the underlying PHY layer. A summary of the WiFi DCF mechanism is provided in Fig. 2.6.

Meanwhile, LTE is based on orthogonal frequency division multiple access (OFDMA) for DL transmission and single-carrier frequency-division multiple access (SC-FDMA) in the UL in order to benefit from a lower peak-to-average power ratio. SC-FDMA results in a better transmit power efficiency and hence a reduced cost for power amplification. LTE can operate in TDD and frequency division duplex (FDD) modes in which the UL and the DL traffic could be transmitted on either the same or different frequencies, respectively. The LTE bandwidth is divided in time and frequency units called resource blocks (RBs) [63,64]. As such,

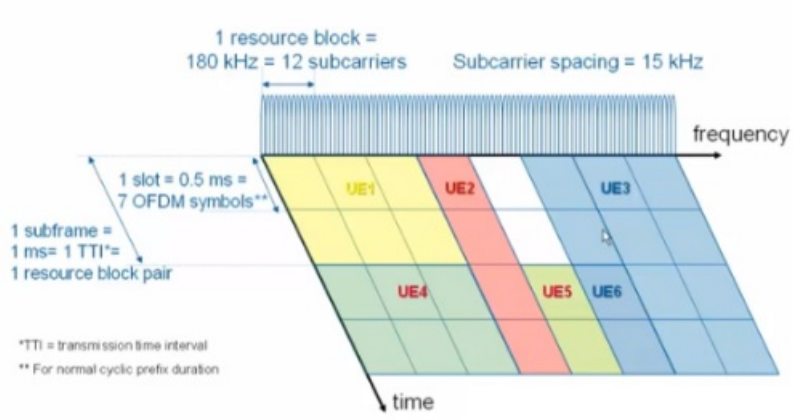

Figure 2.7. The allocation of the bandwidth in LTE OFDMA [7].

a UE may be allocated at least one RB and hence many UEs can transmit at the same time on different portions of the channel during one LTE OFDM symbol (71.4 μs). A summary of the LTE bandwidth division in OFDMA is provided in Fig. 2.7.

Therefore, when WiFi and LTE coexist over the same band, as shown in Fig. 2.8, WAPs will sense the channel as busy and would therefore defer from transmission. Consequently, in the absence of new LTE-WiFi coexistence mechanisms, the aforementioned challenges for WiFi arise thus resulting in an unfair spectrum sharing. Next, we overview some of the main techniques for the enabling the coexistence of LTE-LAA and WiFi in the 5 GHz band.

2.1.2.6 Enabling Techniques for the Coexistence of LTE-LAA and WiFi

In this section, we summarize various LTE-WiFi coexistence mechanisms such as LBT, almost blank subframe (ABS), transmission power control, and channel selection.

2.1.2.6.1 Listen-before-talk LTE-LAA SBSs can essentially adopt an LBT mechanism similar to that of WiFi. By using LBT, it will no longer be possible for LTE-LAA SBSs to transmit immediately since the intended wireless shared channel may be occupied by other SBSs or WAPs. 3GPP has standardized LBT as an effective solution for the coexistence of LTE and WiFi [59]. In particular, 3GPP is considering different requirements for LBT to access spectrum bands in different countries and define parameters to ensure better coexistence of RATs, including WiFi and other LAA-based MAC mechanisms, in the shared unlicensed band [65].

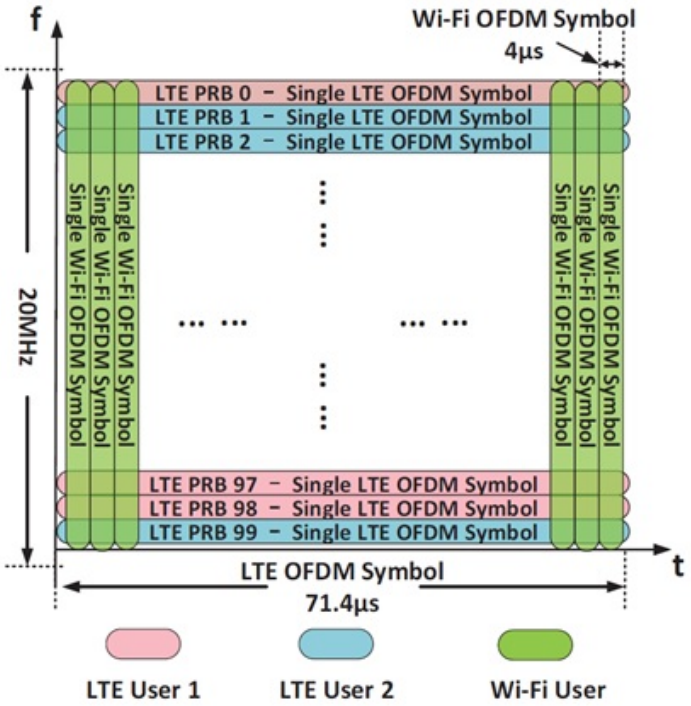

Figure 2.8. PHY and MAC comparisons between LTE and WiFi systems in both time and frequency domains [8].

2.1.2.6.2 Almost Blank Subframe An intuitive way to share the spectrum is to prevent different technologies from accessing a given channel at the same time. Therefore, LTE systems can mute all their operations in certain subframes termed almost blank subframes [65]. These subframes are called "almost blank" because LTE can still transmit some broadcast signals, control signals, and synchronization signals over these subframes. This will in turn yield a reduction in the co-channel interference during the ABS periods.

2.1.2.6.3 Transmission Power Control Another viable solution for the efficient coexistence of LTE and WiFi in the unlicensed band is SBS transmit power reduction. A controlled decrease of the SBSs' transmit powers can reduce the interference level caused to neighbouring WiFi nodes, thus creating more WiFi transmission opportunities as WiFi nodes detect the channel as vacant.

2.1.2.6.4 Channel Selection The uncoordinated nature of WiFi deployments and the limitation of non-overlapping channels in the ISM bands have motivated several studies about channel selection for WiFi networks, which could also be exploited for the coexistence with LTE. For instance, an LTE-LAA SBS can select the least congested unlicensed channel based on WiFi and LTE measurements.

Moreover, instead of having fixed BW channels, adaptive BW channels could be defined and selected in coexistence scenarios. Therefore, by adapting the transmission channel and the corresponding channel BW, LTE-LAA would essentially operate over the less congested channels thus reducing the disruption caused to WiFi.

We note that an LTE-WiFi coexistence approach is not restricted to only one of above enabling techniques. In essence, multiple coexistence schemes can be combined into a single framework in order to utilize the unlicensed spectrum in a more efficient way. In what follows, we elaborate more on alternative LTE solutions in the unlicensed spectrum, such as LTE-Unlicensed and LTE-WiFi link aggregation (LWA).

2.1.2.7 LTE-LAA v/s LTE-Unlicensed v/s LTE-WiFi link aggregation

In this section, we give a general overview on different LTE operation schemes over the unlicensed spectrum, such as LTE-U and LWA.

2.1.2.7.1 LTE-Unlicensed LTE-U is a proposal, originally developed by Qualcomm, for the operation of LTE over the unlicensed spectrum [66]. As opposed to LTE-LAA, LTE-U does not incorporate an LBT mechanism for its coexistence with WiFi and thus does not meet the regulatory requirements for using the unlicensed spectrum in various parts of the world. As such, LTE-U is suitable for countries such as the United States and China in which LBT is not mandatory. To provide fair coexistence with incumbent devices, LTE-U relies on channel selection and carrier-sensing adaptive transmission [67].

2.1.2.7.2 LTE-WiFi link aggregation LWA, also know as LTE-H combines LTE and WiFi links, for existing and new carrier grade WiFi [68]. Unlike LTE-LAA which operates over the 5 GHz band only, LWA transmits on both the 2.4 GHz and the 5 GHz unlicensed bands. In this scenario, WAPs will be connected to the LTE network, like any other SBS, and can fully utilize LTE's core network, encryption, control, authentication, and other systems. LWA allows a user to receive data through an LTE link and a WiFi link simultaneously and thus resulting in a significant increase in the data rates by combining two networks' best achievable rates. LTE-WiFi link aggregation can essentially occur between either collocated or separate (but coordinated) WiFi and LTE nodes. Here, note

that the decision on whether to use both LTE and WiFi or just LTE is the network's decision, not a user's. The traffic that flows over WiFi is collected at the WAP and then tunneled back to the LTE SBS. It is the role of the LTE SBS to control the amount of traffic scheduled over WiFi, and thereby ensure proper load balancing between the LTE and the WiFi links. In fact, LTE-LAA and LWA will be both adopted for future wireless cellular networks – LTE-LAA for new small cell deployments operating over the 5 GHz band, and LWA for existing and new WiFi deployments using both the 2.4 GHz and the 5 GHz unlicensed bands [68].

In this subsection, we provided a general background on the spectrum for future mobile networks. Next, we overview two emerging machine-types devices in the Internet of Things, namely, unmanned aerial vehicles and autonomous connected vehicles.

2.1.3 Internet of Things

Future wireless cellular networks will encompass trillions of machine-type devices that will be connected to the Internet [69]. This IoT environment will enable these devices to connect with each other over wireless links and thus operate in a self-organizing manner. Therefore, IoT devices can collect and exchange real-time information to provide smart services. In this respect, the IoT will allow the delivery of innovative services and solutions in the realms of smart cities, smart grids, smart homes, and connected vehicles that can provide a significant improvement in people's lives. For example, IoT devices can be used to intelligently manage all the city's systems such as local departments' information, schools, libraries, transportation, hospitals, water supply, and electricity systems hence improving service efficiency. Meanwhile, retailers, restaurant chains and makers of consumer goods can use data from smartphones, wearable devices and in-home devices to do targeted marketing and promotions. However, the practical deployment of the IoT system sill faces many challenges such as data analytics, computation, transmission capabilities, connectivity end-to-end latency, security [70], and privacy [71]. In particular, how to provide massive device connectivity with stringent latency requirement will be one of the most important challenges. The current centralized communication models and corresponding technologies may not be able to provide such massive connectivity. Moreover, for each IoT device, energy and computational resources are limited. Hence, how to allocate computational

resources and power for all IoT devices to achieve the data rate and latency requirements is another challenge. In this section, we overview two machine-types devices, UAVs and connected autonomous vehicles, that will be an integral part of IoT in 5G networks.

2.1.3.1 Unmanned aerial vehicles

Due to their rapid and flexible deployment capabilities, mobility, ability to fly above obstacles, and relatively low cost, aerial platforms such as UAVs have received considerable interest for different applications in wireless communications. In addition, compared to ground base stations, UAVs can establish line-of-sight communication links with ground UEs by intelligently adjusting their altitude. These evident benefits make UAV-aided wireless communication a promising integral component of future wireless systems. For instance, in 2013, Amazon announced a research and development initiative focused on its next-generation Prime Air delivery service [72]. The goal of this service is to deliver packages to customers within 30 minutes using UAVs. In 2014, the Phantom and Inspire from DJI, the Loon Project from Google [73], AR Drone and Bebop Drone from Parrot, and IRIS Drone from 3D Robotic have been launched [74].

In particular, UAVs can be broadly classified into two categories, fixed wing and rotary wing, each of which having its own benefits and drawbacks [75]. For example, fixed-wing UAVs are mainly characterized by their high speed and heavy payload. However, such UAVs must maintain continuous forward motion in order to remain aloft thus rendering them unsuitable for stationary applications like close inspection. In contrast, rotary-wing UAVs such as quadcopters, while having limited mobility and payload, are able to move in any direction as well as to stay stationary in the air. Therefore, the choice of UAVs largely depends on the underlying applications. Meanwhile, in the context of wireless communication systems, UAVs mainly have three roles: UAVs as aerial base stations, UAVs as aerial relays, and UAVs as cellular-connected UAV-UEs. In what follows, we summarize these three typical use cases of UAV-aided wireless communications:

- **UAVs as aerial BSs:** UAVs can be used as flying BSs for coverage extension, capacity enhancement, mission critical services, and other scenarios in which no terrestrial infrastructure exists (e.g., in public safety scenarios or in rural areas) or in an event of damaged/overloaded terrestrial BSs. Two

example scenarios are rapid service recovery after partial or complete infrastructure damage due to natural disasters, and offloading of ground networks in extremely crowded areas (e.g., a stadium during a sports event). UAVs can also be used as part of a fully-fledged flying wireless cellular network that can dynamically change its architecture to service the ground UEs.

- **UAV as aerial, cellular-connected UEs:** 3GPP is considering the support of UAVs by LTE [76, 77]. Providing cellular connectivity to the UAVs offers several advantages to other current short-range wireless connections such as WiFi, bluetooth, and radio waves. For instance, cellular technology will enable beyond LoS control, low wireless latency, ubiquitous coverage, and seamless mobility to various UAV operations. This in turn paves the way to a widescale deployment of UAVs over 5G systems, especially for mission-critical use cases. It also enables many IoT applications in agriculture, military, mining operations and industrial inspection services such as real time video streaming, delivery, surveillance, and transmitting telematics and flight information. Nevertheless, providing wireless cellular connectivity for UAV-UEs is contingent upon proper management of their resulting interference. In essence, most UAV communication links are LoS dominated which can therefore result in a high interference level on the ground UEs. As such, network operators will have to limit the admission of aerial vehicles so that the perceived throughput performance of conventional UEs is not deteriorated. Another major challenge for cellular-connected UAVs is the need for efficient handover mechanisms and robust signaling.

- **UAVs as aerial relays:** UAVs can also act as relays between a source and a destination in which a LoS link does not exist. For instance, this could be between the frontline and the command center for emergency responses. UAVs can also form a multihop aerial network for coverage extension or backhaul connectivity to ground SBSs. They can engage in a cooperative transmission scheme thus forming an ad-hoc network in the air. Moreover, a UAV can offer a *mobile relaying* strategy, as opposed to a *static relaying*, which is more suited for delay-tolerant applications [75]. With mobile relaying, the UAV flies continuously between the source and destination aiming to reduce the link distances during both UAV information reception and relaying phases. An alternative strategy to mobile relaying is known as

data ferrying or *load-carry-and-delivery* [75]. With this strategy, the UAV "loads" the data from the source as it reaches the nearest possible location from the source, flies toward the destination with the loaded data until it reaches the nearest possible location to the destination, and then delivers the data to the destination. As data ferrying has less communication time than the proposed mobile relaying, its achievable throughput is expected to be smaller, especially for cases with low UAV speed and/or stringent delay requirements.

Despite the many promising benefits for UAV-enabled wireless networks, several challenges need to be addressed. In what follows, we summarize the main challenges that arise in the context of UAV-based wireless communication systems:

- **3D placement:** UAVs need to be deployed in a 3D system which therefore makes their deployment more complicated as compared to that of the terrestrial BSs. In this regard, the altitude of the UAVs plays a vital role in establishing a LoS communication link with the terrestrial network and should therefore be optimized accordingly. Moreover, the deployment of the UAVs need to be flexible, short-term, and fast.

- **Antenna design:** Current network deployments assume that communication occurs within a 2-dimensional plane, and must support human devices (with low mobility) and ground vehicles (with potentially high mobility). Nevertheless, aerial networks consist of A2G and A2A links and require data to be delivered at different altitudes and orientation angle. As such, directed antenna radiation characteristics are more likely to have high impact on the performance of 3D connectivity.

- **Path planning:** Appropriate path planning schemes are crucial for UAVs in order to guarantee high-capacity performance, particularly, for UAV-UE applications. One useful method for UAV path planning is to approximate the UAV's dynamics by a discrete-time state space, with the state vector typically consisting of the position and velocity in a 3D coordinate system. The UAV's trajectory is then given by the sequence of states which are subject to finite transition constraints that reflect the practical UAV's mobility limitations.

- **Channel modeling:** UAV-based communication is different from terrestrial communication in various aspect which essentially necessitates a statistical RF propagation model for the UAV communication channel [78]. In essence, such a model can mainly vary based on the different communication environments.

- **Backhaul connectivity:** A major limitation for the use of UAVs as aerial BSs is the availability of a reliable wireless backhaul link. As compared to a ground BS that usually has a fixed wired/wireless backhaul connection and can relatively offer very high data rates to a core network, a UAV can only have a wireless backhaul. This in turn limits the achievable peak data rate of a UAV which can also be susceptible to inclement weather conditions.

- **Design of a multi-hop aerial network:** As mentioned earlier, UAVs can form an multihop aerial network for coverage extension or backhaul connectivity for the ground SBSs. However, a major challenge in that scope is the formation of such an aerial network. In particular, the UAVs need to form the A2A and A2G link while taking into account the delay incurred over the formed multihop links.

- **Interference management:** UAV-based communication links are LoS dominated which therefore make the issue of intercell interference more critical than terrestrial communications. As such, novel methods for interference management, and in particular, for the coexistence of the aerial UAVs with the ground network, are indispensable.

- **Power limitation:** UAVs are typically limited in power which therefore result in a limitation for their flight duration. This in turn necessitates energy-aware UAV deployment and operation mechanisms for intelligent energy usage and replenishment.

- **Security issues:** Drone delivery systems are vulnerable to several cyber and physical attacks [79]. On the physical side, UAVs are susceptible to a range of civilian owned hunting rifles due to their low altitude. Moreover, the UAVs are vulnerable to a range of cyber threats targeting their communication links with ground control and with other air units.

- **Spectrum allocation:** Due to the fact that the UAVs' channel experience

less blockage and high probability for establishing a LoS link, the use of mmWave spectrum bands could be a potential candidate for UAV-based communication. Nevertheless, the Doppler spread due to the mobility of the UAVs needs to be compensated in such scenarios.

2.1.3.2 Connected autonomous vehicles

5G access will provide a whole load of horsepower for the auto industry to build advanced solutions to make vehicles smarter, safer, and more energy-efficient. The innovations will not only improve the vehicles' internal functions, but will also allow the vehicles to connect and interact with the outside environment thus enabling better and safer use of the road infrastructure. In fact, cellular-connected vehicles are expected to be the next frontier for automotive revolution and the key to the evolution to next-generation intelligent transportation systems. Cellular-connected vehicles are considered as one of the main emerging wireless-enabled devices in 5G networks and will constitute a main part of future smart city services, e.g., the Waymo google project [80]. Connected autonomous vehicles are expected to anticipate and avoid possible collisions, navigate the quickest route to their destination making use of up-to-the-minute traffic reports, identify the nearest available parking slot, and minimize their carbon emissions. To realize this, connected vehicles and the transportation infrastructure would be equipped with smart sensors that can collect and process a heterogeneous set of data on each vehicle, its passengers, and its environment [81]. Communication can be between different nodes i.e., vehicle-to-infrastructure, vehicle-to-vehicle, or vehicle-to-pedestrian, all together known as vehicle-to-everything. The collection of information from other vehicles and/or the infrastructure must be done at ultra-low latency and in real time in order to support the autonomous feature of connected vehicles. In this context, several challenges need to be addressed in order to reap the benefits of connected autonomous vehicles. In what follows, we summarize some of these challenges.

- **Latency:** Latency is defined as the time interval between the generation of the data packet by the transmitter and the time this data packet is delivered to the recipient. For connected autonomous vehicles, ultra-low latency is required in order to ensure reliable real time reception of the information. To improve reliability and latency requirements, optimal processing

methods must be deployed to combine the heterogeneous data collected from different types of sensors. This in turn would reduce the unnecessary and redundant information before transmission to other vehicles or to the infrastructure.

- **Path planning and autonomous control:** Connected autonomous vehicles would need to find their optimal routing in an effort to reduce the total road travel delay. Nevertheless, the high density of the vehicles in crowded streets and the random decisions taken by human drivers make vehicular path planning and autonomous control algorithms a challenging aspect in such systems.

- **Security:** Connected autonomous vehicles are susceptible to cyber attacks in which an attacker can inject faulty information to autonomous vehicles. This in turn necessitates new techniques for authenticating and securing the vehicle-to-everything communication links.

- **Heterogeneous data integration:** Connected autonomous vehicles will encompass a variety of sensors, from light sensors and cameras to ultrasound sensors, enabling each vehicle to make sense of its surrounding environment. As such, any control action taken by an autonomous vehicle will depend on the different types of sensor data. Integrating such heterogeneous sensor readings into one vector can provide a better assessment of the environment as compared to using each type of data independently. Nevertheless, there exist differences between sensors ranging from sampling rates to the data generation model thus making the integration of data sensors a challenging aspect in this context [81].

- **Resilience:** Autonomous vehicles must be resilient to vehicle accidents, congestion, and natural disasters. For instance, autonomous vehicles must be resilient to possible collisions so as to recover from congestion caused by accidents. Furthermore, they must be capable of adapting to extreme situations such as floods, hurricanes, and other disasters.

Therefore, with the advent of the Internet of Things and the dramatic increase in the mobile data traffic, an ultra-dense network deployment is required for 5G networks, as discussed earlier. Nevertheless, a critical issue in such a dense

deployment is the connectivity of the SBSs to the core network. As such, next we give general background on SBS backhauling in 5G networks, a topic of direct relevance to this thesis.

2.1.4 5G Backhaul

As we have mentioned earlier, SBSs have been introduced as a promising approach to boost the coverage and capacity of current cellular networks. This is mainly accomplished through a paradigm shift in the network design by considering a reduction in the cell size and a denser deployment of the SBSs. However, a fundamental challenge for the efficient operation of such a dense SBS deployment is to provide an economical and ubiquitous backhaul connectivity to the SBSs that would allow the routing of the traffic to/from the SBSs from/to the core network.

To achieve gigabits-per-second transmission capacity and maximum allowed latency in the orders of hundreds of microseconds, fibre optics would be the only backhaul viable solution. Nevertheless, relying on fibre to connect all the SBSs to the core network may be impossible in some cases due to geographical constraints, for instance, and certainly very costly otherwise. To this end, the 5G backhaul research has received much attention, aiming at bridging the gap between the 5G requirements and the realistic backhaul capabilities [38]. The 5G backhaul evolution will essentially include wired and wireless backhauling to and from core network aggregators such as MBSs, cooperation through anchor base stations, multi-hopping at short range links, and cloud-based architecture. Since the backhaul requirements can significantly vary depending on the locations of the SBSs, the cost of implementing backhaul connections, traffic load intensity of the SBSs, latency, and target quality of service requirement of SBS UEs, there is no single optimal approach for the backhauling of SBSs. Therefore, in this section, we summarize the different types of 5G backhauling schemes. In particular, we highlight the benefits and drawbacks of wired and wireless backhauling solutions and elaborate more on novel backhauling solutions such as caching and SDN.

2.1.4.1 Optical Fibre

Current backhaul networks are mostly built with microwave and fibre links with different proportions per operator and country [82]. In fact, fibre-optic-based

backhaul is a leading attractive 5G solution owing to its superior performance relative to other technologies. It is an ideal solution for connecting the backhaul in view of their intrinsic low-latency-high-capacity characteristics. Nevertheless, fiber backhaul network does not exist always and its deployment cost is considered to be relatively high. Moreover, highly reliable wired backhaul connectivity may not be necessary for small cells, which typically serve a relatively reduced traffic load compared to a macrocell. As such, an alternative solution, that is easy and cost-effective to deploy, is crucial in order to bridge the existing fibre-based network and the small cells in an ultra-dense network.

2.1.4.2 Wireless backhauling

Wireless backhauling is considered as a viable and cost-effective approach that allows operators to obtain end-to-end control of their network rather than leasing third party wired backhaul connections [83]. The key wireless backhaul solutions exploit the mmWave spectrum in the 60 GHz and 70-80 GHz bands, microwave spectrum between 6 GHz and 60 GHz bands, sub-6 GHz band, TV white spaces (TVWS), and satellite technologies [83]. An optimal solution of the wireless backhaul solution depends on the propagation environment and a number of system parameters such as locations and deployment density of the SBSs, desired backhaul capacity, interference conditions, cost, coverage, hardware requirements, and spectrum availability. In what follows, we summarize the benefits and limitations of the main wireless backhaul solutions [83].

- **Sub-6 GHz spectrum:** Sub-6 GHz frequencies support NLoS propagation and have better penetration through obstacles as opposed to high frequency bands. As such, sub-6 GHz frequencies can support point-to-multipoint backhaul connectivity, however, at the cost of interference. Moreover, the sub-6 GHz spectrum does not require any new hardware for managing the access and backhaul links. Nonetheless, relying on the sub-6 GHz frequencies for providing wireless backhaul solution is highly vulnerable to interference and traffic congestion, and has a high licensing cost.

- **mmWave:** The propagation characteristics of the mmWave spectrum are attractive for high-capacity short-range communication links. Moreover, as we mentioned earlier, the mmWave band is characterized by the availability of a large bandwidth as well as the ability to minimize interference with

highly directive narrow beamwidth antennas. Nevertheless, mmWaves are affected by atmospheric attenuation to a greater degree than lower frequencies. The power attenuation at 60 GHz is basically due to the oxygen or dry air, whereas 70-80 GHz is more similar to conventional microwave, where attenuation is mainly caused by water molecules in the air. As a result, the 60 GHz band is more heavily attenuated. However, the license-exempt nature of the 60 GHz band makes it more cost effective from the operators' perspective.

- **Microwave:** The microwave frequency band ranges between 6 GHz and 60 GHz. It is characterized by shorter wavelength and hence is suitable for LoS scenarios with fixed antenna alignments on both the transmitter and receiver ends. Moreover, the microwave bands are favorable for short-range communications, such as neighbourhood backhauling in ultra-dense small cell deployment scenarios, due to their high signal attenuation.

- **TVWS:** The TV band is divided into two bands: VHF band (54-60 MHz, 76-88 MHz, 174-216 MHz) and UHF band (470-698 MHz). TVWS refers to the large amount of TV spectrum that became vacant upon the emergence of digital TV transmission. Therefore, TVWS can be exploited for SBS backhauling in a cognitive manner such that the backhaul interference caused to primary TV transmissions does not exceed a prescribed threshold. The TVWS spectrum is characterized by their longer wavelengths and unlicensed nature which results in a lower cost. Moreover, the channels in the TVWS offer much better propagation characteristics compared to low-frequency cellular bands. Nevertheless, this spectrum band would be strictly limited by the transmit power and location of primary TV transmitters when used for SBS backhauling.

- **Satellite frequency:** The degree of attenuation due to weather or rain fade is different for various satellite frequency bands. For instance, lower frequency bands (4-6 GHz) are unaffected by weather, the Ku-band (10-12 GHz) is slightly more affected, and the Ka-band (20-30 GHz) suffer from up to 24 dB of rain fade. The main advantage of satellite-based backhauling is that it is feasible at any location where a convenient satellite is visible. It is also efficient in high mobility scenarios, such as SBSs deployed on airplanes and ships, thus providing continuous backhaul coverage to mobile SBSs.

Nevertheless, the limitation of satellite backhaul solution is time delay and hardware and spectrum cost.

2.1.4.3 Caching

A promising solution for addressing the backhaul bottleneck is to cache the content at the edge of network (SBS or UE) thus transforming the network intelligence from being "reactive" to "proactive" [38]. In other words, if the user's data was predicted and cached in advance during low traffic periods, it can be transmitted during peak hours without burdening the backhaul while still achieving good QoS. Nonetheless, the success of caching remains conditional upon many challenges, such as the SBSs storage capacity, very large catalogue size of users' files, and the need for fast and dynamic learning of cells while making the caching decision.

2.1.4.4 SDN in the backhaul

The network topology for SDN consists of a single controller that manages access to dumb devices or routers in the carrier IP-based mobile backhaul. On the other hand, all control plane functions or intelligence is embedded in the SDN controller. The communication between SDN controller and devices happens through the use of the Open Flow protocol. SDN essentially decouples control from the data forwarding function, in a programmable manner, thus, creating a dynamic, manageable, cost-effective, and adaptable architecture that gives administrators unprecedented automation and control. In fact, an SDN-based backhaul solution expedites the possibilities of adding, extending, and dynamically reallocating the radio resources in the backhaul network. Such an architecture allows multi-operators and multi-technology sharing thus reducing the cost per bit to the UE and maximising the resource usage efficiency. Nonetheless, the separation of control and data forwarding exposes the network to security challenges, especially when used with cloud computing, due to malicious usage or malfunctioning in the system.

2.1.5 Fairness In Wireless Networks

Alongside various network design performance such as maximizing data rate and minimizing wireless latency, the notion of fairness among different wireless devices

is an important design factor that should be accounted for while optimizing the performance of a network. In wireless network, fairness can be analysed from several dimensions such as energy usage, achieving required quality of services, spectrum sharing. Fairness can be considered at the system and device levels. The system fairness addresses the overall fairness amongst all devices in the network, and individual fairness indicates whether a certain device is treated fairly by the network. Moreover, considering the time duration, fairness can be categorized into short-term and long-term. Short-term fairness focuses on resource allocation in a very short time period and thus has a significant impact on QoS, especially in real-time applications because of the focus on the current QoS measurements. On the other hand, long-term fairness measures the resource allocation over a longer time period and is more important when the resources are scarce. Several fairness tools have been introduced for measuring the level of fairness in a wireless network, as given below [84]:

- *Jain's index*: Jain's index is one of the earliest proposed and widely studied fairness measures. Fairness in an allocation can be represented by an index value $f(X)$ and is computed as:

$$f(X) = \frac{(\sum_{i=1}^{n} x_i)^2}{n \sum_{i=1}^{n} x_i^2}, \tag{2.1}$$

  where n corresponds to the number of nodes, x is the metric used for defining fairness, and $0 \leq f(X) \leq 1$. A large value of $f(X)$ represents a fairer resource allocation from the system perspective i.e., a resource allocation scheme tends to be fairer when Jain's index is closer to 1 [84]. Here, note that Jain's index of fairness is mainly used for providing insights into the overall system fairness.

- *Entropy*: Entropy was introduced by Shannon [85]. It assumes that the proportions of resource are allocated to n individuals P=(p_1, p_2, ..., p_n), and

$$p_i = \frac{x_i}{\sum_{i=1}^{n} x_i}, \tag{2.2}$$

  where $0 \leq p_i \leq 1$ and $\sum_{i=1}^{n} p_i = 1$. The uncertainty of the distribution P is called the entropy of the distribution P and is usually measured by $H(P)$:

$$H(P) = -\sum_{i=1}^{n} (p_i \log_2 p_i), \tag{2.3}$$

- *Other measures:* Difference or ratio between the highest and lowest values of particular performance parameter could be a system fairness measure.

Even though $H(P)$ may be employed as a fairness measure, the quality of measuring the fairness is not clear yet. For example, how sensitive they are to the allocation changes and whether they can locate unfairness [84]. It is also worthwhile noting that Jain's index of fairness has received much traction in the literature for assessing the fairness level for resource allocation schemes in wireless networks [84, 86]. As such, given the fact that the main objective of this thesis is to investigate resource allocation schemes for LTE-WiFi coexistence and UAV-based wireless communication, we adopt the notion of Jain's index for evaluating the level of fairness for the outcome of the proposed schemes.

In the first part of this chapter, we gave a general background on specific topics in next-generation cellular networks that are of direct relevance to this thesis. In the remaining part of this chapter, we elaborate more on the underlying analytical techniques adopted as part of the proposed solutions in this thesis. We mainly focus on machine learning techniques, and in particular, neural networks, followed by game theory tools, and specifically, network formation games.

2.2 Machine Learning

In this section, we first provide a brief overview on the basics of machine learning, while motivating the importance of neural networks. We expose the fundamentals of a suite of neural network algorithms and techniques. Then, we elaborate more on two types of artificial neural networks (ANNs); recurrent neural networks (RNNs) and deep neural networks (DNNs). Table 2.1 provides a summary for the description of the main notations used in this section.

2.2.1 Machine Learning Basics

First coined in 1956 by John McCarthy, AI involves machines that can perform tasks that are characteristic of human intelligence. It can be categorized into two groups; general and narrow. General AI would have all of the characteristics of human intelligence. Narrow AI exhibits some facet(s) of human intelligence, and can do that facet extremely well, but is lacking in other areas. For instance, a machine that's great at recognizing images, but nothing else, would be an example

Table 2.1 Variables and notations.

Notation	Description
$\boldsymbol{w}_j$	weight of neuron j
$\boldsymbol{o}_j$	Output signal of neuron j
$\boldsymbol{b}_j$	Bias of neuron j
$\boldsymbol{x}$	Input vector
$\boldsymbol{W}$	Weight matrix
$\boldsymbol{W_{in}}$	Input weight matrix
$\boldsymbol{W_{out}}$	Output weight matrix
γ	Learning rate
α	Leaky parameter
ρ	Spectral radius

of narrow AI. At its core, *machine learning is simply a way of achieving AI.*

Recently, due to growing volumes of generated data – across critical infrastructure, communication networks, and smart cities – and the need for intelligent data analytics, the use of machine learning algorithms has become ubiquitous [87] across many sectors such as financial services, government, health care, technology, marketing, and entertainment. Using machine learning algorithms to build models that uncover connections and predict dynamic system or human behavior, system operators can make intelligent decisions without human intervention. For example, in a wireless system such as the IoT, machine learning tools can be used for big data analytics and edge intelligence. Machine learning tasks often depend on the nature of their training data. In machine learning, *training* is the process that teaches the machine learning framework to achieve a specific goal, such as for speech recognition. In other words, training enables the machine learning framework to discover potentially relationships between the input data and output data of this machine learning framework. There exists, in general, four key classes of learning approaches [88]: a) supervised learning, b) unsupervised learning, c) semi-supervised learning, and d) reinforcement learning.

Supervised learning algorithms are trained using labeled data. When dealing with labeled data, both the input data and its desired output data are known to the system. In contrast, training of unsupervised learning tasks is done without labeled data. The goal of unsupervised learning is to explore the data and infer some structure directly from the unlabeled data. Semi-supervised learning is used for the same applications as supervised learning but it uses both labeled and unlabeled data for training. This type of learning can be used with methods such

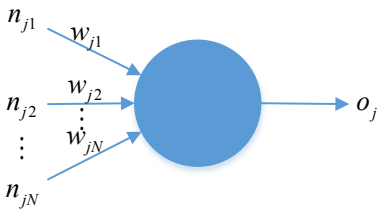

Figure 2.9. A mathematical neuron model.

as classification, regression and prediction. Semi-supervised learning is useful when the cost of a fully labeled training process is relatively high. In contrast to the previously discussed learning methods that need to be trained with historical data, RL is trained by the data from implementation. The goal of RL is to learn an environment and find the best strategies for a given agent, in different environments. RL algorithms are used for robotics, gaming, and navigation [89]. To perform these learning tasks, several frameworks have been developed. Among those frameworks, artificial neural networks [90] constitute one of the most important pillars of machine learning, as they are able to mimic human intelligence, to model complex relationships between inputs and outputs, to find patterns in data, or to extract the statistical structure in an unknown joint probability distribution from the observed data. ANNs can also be used in a self-organizing manner to learn how to perform tasks based on the data given for training or initial experience. Next, we introduce the basic concepts and general architecture of ANNs.

2.2.1.1 General Architecture of Artificial Neural Networks

The architecture of ANNs consists of a number of simple, highly interconnected processing elements known as *neurons*, which are used to mimic how the human brain learns. ANNs are essentially an artificial model of a human nervous system whose base elements are also neurons used to process information in the sense of cognition and transmit this information signal in the nervous system [91]. In essence, ANNs use artificial neurons to replicate the operation of the human nervous system, thus enabling artificial intelligence. Mathematically, an artificial neuron consists of the following components: (a) a number of incoming connections; (b) a number of outcoming connections; and (c) an activation value assigned to each neuron. The connection strength between two neurons is mainly captured by a weight value. The basic model for a neuron j is shown in Fig. 2.9

and mathematically given by:

$$o_j\left(\boldsymbol{w}_j, b_j, \boldsymbol{n}_j\right) = f\left(b_j + \sum_{j=k}^{N} n_{jk} \cdot w_{jk}\right), \tag{2.4}$$

where n_{jk} is the input signal from neuron j to neuron i, $\boldsymbol{n}_j = [n_{j1}, n_{j2}, \ldots, n_{jN}]$ is a vector of the input signal of neuron j, w_{jk} is the corresponding input weight value, $\boldsymbol{w}_j = [w_{j1}, w_{j2}, \ldots, w_{jN}]$ is a vector of input weight of neuron j, o_j is the output signal of neuron j, b_j is the bias of neuron j, and $f(\cdot)$ is a nonlinear activation function. A bias value can shift the activation function, which is critical for successful learning. The activation function in a neural network will represent the rate of action potential firing in the cell of a neuron. The simplest form of an activation function [92] is binary such as a Heaviside step function, which is used to indicate whether a neuron is firing or not. However, using linear activation functions, many neurons must be used in the computation beyond linear separation. Meanwhile, an ANN constructed using linear activation functions in (2.4) cannot reach a stable state after training since the value of the activation function will increase without bound. To avoid this drawback, one can choose, $f(\cdot)$ in (2.4) to be a normalizable activation function such as a sigmoid activation function rather than a linear activation function. The selection of a type of activation functions in ANNs depends on the sought objectives such as analytic tractability, computational power, and the type of the desired output signal (logistic or continuous). In essence, an ANN is a composition of multiple neurons connected in different ways and operating using different activation functions. In general, the main components of an ANN that consists of multiple neurons will include the following:

- *Input layer* that consists of a number of neurons used to represent the input signal.
- *Output layer* that consists of a number of neurons used to represent the output signal.
- *Hidden layer* that consists of a number of neurons used to mimic the human brain.
- *Input weight matrix* that represents the strengths of the connections between the neurons in the input layer and the neurons in the hidden layer.

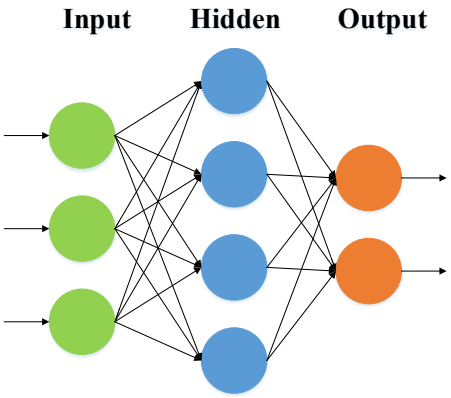

Figure 2.10. Feed-forward neural network.

- *Neuron weight matrix* that represents the strengths of the connections between the neurons in the hidden layer.
- *Output weight matrix* that represents the strengths of the connections between the neurons in the hidden layer and the neurons in the output layer.

The connection strength in all weight matrices can be used to calculate the value of the activation function as per (2.4). For example, if $\boldsymbol{n}_j = [n_{j1}, n_{j2}, \ldots, n_{jN}]$ in (2.4) is an input signal vector, then $\boldsymbol{w}_j = [w_{j1}, w_{j2}, \ldots, w_{jN}]$ will represent the value of the input weight, and, thus, the value of the activation function can be calculated by (2.4). The hidden layer is used to analyze the relationship between the input signal in the input layer and the output signal in the output layer.

One of the simplest forms of an artificial neural network is the *feed-forward neural network (FNN)* [93], as shown in Fig. 2.10. An FNN consists of the following components: (a) input layer, (b) hidden layer(s), and (c) output layer. In an FNN, the connection between the neurons is unidirectional and there is no connection between the neurons in a layer. Each neuron in the hidden layer calculates its output using an activation function such as the function in (2.4). Moreover, each neuron in the hidden layer has incoming connections only from the previous layer and outgoing connections only to the next layer, and, hence, this architecture is named *feed-forward neural network*.

Having introduced the general architecture of an ANN, next, we discuss the training methods that ANNs can use to perform their learning tasks.

2.2.1.2 Training in Artificial Neural Networks

To learn information from their input data, ANNs must adjust the weights of the connections between the neurons in the system. The process of adjusting and

updating the weights is known as the *training process*. Different learning tasks require different training algorithms. Here, we mainly focus on the general training algorithms for supervised learning tasks, as they constitute the foundation for other types, such as unsupervised learning.

For supervised learning tasks, the objective of training ANNs is to minimize the errors between the desired output signal and actual output signal. This error can be typically defined as:

$$E\left(\boldsymbol{W},\boldsymbol{b}\right)=0.5\cdot\sum_{\mathcal{S}}\left(\left\|\boldsymbol{y}\left(\boldsymbol{W},\boldsymbol{b},\boldsymbol{x}\right)-\boldsymbol{y}_{D}\right\|_{2}\right), \tag{2.5}$$

where $\mathcal{S}$ is the training dataset, $\boldsymbol{x}$ is a vector of input signals, $\boldsymbol{W}$ is the weight matrix that is a combination of the input weight matrix, hidden weight matrix, and output weight matrix, $\boldsymbol{b}$ is a vector of bias factor, and $\boldsymbol{y}_D$ is the desired output. $\boldsymbol{y}\left(\boldsymbol{W},\boldsymbol{b},\boldsymbol{x}\right)$ is the actual output signal for each neuron, which can be calculated based on (2.4). In (2.5), the error is scaled by $\frac{1}{2}$ to facilitate differentiation. In general, the most commonly used supervised learning algorithms for ANNs include gradient descent and backpropagation [94], which is a special case of gradient descent. Hence, we overview these learning algorithms as they constitute a building block for any other learning algorithm.

In order to minimize $E\left(\boldsymbol{W},\boldsymbol{b}\right)$, the weights of each neuron are updated via the *gradient descent algorithm*, as explained next. For a given neuron j, the error between the desired output signal $o_{D,j}$ and actual output signal o_j will be given by $E_j\left(\boldsymbol{w}_j,b_j\right)=0.5\cdot\left\|o_j\left(\boldsymbol{w}_j,b_j,\boldsymbol{n}_j\right)-o_{D,j}\right\|_2$. The gradient descent algorithm is used to minimize $E_j\left(\boldsymbol{w}_j,b_j\right)$. For every element w_{jk} of vector $\boldsymbol{w}_j$, the minimization of $E_j\left(\boldsymbol{w}_j,b_j\right)$ using gradient descent algorithms follows from the following equations:

$$w_{jk,\text{new}}=w_{jk,\text{old}}-\gamma\frac{\partial E_j\left(\boldsymbol{w}_j,b_j\right)}{\partial w_{jk}}, \tag{2.6}$$

$$b_{j,\text{new}}=b_{j,\text{old}}-\gamma\frac{\partial E_j\left(\boldsymbol{w}_j,b_j\right)}{\partial b_j}, \tag{2.7}$$

where γ is the learning rate. Based on (2.6) and (2.7), ANNs can update the weight matrix and bias to find the optimal $\boldsymbol{w}_j$ and b_j that will minimize $E_j\left(\boldsymbol{w}_j,b_j\right)$. From (2.6) and (2.7), we can see that, the update of w_{jk} and b_j is easy to compute and, hence, the gradient descent algorithm is known to be computationally fast, even on large datasets [95]. The gradient descent algorithm mentioned above only focuses on the update of a single neuron. However, in an ANN, the signal is transmitted from one neuron to another neuron and, hence,

we must design a rule to train all of these neurons. Backpropagation is the most widely used algorithm to calculate the gradient of the error such as $\frac{\partial E_j(\boldsymbol{w}_j,b_j)}{\partial w_{jk}}$ in (2.6) and $\frac{\partial E_j(\boldsymbol{w}_j,b_j)}{\partial b_j}$ in (2.7), so as to effectively minimize $E(\boldsymbol{W},\boldsymbol{b})$ for an ANN. In fact, backpropagation is a method to compute the gradient of each neuron for an ANN, which is just a chain rule.

Next, we provide a step-by-step explanation of the backpropagation algorithm. We first assume that neuron j is at layer L and neuron i is at layer $L+1$, which is closer to the output layer than layer L. The backpropagation procedure that can be used to update the weight value of w_{ij} proceeds as follows:

- An input signal is transmitted from the input layer to the hidden layer of the ANN, until it reaches the output layer. If an ANN does not have hidden layers, (e.g., a perceptron), the input signal will be directly transmitted to the output layer. Then, the activations of the neurons in all of the layers will be computed based on (2.4).

- The ANN will next compute the error between the targeted output and actual output based on (2.5) and will derive an error propagation value δ for each neuron. δ_i of neuron i can be given by [96]:

$$\delta_i = \frac{\partial E(\boldsymbol{W}_i,\boldsymbol{b})}{\partial o_i} \cdot \frac{\partial o_i}{\partial n_{\text{sum},i}}, \tag{2.8}$$

  where o_i is the output of neuron i and $n_{\text{sum},i} = b_i + \sum_{k=1}^{N} n_{ik} \cdot w_{ik}$ is the summation of the input signal of neuron i and its bias. In particular, if the activation function is a logistic function, $f(x) = \frac{1}{1+e^{-x}}$, then the error propagation of neuron j can be given by [94]:

$$\delta_i = \begin{cases} (o_i - o_{D,i})\, o_i\, (1-o_i)\,, & \text{neuron } i \text{ in the output layer,} \\ \left(\sum_{l \in \mathcal{L}_{L+1}} \delta_l w_{li}\right) o_i\,(1-o_i)\,, & \text{neuron } i \text{ in the hidden layer,} \end{cases} \tag{2.9}$$

  where $\mathcal{L}_{L+1}$ represents the set of neurons at layer $L+1$ (layer $L+1$ is closer to the output layer than layer L). From (2.9), we can see that error propagation δ_i of a neuron in layer L depends on the error propagation $\delta_l, l \in \mathcal{L}_{L+1}$, of a neuron at layer $L+1$. Therefore, each neuron must transmit its error propagation parameter to the neurons at the former layer. This is the central definition of *backpropagation*.

- Next, the ANN updates the weight value of w_{ij}, which can be given by $w_{ij} = w_{ij} - \gamma \delta_i o_j$.

- The above process is repeated until all weight values reach the minimum of $E(\boldsymbol{W}, \boldsymbol{b})$. Note that backpropagation is not guaranteed to find a global minimum as it typically converges to a local optimum since the dataset used for training ANNs is finite and, hence, it must have some blindness in exploration.

In backpropagation, the gradient is computed based on the complete labeled data. However, if the size of the labeled data is very large, then using backpropagation may be time consuming. To reduce the time used for training when the size of the labeled data is very large, *a stochastic gradient descent (SGD)* algorithm [97] can be employed to update the weight values and bias. The stochastic gradient descent algorithm performs a weight value update for each training example. However, the SGD algorithm will often update frequently, which can lead to overshooting – the weight values are larger or smaller than the optimum. To overcome these drawbacks of SGD, the *mini batch gradient descent* [98] can be used. The mini batch gradient descent is an algorithm that strikes a balance between stochastic gradient descent and batch gradient descent [98]. In mini-batch gradient descent, the gradient is computed based on a small number of samples, e.g., of around 10-500. One benefit of mini-batch gradient descent is that it can be performed in a distributed manner and, hence, it can train ANNs in a time efficient manner.

In summary, gradient descent algorithms enable an ANN to be trained in a computationally simple manner, and hence, they can quickly converge to a local minimum value for the error, even on a large dataset. However, choosing a proper learning rate for the update of the weights and bias can be difficult. In fact, the learning rate determines the step size that the algorithm uses to reach the minimum and, thus, it has an impact on the convergence rate. In particular, a learning rate that is too large can cause the algorithm to diverge from the optimal solution. This is due to the fact that choosing very large initial learning rates will decay the loss function faster thus not allowing the model to explore better the optimization space. On the other hand, a learning rate that is too small will result in a slow speed of convergence. In particular, the optimal value of the initial learning rate is dependent on the dataset under study, whereby for each dataset,

there exists an interval of good learning rates at which the performance does not vary much [99]. Moreover, gradient descent algorithms can often converge to a sub-optimal local minimum rather than the global minimum. To solve these challenges, several algorithms have been proposed, such as momentum SGD, RMSProp [98], nesterov accelerated gradient [100], Adagrad [101], and AdaDelta [102]. For instance, Adagrad and Adam are independent of the initial value of the learning rate while RMSProp relies heavily on a good choice of an initial learning rate. Moreover, one can use pruning techniques [103] to minimize the number of neurons in ANNs and make the ANNs become smaller and faster. The basic concept of pruning is that to eliminate ANN neurons that may not contribute to the output as they are not relevant to the learning task.

It is worth noting that two central problems in training ANNs are *overfitting* and *underfitting*. Overfitting corresponds to the case in which the model learns the random fluctuations and noise in the training dataset to the extent that they negatively impact the model's ability to generalize when fed with new data. This occurs mainly when the dataset is too small compared to the number of model parameters that must be learned. On the other hand, underfitting occurs when a learning algorithm cannot capture the underlying trend of the data. Intuitively, underfitting occurs when the learning algorithm does not fit the data well enough. Therefore, one must carefully choose the architecture of an ANN along with the proper training methods to avoid overfitting and underfitting.

Using the aforementioned training algorithms, the values of the weight matrix and bias can be updated to their optimal values, and, hence, a trained ANN can output the desired output signal. However, each type of ANNs is suitable for a particular type of data. For instance, RNNs are more convenient for time series data while spiking neural networks are good at modeling continuous data. Next, we elaborate more on two types of ANNs – RNNs and DNNs.

2.2.2 Recurrent Neural Networks

In this section, we overview the basics of RNNs. Then, we discuss the training algorithms that are generally used for training RNNs.

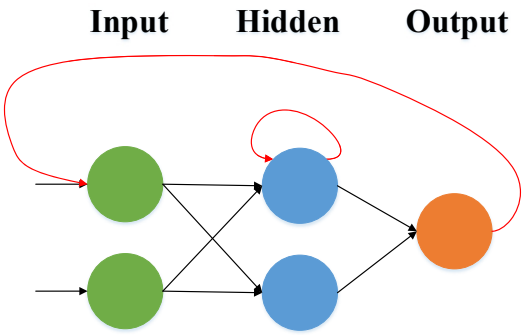

Figure 2.11. Recurrent neural network.

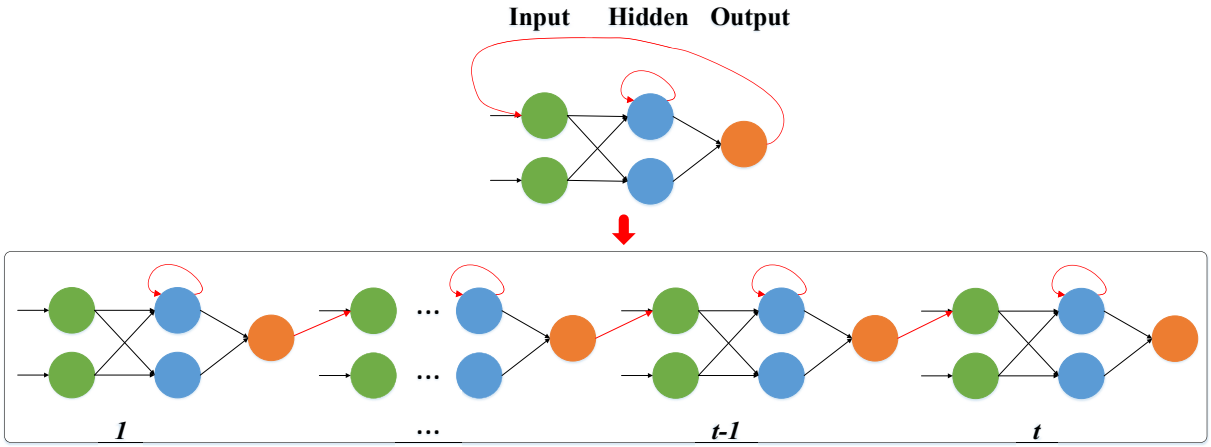

Figure 2.12. Architecture of an unfolded recurrent neural network.

2.2.2.1 Architecture of Recurrent Neural Networks

Unlike traditional ANNs that assume that all inputs are independent from each other or all outputs are independent from each other, *recurrent neural networks* [104] allow neuron connections from a neuron in one layer to neurons in previous layers, as shown in Fig. 2.11. This in turn enables the output of a neural network to depend, not only on the current input, but also on the historical input, as shown in Fig. 2.12. This allows RNNs to make use of sequential information and exploit dynamic temporal behaviors such as those faced in mobility prediction, handwriting recognition, or speech recognition. RNNs can also be seen as an ANN that has a "memory", which in essence allows them to store historical information and thus perform time-related tasks such as users' mobility pattern predictions compared to traditional ANNs (e.g., FNNs). In terms of architecture, the key components of a given RNN can be specified as follows:

- *Input signal* $\boldsymbol{x}_t$: this signal represents the input data to a given RNN at time t.

- *Input weight matrix* $\boldsymbol{W}_{in}$: this matrix represents the strengths of the connections between the neurons in the input layer and the neurons in the hidden layers.

- *Output weight matrix* $\boldsymbol{W}_{out}$: this matrix is used to represent the strengths of the connections between the neurons in the output layer and the neurons in the hidden layers.

- *Recurrent weight matrix* $\boldsymbol{W}$: The hidden weight matrix is defined as the recurrent weight matrix, which captures the strengths of the connections between the neurons in the hidden layers of the RNN.

- *Hidden state* $\boldsymbol{s}_t$: this is effectively the hidden state of a neuron in the hidden layer at time t. The hidden state represents the value of the activation function at time t, which is calculated based on the previous hidden state $\boldsymbol{s}_{t-1}$ and the input at time t. $\boldsymbol{s}_t$ can be computed using different methods for different recurrent neural networks. For most commonly used RNNs, we have $\boldsymbol{s}_t = f(\boldsymbol{W}\boldsymbol{s}_{t-1} + \boldsymbol{W}_{\text{in}}\boldsymbol{x}_t)$ where $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ or $f(x) = \max(0, x)$. However, in more elaborate types of RNNs, such as long short-term memory algorithm [19], each neuron needs to decide what to keep in and what to erase from the hidden state.

- *Output signal*: $\boldsymbol{y}_t$ is the output of a RNN at time t, representing the output signal.

Clearly, we can see that the basic architecture of RNNs is similar to that of FNNs except for the generation of the input, output, and recurrent weight matrices. Moreover, the hidden state in RNNs depends on both current and historical inputs, which enables RNNs to store the historical information. However, when the architecture of an ANN changes from FNNs to RNNs, traditional training methods may not be applicable to RNNs. Hence, next, we introduce training methods suitable for RNNs.

2.2.2.2 Training in Recurrent Neural Networks

In the RNN architecture, the connections between units will form a directed cycle and, hence, the feedforward gradient descent algorithms such as backpropagation cannot be directly used. This is due to the fact that the error backpropagation

pertaining to a backpropagation algorithm requires no cycles in the connections between the ANN neurons. In consequence, the *backpropagation through time (BPTT) algorithm* [105] is more commonly used for training RNNs. The BPTT approach unfolds the recurrent network in time, by stacking identical copies of it, and redirecting connections within the network to obtain connections between subsequent copies, as shown in Fig. 2.12. In consequence, the BPTT algorithm actually transforms an RNN into an FNN, which is amenable for training by a backpropagation algorithm. However, due to the cycle connections in RNNs, BPTT may get more often trapped in numerous sub-optimal local minima compared to the gradient descent algorithms used for training FNNs. Moreover, like backpropagation, the gradient in BPTT is computed based on the complete training set, which may become time consuming if the size of the training set is very large.

To overcome these drawbacks in BPTT training, real-time recurrent learning (RTRL) [106] can be used to compute the exact error gradient at every time step, which is suitable for online learning tasks. In contrast to the BPTT that unfolds RNNs in time, RTRL propagates error forward in time. From (2.6), we can see that the gradient value with respect to w at time t is $\frac{\partial E(t)}{\partial w}$. In RTRL, the update of weight w depends not only on the gradient value at time t but also on the gradient value at the previous time instant, i.e., $w_{t+1} = w_t - \gamma \sum_{k=0}^{t} \frac{\partial E(k)}{\partial w_k}$. In RTRL, the gradient of errors propagates forward in time rather than backward in time as in the BPTT algorithm and, therefore, there is no need to unfold RNNs as needed by the BPTT algorithm. However, the time complexity of RTRL is $O\left(N_w^4\right)$ where N_w is the number of neurons in the considered RNN. In contrast, BPTT has a time complexity of $O\left(N_w^2 G\right)$ where G is length of the input data.

Next, we elaborate more on a particular type of RNNs, namely the echo state network.

2.2.2.3 RNN Example: Echo State Network

Here, we introduce a type of RNNs that is conceptually simple and easy to implement, called *echo state networks* [35]. Since their inception, ESNs proved to be a highly practical type of RNNs due to their effective approach for training the neural network [107]. In ESN, the input weight matrix and hidden weight matrix are randomly generated without any specific training. Therefore, one would need

to train the output weight matrix only for ESN. Moreover, ESNs belong to a class of algorithms in the framework of *reservoir computing* (RC) [108]. In general, an ANN algorithm is considered part of the framework of RC if its input signals are mapped to a higher dimensional dynamics of the randomly generated hidden layers, known as a *reservoir*, and the dynamics of the reservoir are mapped to the desired output signals by using a simple training method such as backpropagation. The main benefit of RC is that the neural network training is performed only at the readout stage while the input and hidden weight matrices are fixed. ESNs can, in theory, approximate arbitrary nonlinear dynamical system with arbitrary precision. In wireless networks, ESNs admit many natural applications, such as content prediction, resource management, and mobility pattern estimation. Next, the specific architecture and training methods for ESNs are introduced.

• *Architecture of an Echo State Network:* ESNs use an RNN architecture with only one hidden layer. The generation of an ESN can be given as follows:

- *Generation of a Reservoir:* Generate a large random reservoir that is represented by the tuple $(\boldsymbol{W}_{\text{in}}, \boldsymbol{W}, \alpha)$ where α is known as the leaking rate which can be seen as the speed of the reservoir update dynamics, discretized in time. As we mentioned previously, the dynamical system in RC is known as a reservoir. In ESN, the input and hidden weight matrices are jointly known as the reservoir. Setting the leaking rate α must match the speed of the dynamics of hidden state $\boldsymbol{s}_t$ and output $\boldsymbol{y}_t$. Here, $\boldsymbol{W}_{\text{in}}$ and $\boldsymbol{W}$ is generated randomly. In particular, $\boldsymbol{W}$ is a sparse matrix while $\boldsymbol{W}_{\text{in}}$ is a dense matrix. The generation of $\boldsymbol{W}_{\text{in}}$ and $\boldsymbol{W}$ are determined by the training data and other ESN parameters. If one ESN uses discrete bi-valued distribution, i.e., $(-0.5, 0.5)$, to generate $\boldsymbol{W}_{\text{in}}$ and $\boldsymbol{W}$, then the ESN tends to have a slightly less rich signal space (there is a non-zero probability of identical neurons), but might render the analysis of what is happening in the reservoir easier. To allow ESNs to store historical information, the reservoir should satisfy the so-called *echo state property* (ESP) which means that the hidden state $\boldsymbol{s}_t$ should be uniquely defined by the fading history of the input $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_t$. This is in contrast to traditional ANNs such as FNNs that need to adjust the weight values of the neurons in hidden layers, ESNs only need to guarantee the ESP. To guarantee the echo state property of an ESN, the spectral radius of $\boldsymbol{W}$ should be smaller than 1. The scaling of $\boldsymbol{W}_{\text{in}}$ is another key method to optimize an ESN. In order to have a small

number of freely adjustable parameters, all elements in $\boldsymbol{W}_{\text{in}}$ are scaled using a single scaling value. If the input signals are of different types and have different contributions to the learning task, one should separately optimize the scalings of $\boldsymbol{W}_{\text{in}}$. For very linear tasks, $\boldsymbol{W}_{\text{in}}$ should be small, letting the neruons operate around 0. For large $\boldsymbol{W}_{\text{in}}$, the neurons will get easily close to their -1 and 1 values acting in a more nonlinear, binary switching manner.

- *ESN Implementation:* Run ESN using the training input $\boldsymbol{x}_t$ and collect the corresponding hidden states $\boldsymbol{s}_t$. The input $\boldsymbol{x}_t$ can be a vector or a scalar, which depends on the training dataset. Normalization of input data $\boldsymbol{x}_t$ can keep the input $\boldsymbol{x}_t$ bounded and avoid the hidden weight matrix being infinity. In general, the input data from the beginning of the training will be discarded and not used for training $\boldsymbol{W}_{\text{out}}$ since it may introduce an unnatural starting state which is not normally visited once the network has learnt the task. The amount of input data to discard depends on the memory of the ESN.

- *Training Output weight matrix:* Compute the output weight matrix $\boldsymbol{W}_{\text{out}}$ from the reservoir using a training algorithm such as gradient descent or ridge regression (explained next) to minimize the mean square error (MSE) between the targeted output and action output.

- *Generate Output:* Use the trained network on new input data $\boldsymbol{x}$ computing $\boldsymbol{y}_t$ by employing the trained output weights $\boldsymbol{W}_{\text{out}}$.

Given the components of ESNs, we will describe the activation value of each neuron. Even though the input and hidden weight matrices are fixed (randomly), all neurons of an ESN will have their own activation values (hidden state). As opposed to the classical RNNs in which the hidden state depends only on the current input, in ESNs, the hidden state will be given by:

$$\tilde{\boldsymbol{s}}_t = f(\boldsymbol{W}[1; \boldsymbol{s}_{t-1}] + \boldsymbol{W}_{\text{in}}\boldsymbol{x}_t)\,, \tag{2.10}$$

$$\boldsymbol{s}_t = (1-\alpha)\,\boldsymbol{s}_{t-1} + \alpha\tilde{\boldsymbol{s}}_t, \tag{2.11}$$

where $f(x)$ is the tanh function and $[\cdot;\cdot]$ represents a vertical vector (or matrix) concatenation. The model is also sometimes used without the leaky integration, which is a special case for $\alpha = 1$ which yields $\tilde{\boldsymbol{s}}_t = \boldsymbol{s}_t$. From (2.10), we can

see that the scaling of $\boldsymbol{W}_{\text{in}}$ and $\boldsymbol{W}$ determines the proportion of how much the current state $\boldsymbol{s}_t$ depends on the current input $\boldsymbol{x}_t$ and how much on the previous state $\boldsymbol{s}_{t-1}$. Here, a feedback connection from $\boldsymbol{y}_{t-1}$ to $\boldsymbol{s}_t$ can be applied to the ESNs, which is defined as a weight matrix $\boldsymbol{W}_{\text{fb}}$. Hence, (2.10) can be rewritten as $\tilde{\boldsymbol{s}}_t = f\big(\boldsymbol{W}[1; \boldsymbol{s}_{t-1}] + \boldsymbol{W}_{\text{in}}\boldsymbol{x}_t + \boldsymbol{W}_{\text{fb}}\boldsymbol{y}_{t-1}\big)$.

Based on the hidden state $\boldsymbol{s}_t$, the output signal of ESN can be given by:

$$y_t = \boldsymbol{W}_{\text{out}}\left[1; \boldsymbol{s}_t; \boldsymbol{x}_t\right]. \tag{2.12}$$

• *Training in Echo State Networks*: The objective of the training process in ESNs is to minimize the MSE between the targeted output and actual output. When this MSE is minimized, the actual output will be the target output which can be given by $\boldsymbol{y}_t^{\text{D}} = \boldsymbol{W}_{\text{out}}\left[1; \boldsymbol{s}_t; \boldsymbol{x}_t\right]$ where $\boldsymbol{y}_t^{\text{D}}$ is the targeted output. Therefore, the training goal is to find an optimal $\boldsymbol{W}_{\text{out}}$ that enables $\boldsymbol{W}_{\text{out}}\left[1; \boldsymbol{s}_t; \boldsymbol{x}_t\right]$ to be equal to $\boldsymbol{y}_t^{\text{D}}$. In contrast to conventional recurrent neural networks that require gradient-based learning algorithms, such as BPTT mentioned in Subsection 2.2.2.2 to adjust all input, hidden, and output weight matrices, ESNs only need to train the output weight matrix with simple training methods such as the *least mean squares* (LMS) method. The LMS algorithm, which is a stochastic gradient descent algorithm, allows the training of $\boldsymbol{W}_{\text{out}}$ in an *online* manner [109]. At every time step t, the LMS algorithm changes $\boldsymbol{W}_{\text{out}}$ in the direction of minimizing the instantaneous squared error $\left\|\boldsymbol{y}_t^{\text{D}} - \boldsymbol{y}_t\right\|^2$. LMS is a first-order gradient descent method, locally approximating the error surface with a hyperplane. However, this approximation in LMS is not always accurate. In particular, the curvature of the error surface is very different in different directions. To overcome this disadvantage, a learning algorithm named the recursive least squares (RLS), can be used for training ESNs. RLS, is insensitive to the detrimental effects of eigenvalue spread and exhibits a much faster convergence. Demonstrations of RLS for ESNs are presented in [110] and [111]. The backpropagation-decorrelation in [112] and the so-called FORCE learning algorithm in [113] are two other powerful methods for online training of single-layer output with feedback connections. Hence, the output weight matrix of each ESN can be optimized using different training methods. One can select the most suitable ESN training algorithm according to the scenario and target performance needed.

Next, we give a general overview on DNNs, which in essence, are ANNs with multiple hidden layers. We investigate two types of DNNs, namely, LSTM and

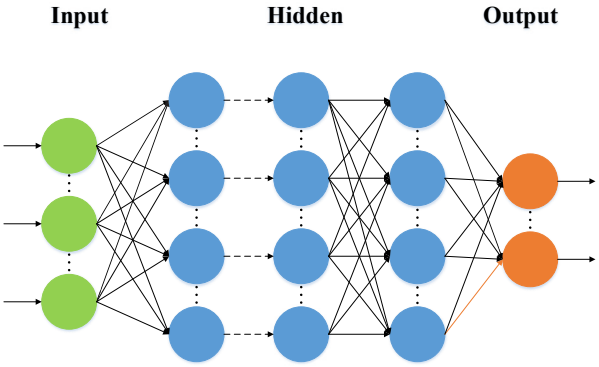

Figure 2.13. Architecture of a DNN.

deep ESNs.

2.2.3 Deep Learning

A *deep neural network* is an ANN with multiple hidden layers between the input and output layers [114], as shown in Fig. 2.13. Therefore, a DNN models high-level abstractions in data through multiple nonlinear transformations and thus learning multiple levels of representation and abstraction [114]. Several types of DNNs exist such as deep convolutional networks, deep RNNs, deep belief networks, deep feedforward networks, deep SNNs, deep Q-learning, deep ESN, deep residual network (ResNet), and long-short term memory [114]. The main reasons that have enabled a paradigm shift from conventional, shallow ANNs, towards DNN possible and desirable include recent advances in computing capacity due to the emergence of capable processing units, the wide availability of data for DNN training, and the emergence of effective DNN training algorithms such as those made possible by the use of rectified linear units (ReLUs) instead of sigmoid or tanh functions [115]. As opposed to shallow ANNs that have only one hidden layer, a DNN having multiple layers is more beneficial due to the following reasons:

- *Number of neurons:* Generally, a shallow network would require a lot more neurons than a DNN for the same level of performance. In fact, the number of units in a shallow network grows exponentially with the complexity of the task.

- *Task learning:* While shallow networks can be effective to solve small-scale problems, they can be ineffective when dealing with more complex problems, such as proactive resource management in wireless networks. In fact, the main issue is that shallow networks are very good at memorization, but

not so good at generalization. As such, DNNs are more suitable for many real-world tasks which often involve complex problems that are solved by decomposing the function that needs to be learned into a composition of several simpler functions thus making the learning process effective.

It is worth noting that, although DNNs have a large capacity to model a high degree of nonlinearity in the input data, a central challenge is that of overfitting. While overfitting can be a challenge in any type of ANN, typically, it can be overcome by simple regularization methods [116]. However, in DNNs, it becomes particularly acute due to the presence of a very large number of parameters. To overcome this issue, several advanced regularization approaches, such as dataset augmentation, early stopping, and weight decay [116] have been designed. These methods modify the learning algorithm so that the test error is reduced at the expense of increased training error. For instance, data augmentation overcomes overfitting by synthetically creating more data while early stopping aims at interrupting the training process once the performance of the model on a validation set gets worse. A validation set is a set of examples that are not used for either the training or the testing process.

2.2.3.1 Training Deep Neural Networks

DNNs are often much harder to train than shallow ANNs due to the instability of their gradient that occurs when training them with gradient-based methods such as those described earlier. In such conventional methods, each one of the ANN's weights receives an update proportional to the gradient of the error function with respect to the current weight in each iteration of training. In particular, the weights and the activation functions (or more precisely, their derivatives) that the gradient passes through will affect the magnitude of the gradients. Here, note that the gradient by the backpropagation algorithm is computed by the chain rule. Therefore, multiplying n of the gradients at each layer makes the gradients at the "front" layers, in an n-layer DNN, exponentially decrease or increase with n for small gradient values within range (-1, 1) or for large gradient values, respectively. This is obviously not a major problem in conventional shallow ANNs, as they have only one single layer. For example, the tanh derivative is < 1 for all inputs except 0 and the sigmoid is always ≤ 0.25 when used as activation functions $f(\cdot)$ in (2.4). These two problems are known as the *vanishing gradient*

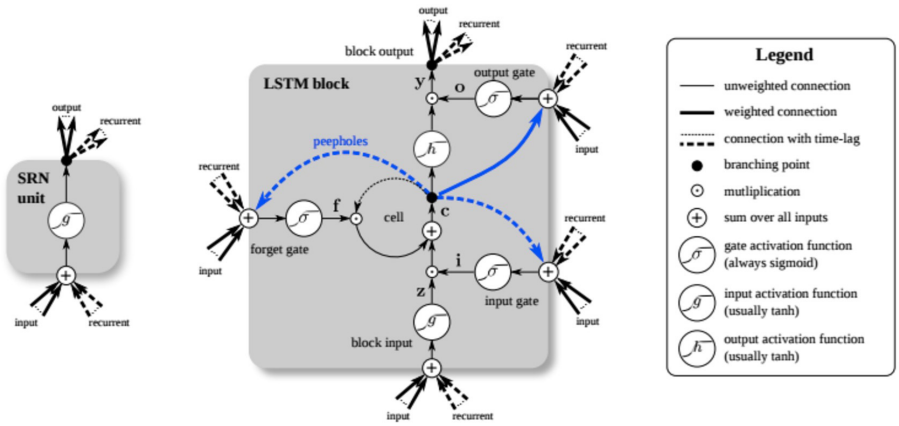

Figure 2.14. Architecture of an LSTM [9].

problem and the *exploding gradient problem*, respectively, and they can result in having different layers in DNNs learn at vastly different speeds. For instance, for a vanishing gradient problem, when latter layers in the DNN are learning well, early layers often learn almost nothing during training. To overcome this instability, several techniques for training DNNs have been proposed in the literature [19,117–119] such as adaptive learning rate algorithms (e.g., Adagrad, Adam, AdaDelta, RMSProp [101, 117, 120]), multi-level hierarchy [118], LSTM [19] and ResNets [119]. For instance, multi-level hierarchy of DNNs [118] is based on pre-training one level at a time through unsupervised learning and then fine-tuning through backpropagation. Meanwhile, the LSTM is a special type of deep RNN with the identity function with a derivative of 1 being its activation function [19]. Therefore, in LSTM, the backpropagated gradient neither vanishes nor explodes when passing through, but remains constant, and, thus, iterative gradient descent such as backpropagation through time can be used for training LSTMs.

The above discussion gives a brief overview on general DNNs. Next, we elaborate more on LSTM, a special kind of DNN that is capable of storing information for long periods of time by using an identity activation function for the memory cell. This in turn makes LSTM suitable for various wireless communication problems such as channel selection in which SBSs or WAPs need to learn a sequence of future traffic patterns on each channel and thus allocate the available radio resources accordingly.

2.2.3.2 DNN Example (1): LSTM

LSTMs are a special kind of "deep learning" RNNs that are capable of storing information for either long or short periods of time. In particular, the activa-

Table 2.2 Various behaviors of an LSTM cell.

Input gate	**Forget gate**	**Behavior**
0	1	remember the previous value
1	1	add to the previous value
0	0	erase the value
1	0	overwrite the value

tions of an LSTM network correspond to short-term memory, while the weights correspond to long-term memory. Therefore, if the activations can preserve information over long duration of time, then this makes them long-term short-term memory. Although both ESN and LSTM are good at modeling time series data, LSTM cells have the capability of dealing with long term dependencies. An LSTM network contains LSTM units each of which having a cell with a state c_t at time t. Access to this memory unit, shown in Fig.2.14, for reading or modifying information is controlled via three gates:

- *Input gate (i_t):* controls whether the input to is passed on to the memory cell or ignored.
- *Output gate (o_t):* controls whether the current activation vector of the memory cell is passed on to the output layer or not.
- *Forget gate (f_t):* controls whether the activation vector of the memory cell is reset to zero or maintained.

Therefore, an LSTM cell makes decisions about what to store, and when to allow reads, writes, and erasures, via gates that open and close. At each time step t, an LSTM receives inputs from two external sources, the current frame $\boldsymbol{x}_t$ and the previous hidden states of all LSTM units in the same layer $\boldsymbol{s}_{t-1}$, at each of the four terminals (the three gates and the input). These inputs get summed up, along with bias factors b_f, b_i, b_o, and b_c. The gates are activated by passing their total input through the logistic function. Table 2.2 summarizes the various behaviors an LSTM cell can achieve depending on the values of the input and forget gates. Moreover, the update steps of a layer of LSTM units are summarized in the following equations:

$$g_t = f_g(\boldsymbol{W}_f \boldsymbol{x}_t + \boldsymbol{U}_f \boldsymbol{s}_{t-1} + \boldsymbol{b}_f), \tag{2.13}$$

$$i_t = f_g(\boldsymbol{W}_i \boldsymbol{x}_t + \boldsymbol{U}_i \boldsymbol{s}_{t-1} + \boldsymbol{b}_i), \tag{2.14}$$

$$o_t = f_g(\boldsymbol{W}_o \boldsymbol{x}_t + \boldsymbol{U}_o \boldsymbol{s}_{t-1} + \boldsymbol{b}_o), \tag{2.15}$$

$$c_t = g_t \odot c_{t-1} + i_t \odot f_c(\boldsymbol{W}_c \boldsymbol{x}_t + \boldsymbol{U}_c \boldsymbol{s}_{t-1} + \boldsymbol{b}_c), \tag{2.16}$$

$$\boldsymbol{s}_t = o_t \odot f_h(c_t), \tag{2.17}$$

where g_t, i_t, and o_t are the forget, input and output gate vectors at time t, respectively. $\boldsymbol{x}_t$ is the input vector at time t. c_{t-1} is the cell state vector (i.e., internal memory) and $\boldsymbol{s}_{t-1}$ is the hidden/output vector at time $(t-1)$. $\boldsymbol{W}_f$ and $\boldsymbol{U}_f$ are the weight and transition matrices of the forget gate, respectively. f_g, σ_c, and f_h are the activation functions and correspond respectively to the sigmoid, tanh and tanh functions. $\odot$ denotes the Hadamard product. Compared to a standard RNN, LSTM uses additive memory updates and separates the memory c from the hidden state $\boldsymbol{s}$, which interacts with the environment when making predictions. To train an LSTM network, the stochastic gradient descent algorithm that was introduced in Subsection 2.2.1.2 can be employed.

LSTM is thus suitable for applications involving sequential learning; it can classify, process and predict time series given time lags of unknown size and duration between important events. Several variants of LSTM exist such as bidirectional LSTM, sequence-to-sequence LSTM (a.k.a. encoder-decoder LSTM) [121], peephole LSTM [122], and gated recurrent unit (GRU) [123]. For instance, the encoder-decoder LSTM is suitable for solving problems with sequences whose lengths are not known a-priori [121]. The authors in [124] show that some modifications to the LSTM architecture such as coupling the input and the forget gates or removing peephole connections simplify the LSTM architecture without significantly degrading its performance.

In this subsection, we discussed LSTMs that are particularly suitable for sequence learning. Next, we elaborate more on another type of DNNs, namely deep ESNs, which are characterized by their ability to represent features at different levels of abstraction while preserving the RC training efficiency.

2.2.3.3 DNN Example (2): Deep ESN

Here, we introduce a different type of DNNs, namely, deep ESNs. In essence, multiple non-linear reservoir layers can be stacked on top of each other resulting in a *deep ESN architecture*. Deep ESNs exploit the advantages of a hierarchical temporal feature representation at different levels of abstraction while preserving the RC training efficiency. They can learn data representations at different levels

of abstraction, hence disentangling the difficulties in modeling complex tasks by representing them in terms of simpler ones hierarchically. Let $N_R^{(n)}$ be the number of internal units of the reservoir at layer n, N_U be the external input dimension, and N_L be the number of layers in the stack. Next, we define the following deep ESN components:

- $\boldsymbol{x}_t \in \mathbb{R}^{N_U}$ the external input at time t,
- $\boldsymbol{s}_t^{(n)} \in \mathbb{R}^{N_R^{(n)}}$ as the state of the reservoir at layer n and time t,
- $\boldsymbol{W}_{\text{in}}^{(n)}$ as the input-to-reservoir matrix at layer n, where $\boldsymbol{W}_{\text{in}}^{(n)} \in \mathbb{R}^{N_R^{(n)} \times N_U}$ for $n = 1$, and $\boldsymbol{W}_{j,\text{in}}^{(n)} \in \mathbb{R}^{N_R^{(n)} \times N_R^{(n-1)}}$ for $n > 1$,
- $\boldsymbol{W}^{(n)} \in \mathbb{R}^{N_R^{(n)} \times N_R^{(n)}}$ as the recurrent reservoir weight matrix at layer n,
- $\boldsymbol{W}_{\text{out}} \in \mathbb{R}^{|\mathcal{Z}| \times (N_U + \sum_n N_R^{(n)})}$ as the reservoir-to-output matrix for layer n only.

The objective of the deep ESN architecture is to approximate a function $\boldsymbol{F} = (F^1, F^2, \cdots, F^{N_L})$ based on the input vector at each time t. For each $n = 1, 2, \cdots, N_L$, the function $F^{(n)}$ describes the evolution of the state of the reservoir at layer n, i.e., $\boldsymbol{s_t^{(n)}} = F^{(n)}(\boldsymbol{x}, \boldsymbol{s}_{t-1}^{(n)})$ for $n = 1$ and $\boldsymbol{s_t^{(n)}} = F^{(n)}(\boldsymbol{s}_t^{(n-1)}, \boldsymbol{s}_{t-1}^{(n)})$ for $n > 1$. Here, the initialization of $\boldsymbol{W}^{(n)}$ should satisfy the echo state property of deep ESN. In essence, the echo state property of deep ESN states that for each input sequence of length N, $\boldsymbol{S}_N = [\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots, \boldsymbol{x}(N)]$, and for all couples of deep ESN initial states $\boldsymbol{s}, \boldsymbol{s}' \in \mathbb{R}^{\sum_n N_R^{(n)}}$, $\| \widehat{F}(\boldsymbol{S}_N, \boldsymbol{s}) - \widehat{F}(\boldsymbol{S}_N, \boldsymbol{s}') \| \to 0$ as $N \to \infty$, where $\widehat{F}$ denotes the iterated version of the deep ESN state transition function F. Equivalently, the current state $\boldsymbol{s}$ of the network is a function of its past input history independently of initial state values [125]. To satisfy the ESP, a necessary condition is that the spectral radius of $\boldsymbol{W}^{(n)}$ (i.e., the largest eigenvalue in absolute value), $\rho^{(n)}$, is strictly less than 1 [125]. If this condition is violated, the dynamical reservoir is locally asymptotically unstable at the zero state $0 \in \mathbb{R}^{N_R^{(n)}}$ and echo states cannot be guaranteed if the null sequence is an admissible input for the system. In fact, the value of $\rho^{(n)}$ is related to the variable memory length and the degree of contractivity of reservoir dynamics, with larger values of $\rho^{(n)}$ resulting in longer memory length. Varying the values of $\rho^{(n)}$ implies a variability of contractivity and memory length among the state dynamics of different layers.

Here, note that for deep ESN architecture, we distinguish between two types of inputs: external input, $\boldsymbol{x}_t$, that is fed to the first layer of the deep ESN and

corresponds to the current state of the network and input that is fed to all other layers for $n > 1$. The input to any layer $n > 1$ at time t corresponds to the state of the previous layer, $\boldsymbol{s}_t^{(n-1)}$. Therefore, the state transition function of the first layer $\boldsymbol{s}_t^{(1)}$ will be:

$$\boldsymbol{s}_t^{(1)} = (1 - \alpha^{(1)})\boldsymbol{s}_{t-1}^{(1)} + \alpha^{(1)}\tanh(\boldsymbol{W}_{\text{in}}^{(1)}\boldsymbol{x}_t + \boldsymbol{W}^{(1)}\boldsymbol{s}_{t-1}^{(1)}), \tag{2.18}$$

where $\alpha^{(n)} \in [0, 1]$ is the leaky parameter at layer n which relates to the speed of the reservoir dynamics in response to the input, with larger values of $\alpha^{(n)}$ resulting in a faster response of the corresponding n-th reservoir to the input. The state transition, $\boldsymbol{s}_t^{(n)}$, for $n > 1$ is given by:

$$\boldsymbol{s}_t^{(n)} = (1 - \alpha^{(n)})\boldsymbol{s}_{t-1}^{(n)} + \alpha^{(n)}\tanh(\boldsymbol{W}_{\text{in}}^{(n)}\boldsymbol{s}_t^{(n-1)} + \boldsymbol{W}^{(n)}\boldsymbol{s}_{t-1}^{(n)}), \tag{2.19}$$

The output y_t of the deep ESN at time t is used to estimate the reward after training $\boldsymbol{W}_{\text{out}}$ and can be computed as:

$$y(\boldsymbol{x}_t) = \boldsymbol{W}_{\text{out},t}[\boldsymbol{x}(t), \boldsymbol{s}^{(1)}(t), \boldsymbol{s}_t^{(2)}, \cdots, \boldsymbol{s}_t^{(n)}]. \tag{2.20}$$

Here, note that training deep ESN networks can be achieved via similar methods as training shallow ESN. Moreover, the authors in [126] compare the performance of a deep ESN architecture with decreasing leaky parameter, $\alpha^{(n)}$, for increasing layer depth with its shallow counterpart. Results show that the variability of the leaky parameter has a great impact on the separation among the emerging time-scales dynamics, reaching longer times-scales than the shallow ESN with the slowest dynamics. In fact, this characterization is a result of the interplay between layering and leaky integration variability, and indeed it is lost when non-stacked architectures are considered.

Next, we give general background on game theory, and in particular, network formation games. Game theory is a mathematical framework for modelling the conflict and cooperation between intelligent rational decision-makers. Network formation games are a branch of game theory which essentially deal with the formation of a network among different players in a given network.

2.3 Game Theory

In this section, we provide a general overview on game theory and some of its general concepts. Then, we elaborate more on a particular type of games, namely network formation games.

2.3.1 Basics of Game Theory

Game theory is a formal analytical framework with a set of mathematical tools to study the *strategic* interactions among independent *rational* players [127]. Throughout the past decades, game theory has made a revolutionary impact in different disciplines ranging from economics, politics, philosophy, or even psychology. Likewise, the emergence of large-scale distributed wireless networks, as well as the recent interest in mobile flexible networks where the nodes are autonomous decision makers has made game theory an interesting tool for various competitive and cooperative problems in wireless communications [21, 46, 127]. In essence, the application of game-theoretic concepts in wireless communication networks lies at the heart of the need for self-organizing, self-configuring, and self-optimizing networks [127]. In its basic form, a game consists of:

- **Players:** with conflicting interests or mutual benefit.
- **Actions or strategies:** set of actions available for each player.
- **Utility or payoff:** corresponds to the benefit that a player can obtain upon taking a particular action.

Different types of games exists. Broadly speaking, these games can be classified into two main categories: non-cooperative games, in which different players have conflicting interests, and cooperative games, in which different players may have mutual benefits. In these types of games, it is important to characterize the solution that the players of each game aim at reaching. As such, different solution concepts exist for different types of games. In this thesis, we limit our discussion to the most commonly adopted solution concepts for non-cooperative games. In what follows, we summarize such concepts:

- *Nash equilibrium*: The NE is a fundamental solution concept for strategic-form games, on which many other concepts are built. It corresponds to the stable state of a system involving the interaction of multiple players, in which no player can gain by a unilateral deviation of his strategy if the strategies of the other players remain unchanged. Therefore, at NE, each player's choice of action is a best response (BR) to the actions taken by his opponents and, thus, the NE outcome can be regarded as a steady state of a strategic interaction. Here, it is important to distinguish between pure,

mixed and behavioral NE. A pure strategy is an unconditional and defined choice that a player makes in a game, i.e., each player plays one specific strategy. On the contrary, a mixed strategy is an assignment of a probability to each pure strategy, i.e., at least one player in the game randomizes over some or all of their pure strategies. Meanwhile, a behavioral strategy exists in dynamic games. It allows each player to assign independent probabilities to the set of actions at each state of the game that is independent across the different states.

- *Pareto optimality*: The outcome of a game is Pareto optimal if there is no other outcome that makes every player at least as well off and at least one player strictly better off. In other words, it implies an operating point where no player can do better without "hurting" the others.

Given the above basic concepts in game theory, next, we elaborate more on a particular type of game, namely, the network formation game. In essence, network formation games correspond to situations where network structure plays a key role such as trading agreements, personal relationships, and wireless sensor networks. As such, these type of games rely on theories on how network structures matter and how they are formed. For these games, one needs to predict which networks can form and assess the stability and efficiency of these networks. Next, we define the basics of network formation games and the solution concepts adopted in such type of games.

2.3.2 Network Formation Games

Network formation games involve a number of independent decisions makers that interact with each other in order to form a suited network graph that connects them [44]. Therefore, these type of games are mainly suited for applications in which network structure plays a key role thus making them suitable for various wireless communication problems. In essence, network formation games capture two conflicting objectives of self-interested nodes in a network. On the one hand, such a node wishes to be able to reach all other nodes in the network; on the other hand, it wishes to minimize its cost of participation. A network formation game consists of:

- **Players:** that are connected in some network relationship.

- **Links:** a link represents a relationship between two players. It can be either a directed or an undirected link. The former refers to links with directed edges, i.e., directional links, while the latter corresponds to bidirectional links. An undirected graph can be simulated using a directed graph by using pairs of edges in both directions.

- **Network:** is a list of which pairs of players are linked to each other and thus corresponds to a relationship between the players. It is typically represented by a graph $\mathcal{G}$ with the players being the vertices. For any pair of players i and j, a formed link is denoted ij.

Here, it is important to define the basic notations for network formation games. A graph is denoted by g where $g + ij$ represents adding a link to a network and $g - ij$ represents deleting a link from a network. A *path* in a network g is a sequence of players $i_1, \ldots, i_K$ such that $i_k i_{k+1} \in g$ for each k. A *component* of a network g is a subnetwork g^c, i.e., every i and j that have at least one link in g^c are connected by a path and any link ij that is in g^c is also in g. A value function v is a function that provides the total amount accrued from a certain graph structure and thus an allocation rule defines how this value is divided between the players. Therefore, a network game can be defined as the pair (N, v) where N is the set of players and v is the value.

For a given network formation game, the main objective of the players is to form a network among them. Several approaches have been proposed for forming a network, each of which considering a different solution concept. In what follows, we overview the main considered solution concepts [44]:

- **Nash network (Baya and Goyal, 2000):** The most basic dynamic network formation process is BR dynamics, where at each round of the process, a profitable deviation is undertaken by one or a pair of nodes at a time. In these games, the strategy of a player is the other player with which it wants to link to. The outcome of such an approach is the Nash network, at which each player plays their BRs. Here, note that the notion of Nash network is applicable to directed graphs only. Although BR dynamics are attractively simple, they may fail to converge; further, BR dynamics can lead to inefficient equilibria or even to multiplicity of Nash networks. For instance, consider a game where each player simultaneously announces which other player he or she is willing to link to. It is always a NE for each player to say

that he or she does not want to form any links, anticipating that the other player will do the same i.e., the empty graph is always a NE. Generally, this allows for a large multiplicity of equilibria, many of which make little sense from a network perspective in which mutual interest is essential. Therefore, although suitable in some social networks applications, the model is unsuitable for complex problems such as communication networks. The concept of NE does not account for the fact that the players can discuss their decisions, which in fact is essential for the formation of networks. To model such a situation a stability concept that accounts for the consent of both players is required.

- **Pairwise stable networks (Jackson and Wolinsky, 1996):** Given the fact that, in network formation games, the consent of two nodes is required to form a single link, the stability of the outcome can be more accurately characterized than in Nash networks by considering bilateral deviations [128]. As such, a network is said to be *pairwise stable* if there is no incentive for any player to break a link that is formed with another player (unilateral deviation) and no pair of players have an incentive to establish a new link (bilateral deviation). The notion of pairwise stability accounts for the mutual consent of both players and is therefore applicable to undirected graphs only. Link stability, on the other hand, is a similar concept to pairwise stability but for directed graphs. Watts proposed dynamic model for forming pairwise stable networks [44]. The Watts process starts with an empty network. At each time t, a link is identified, and is added or deleted only if the resulting network defeats the previous network. For this process, myopic players are adopted, i.e., players update their strategic decisions considering only the current state of the network without taking into account the future evolution of the network. The process ends in a stable state if no more links can be added or deleted and a pairwise stable network is reached in that case. The stochastic dynamic model is another process that would reach a pairwise stable network and was proposed by Jackson and Watts in 2001. This process is in fact an improvement over the Watts model, by allowing the reverse of the players intentions to occur with probability ϵ. Nevertheless, like Nash network, a pairwise stable network may not always exist. In particular, this occurs when each network is defeated

by some adjacent network, and that these *improving paths* form cycles with no undefeated networks existing. Here, an *improving paths* is a sequence of networks $\{G_1, G_2, \ldots, G_k\}$ where G_k is defeated by the subsequent network G_{k+1}. Moreover, a *cycle* is an improving path $\{G_1, G_2, \ldots, G_k\}$ such that $G_1 = G_k$. Consequently, a network is pairwise stable if and only if it has no improving paths emanating from it [129]. Another drawback that makes the pairwise stability concept weak, is the fact that it is does not consider changes in multiple links at a time, but it only looks at changes that happen between single links. In other words, it does not consider deviations where players delete some links and add other links at the same time.

- **Pairwise Nash stable:** Pairwise-Nash equilibrium is a variation of NE where players are allowed to deviate by pairs. Therefore, as opposed to pairwise stable networks which are robust to one-link deviations, pairwise-Nash networks are robust to the bilateral one-link creation and to unilateral multi-link severance [130]. In other words, a network is said to be a pairwise-Nash equilibrium network if it is both pairwise stable and a NE outcome.

Given the above solution processes and concepts, it is important to characterize the efficiency of the resultant network. The proposed solution concepts exhibit different tradeoff between stability and efficiency. In essence, achieving both efficiency and stability is challenging for network formation games. For instance, the star network is an efficient network but is not pairwise stable while the empty network can be pairwise stable but non-efficient.

Chapter 3

Deep Learning for Proactive Resource Management for LTE in Unlicensed Spectrum

3.1 Introduction

As we have mentioned in Chapter 2, future mobile networks will essentially operate over the unlicensed band along with the conventional licensed spectrum. LTE-LAA has mainly emerged due to the dramatic growth in mobile data traffic on one hand and the scarcity of the licensed spectrum on the other hand. Therefore, to reap the benefits of LTE-LAA and thus allow a higher system capacity and better user experience, the main LTE-LAA challenges that were highlighted in Chapter 2 need to be addressed. These challenges include effective coexistence with existing WiFi networks, resource allocation, multiple access, traffic offloading from licensed to unlicensed spectrum, and inter-operator spectrum sharing [131]. In this chapter, we address the problem of LTE-WiFi coexistence in the unlicensed band. In particular, we propose a novel proactive resource management scheme for the coexistence of LTE and WiFi in the unlicensed spectrum. The proposed scheme allows different SBSs to autonomously update their channel selection and channel access probabilities based on the traffic load of WLAN on each of the unlicensed channels. The work presented in this chapter has appeared in the Proceedings of European Wireless Conference [20] and an extended journal version has been accepted for publication at the IEEE Transactions on Wireless Communications [21]. Table 3.1 provides a summary for the description of the

Table 3.1 Variables and notations.

Notation	Description
J	Number of LTE-LAA SBSs
W	Number of WAPs
C	Number of unlicensed channels
K_j	Number of LTE-LAA UEs associated with SBS j
P_t	Transmit power of SBS
T	Time window
t	Time epoch
τ_j	Stationary probability of WAP w
$\tau_{j,c,t}$	Stationary probability of SBS j on channel c at t
$t_{\max}$	Maximum fraction of time an unlicensed channel can be used
$\mathrm{CW}_{j,c,t}$	CW of SBS j on channel c at t
$\alpha_{j,c,t}$	Fraction of time during which SBS j transmits on channel c during t
$x_{j,c,t}$	The channel c that SBS j transmits over during t
M_c	Maximum channels that can be aggregated by SBS
λ	Learning rate
γ	Learning rate decay
$h_{m,t}$	Hidden vector of the traffic encoder m
$h_{j,t}$	Hidden vector of the action decoder of SBS j
ϕ	LSTM cell function
θ_t	Model parameters at t

main notations used in this chapter. Next, we give an overview on the related literature and then we present our problem statement and contribution.

3.1.1 Related Work

There has been a number of recent works [15, 16, 18, 25, 132–135] that study the problem of enhanced LTE-LAA and WiFi coexistence. This existing body of works can be categorized into two groups: channel access [25, 132–134] and channel selection [15, 18, 135]. The authors in [132–134] propose different channel access mechanisms based on LBT that rely on either an exponential backoff [132], a fixed/random CW size [133], or an adaptive CW size [134]. Nevertheless, an exponential backoff approach leads to unnecessary retransmissions while a fixed CW size cannot handle time-varying traffic loads thus yielding unfair outcomes. The authors in [25] develop a holistic approach for both traffic offloading and resource sharing across the licensed and unlicensed bands but considering one SBS. In [16], the authors study the problem of resource allocation with UL-DL

decoupling for LTE-LAA. The authors in [17] propose an inter-network coordination scheme with a centralized radio resource management for the LTE-WiFi coexistence. However, this prior art is limited to one unlicensed channel and does not jointly account for channel selection and channel access. In other words, these works do not analyze the potential gains that can be obtained upon aggregating or switching between different unlicensed channels. Operating on a fixed unlicensed channel limits the amount of cellular data traffic that can be offloaded to the unlicensed band and leads to an increase in the interference level caused to neighbouring WAPs operating on that same channel.

In terms of LTE-LAA channel selection, the authors in [18] propose a distributed approach based on Q-learning. A matching-based solution approach is proposed in [135], which is both distributed and cooperative. Moreover, the work in [15] combines channel selection along with channel access. Despite the promising results, all of these works [15, 18, 135] consider a reactive approach in which data requests are first initiated and, then, radio resources are allocated based on their corresponding delay tolerance value. Nevertheless, this sense-and-avoid approach can cause an underutilization of the spectrum due to the impulsive reconfiguration of the spectrum usage that does not account for the future dynamics of the network. Despite the predominance of the reactive LTE-WiFi coexistence solutions, cellular data traffic networks are known to exhibit statistically fluctuating and periodic demand patterns, especially applications such as file transfer, video streaming and browsing [136], therefore providing an opportunity for the network to exploit the predictable behavior of the UEs to smooth out the traffic over time and reduce the difference between the peak and the average load. Therefore, in a *proactive* approach, rather than reactively responding to incoming demands and serving them when requested, an SBS can predict traffic patterns and determine future off-peak times so that incoming traffic demand can be properly allocated over a given time window.

3.1.2 Problem Statement and Contribution

The main objective of this chapter is to propose a proactive LTE-WiFi coexistence scheme that would allow a better utilization of the unlicensed spectrum. This is mainly accomplished by either serving a fraction of the LTE-LAA traffic when requested or shifting part of it to the future, over a given time window, so as

to balance the occupancy of the unlicensed spectrum usage across time and, consequently, improve its degree of utilization. From the LTE-LAA network perspective, this will increase its transmission opportunities on the unlicensed spectrum, reduce the collision probability with WAPs and other SBSs and, hence, provide a boost for its throughput. Moreover, a proactive resource allocation scheme can exploit the inherent predictability of the future channel availability status so as to allocate the radio resources in a window of time slots based on the predicted requests. This, in turn, can lead to a decrease in the probability of occurrence of a congestion event while ensuring a degree of fairness to the WLAN.

The main contribution of the work presented in this chapter is a novel deep RL algorithm based on long short-term memory (RL-LSTM) cells for proactively allocating LTE-LAA resources over the unlicensed spectrum. The LTE-LAA resource allocation problem is formulated as a noncooperative game in which the players are the SBSs. To solve this game, we propose an RL-LSTM framework using which the SBSs can autonomously learn which unlicensed channels to use along with the corresponding channel access probability on each channel taking into account future environmental changes, in terms of WLAN activity on the unlicensed channels and LTE-LAA traffic loads. Unlike previous studies which are either centralized [15] or rely on the coordination among SBSs [134], the proposed scheme in this chapter is based on a self-organizing proactive resource allocation scheme in which the SBSs utilize past observations of the network state to build predictive models on spectrum availability and to intelligently plan channel usage over a finite time window. The use of long short term memory cells enables the SBSs to predict a sequence of interdependent actions over a long-term time horizon thus achieving long-term fairness among different underlying technologies. We show that, upon convergence, the proposed algorithm reaches to a mixed-strategy distribution which constitutes a mixed-strategy NE for the studied game. We also show that the gain of the proposed proactive resource allocation scheme and the optimal size of the prediction time window is a function of the traffic pattern of the dataset under study. Simulation results show that the proposed approach yields significant rate improvements compared to conventional reactive solutions such as instantaneous equal weighted fairness, proportional fairness and total network throughput maximization. The results also show that the proposed scheme prevents disruption to WLAN operation in the case large number of LTE operators selfishly deploy LTE-LAA in the un-

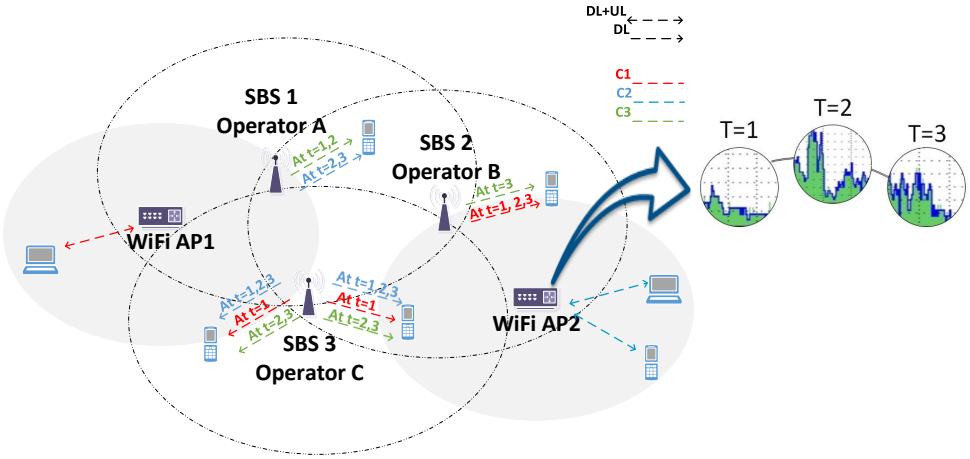

Figure 3.1. Illustration of the system model. In the above example, 3 SBSs belonging to different operators and 3 unlicensed channels are only shown for simplicity. The channel selection vector over a time window of 3 epochs is also shown.

licensed spectrum. In terms of priority fairness, results show that an efficient utilization of the unlicensed spectrum is guaranteed when both technologies are given equal weighted priorities for transmission on the unlicensed spectrum.

The rest of this chapter is organized as follows. In Section 3.2, we present the system model. Section 3.3 describes the proposed coexistence game model. The LSTM-based algorithm is proposed in Section 3.4. In Section 3.5, simulation results are analyzed. Finally, a summary is provided in Section 3.6.

3.2 System Model

Consider the downlink of an LTE-LAA network composed of a set $\mathcal{J}$ of J LTE-LAA SBSs belonging to different LTE operators, a set $\mathcal{W}$ of W WAPs, and a set $\mathcal{C}$ of C unlicensed channels as shown in Fig. 3.1. Each SBS $j \in \mathcal{J}$ has a set $\mathcal{K}_j$ of K_j LTE-LAA UEs associated with it. We consider a network scenario corresponding to environments such as work offices, a university campus, and airports in which the traffic load of a given WAP or SBS can be characterized through a particular model that typically remains unchanged over coarse periods of time (e.g., one day) [136]. We focus on the operation of the SBSs over the unlicensed band, while the licensed spectrum resources are assumed to be allocated in a conventional way [137]. Both SBSs and WAPs adopt the LBT access scheme and, thus, at a particular time, a given unlicensed channel is occupied by either an SBS or a WAP. We consider the LTE carrier aggregation feature using which the SBSs can aggregate up to five component carriers belonging to the same or different operating frequency bands [138]. This, in turn, would enable the SBSs to operate

on multiple unlicensed channels simultaneously thus maximizing their data rate during a particular transmission opportunity.

Our goal is to jointly determine the dynamic channel selection, carrier aggregation, and fractional spectrum access for each SBS, while guaranteeing long-term airtime fairness with WLAN and other LTE-LAA operators. The main motivation for adopting a long-term fairness approach is to avoid the underutilization of the unlicensed spectrum by either serving part of the LTE-LAA traffic when requested or shifting part of it in the future over a given time window in a way that would balance the occupancy of the unlicensed spectrum usage across time and, consequently, improve its degree of utilization. This will subsequently result in an increase in the transmission opportunities for LTE-LAA as well as a decrease in the collision probability for the WLAN. To realize this, we need to dynamically analyze the usage of various unlicensed channels over a particular time window. To this end, we divide our time domain into multiple time windows of duration T, each of which consists of multiple time epochs t, as shown in Fig. 3.2. Our objective is to proactively determine the spectrum allocation vector for each SBS at $t = 0$ over T while guaranteeing long-term equal weighted airtime share with WLAN. To guarantee a fair spectrum allocation among SBSs belonging to different operators, we consider inter-operator interference along with inter-technology interference. In fact, inter-operator interference is the consequence of the selfish behavior of different operators and could result in a degradation in the spectral efficiency if not managed. Next, we define $x_{j,c,t} = 1$ if channel c is selected by SBS j during time epoch t, and 0, otherwise, and $\alpha_{j,c,t} \in [0, 1]$. $x_{j,c,t}$ determines the channel c that is used by SBS j during time t and $\alpha_{j,c,t}$ is the channel access probability of SBS j on the unlicensed channel c at time t.

Since the 3GPP has identified LBT as an access mechanism for standardizing a global solution for the operation of LTE in the unlicensed spectrum, we consider a contention-based protocol for our proposed channel access mechanism [139]. In this protocol, prior to transmission, an SBS applies clear channel assessment for the duration of DIFS to detect the state of the channel (idle or busy) based on the detected energy level. If the channel is idle, the SBS would backoff for a random number between [0, CW] and if the medium was still free, it gets a transmit opportunity for up to 10 LTE sub-frames (considering priority class 1 devices [140]); it sends a reservation signal, e.g., clear-to-send (CTS), with the duration of its transmission period along with the remaining time period until the beginning of

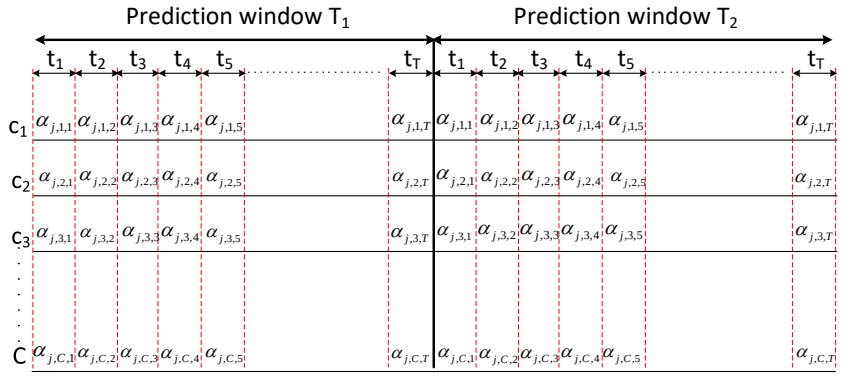

Figure 3.2. The division of the time domain into multiple time windows T, each of which consists of multiple time epochs t.

its next subframe. This allows prevention of other competing devices from getting access to the unlicensed channel until the beginning of the next subframe of the corresponding SBS and hence reserving the channel for transmission. On the other hand, if the channel was busy, the SBS keeps monitoring the channel until it becomes idle. Here, we note that our proposed algorithm is not fully compliant with the regulations in terms of CW size adjustment. In particular, we consider an exponential backoff scheme for WiFi while the SBSs adjust their CW size (and thus the channel access probability) on each of the selected channels in a way that would guarantee a long-term equal weighted fairness with WLAN and other SBSs. In essence, the exponential backoff access method that has been adopted by 3GPP for SBSs can lead to short-term unfairness [141]. This results from the fact that, after each collision, the colliding hosts double their CWs and, thus, have higher probability of choosing a larger backoff during which other hosts may benefit from channel access. This also means increased delay for hosts that doubled their CW. Therefore, the standard DCF method controls the load on the channel by reducing the number of contending hosts, because the hosts that have failed their transmission are likely to attempt to access the channel in the future. Moreover, hosts consider all failed transmissions as collisions in DCF, however, this is not always the case. Thus, DCF bases its load control on a biased indicator, which can potentially lead to lower performance and increased unfairness [141]. On the other hand, by having a fixed CW size for each SBS during each time epoch t, we can alleviate these problems and, more importantly, we can decouple the load control from handling failed transmissions. It is also worth noting that small CW sizes lead to an increase in the collision probability while

large CW sizes result in too much time spent waiting in idle slots. Therefore, an efficient access method should adapt the value of the CW of each SBS to the traffic conditions of the network.

To derive the throughput achieved by an LTE-LAA UE and a WAP, we first define the stationary probability of each WAP w and each SBS j, τ_w and $\tau_{j,c,t}$ respectively. The stationary probability is the probability with which a given base station attempts to transmit in a randomly chosen slot. Considering an exponential backoff scheme for WiFi, the stationary probability with which WAPs transmit a packet during a particular WiFi time slot, τ_w, will be given by [142]:

$$\tau_w = \frac{2(1-2q_w)}{(1-2q_w)(\mathrm{CW}_{\min}+1)+q_w\mathrm{CW}_{\min}(1-(2q_w)^m)}, \tag{3.1}$$

where q_w is the collision probability of a WAP, m is the maximum backoff stage where $\mathrm{CW}_{\max} = 2^m\mathrm{CW}_{\min}$. $\mathrm{CW}_{\min}$ and $\mathrm{CW}_{\max}$ are the minimum and maximum contention window size, respectively. For LTE-LAA, m=0 since no exponential backoff is considered, and, thus, the stationary probability of an SBS on a given unlicensed channel c during time epoch t will be $\tau_{j,c,t} = \frac{2}{\mathrm{CW}_{j,c,t}+1}$, where $\mathrm{CW}_{j,c,t}$ is the contention window size of SBS j on channel c during time epoch t. Therefore, we do not consider a contention stage for LTE-LAA and, thus, the CW size of the SBSs is not doubled after each unsuccessful transmission. Instead, the SBSs consider a fixed value for the CW for each time epoch t and this value is adjusted adaptively from one time epoch t to another in order to control the corresponding channel access probabilities over the unlicensed band for different time epochs. The collision probability of a WAP is defined as $q_w = 1-\prod_{v=1,v\neq w}^{W}(1-\tau_v)\prod_{j=1}^{J}(1-\tau_{j,c,t})$, where c is the channel used by WAP w. The throughput R_w of a WAP w during a particular WiFi time slot will be:

$$R_w = \frac{P_{w,\mathrm{succ}}\cdot E[D_w]}{P_{w,\mathrm{idle}}\cdot\theta + P_{w,\mathrm{busy}}\cdot T_b}, \tag{3.2}$$

where $E[D_w]$ is the expected payload size for WAP w, $P_{w,\mathrm{succ}} = \tau_w\prod_{v=1,v\neq w}^{W}(1-\tau_v)\prod_{j=1}^{J}(1-\tau_{j,c,t})$ is the probability of a successful transmission, $P_{w,\mathrm{idle}} = \prod_{j=1}^{J}(1-\tau_{j,c,t})\prod_{w=1}^{W}(1-\tau_w)$ is the probability of an idle slot, and $P_{w,\mathrm{busy}} = 1-\prod_{j=1}^{J}(1-\tau_{j,c,t})\prod_{w=1}^{W}(1-\tau_w)$ is the probability of a busy slot, regardless of whether it corresponds to a collision or a successful transmission. θ and T_b are, respectively, the average durations of an idle and a busy slot and, thus, the denominator in (3.2) corresponds to the mean duration of a WiFi MAC slot.

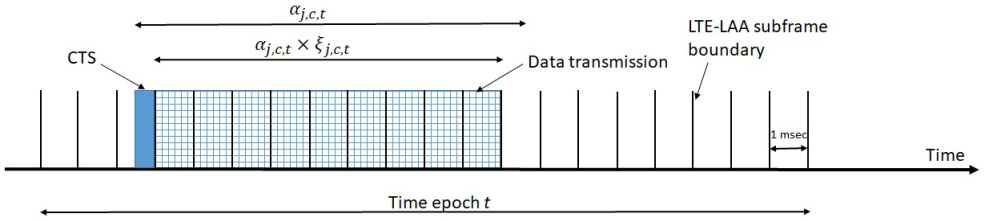

Figure 3.3. An illustrative example for computing the actual data transmission time of an LTE-LAA SBS.

The achievable airtime fraction for an LTE-LAA SBS j on channel c at time t is:

$$\alpha_{j,c,t} = \tau_{j,c,t} \prod_{i=1, i\neq j}^{J} (1 - \tau_{i,c,t}) \prod_{w=1}^{W} (1 - \tau_w). \tag{3.3}$$

The airtime fraction represents the time allocated for an SBS on channel c during time t which essentially accounts for both the data transmission time as well as the reservation signal overhead. Here, it is important to account for the reservation signal overhead during a transmission burst of an SBS when computing the throughput, as done in [143]. As such, we let $\xi_{j,c,t}$ be the average fraction of time of $\alpha_{j,c,t}$ during which LTE-LAA SBS is transmitting data. Fig. 3.3 provides an illustrative example for computing the fraction of time of t during which LTE-LAA SBS is transmitting data. Thus, the total throughput of all $K_{j,t}$ UEs that are served by SBS j during time epoch t is:

$$R_{j,t} = \sum_{c=1}^{C} \alpha_{j,c,t} \xi_{j,c,t} r_{j,c,t}, \tag{3.4}$$

where

$$r_{j,c,t} = \sum_{k=1}^{K_{j,t}} B_c \log\left(1 + \frac{P_{j,c,t} h_{j,k,c,t}}{I_{j,c,t} + B_c N_0}\right). \tag{3.5}$$

Here, $I_{j,c,t} = \sum_{w=1}^{W} P_{w,c,t} h_{w,k,c,t} + \sum_{i=1, i\neq j}^{J} P_{i,c,t} h_{i,k,c,t}$ is the interference level on SBS j when operating on channel c during time t and B_c is the BW of channel c. $P_{j,c,t}$ is the transmit power of SBS j on channel c during time t. $h_{j,k,c,t}$ is the channel gain between SBS j and UE k on channel c during time t. N_0 is the power spectral density of additive white Gaussian noise. Since SBSs and WAPs both adopt LBT, then one cell may occupy the entire channel at a given time thus transmitting *exclusively* on a given channel c. However, hidden and exposed

terminals could be present on a given channel which can result in interference and thus a degradation in the throughput.

Given this system model, next, we develop an effective spectrum allocation scheme that can allocate the appropriate unlicensed channels along with the corresponding channel access probabilities to each SBS simultaneously over T, at $t = 0$.

3.3 Proactive Resource Allocation Scheme for Unlicensed LTE

3.3.1 Proactive Resource Allocation Game

We formulate the resource allocation problem as a noncooperative game $\mathcal{G}$=($\mathcal{J}, \mathcal{A}_j, u_j$) with the SBSs in $\mathcal{J}$ being the players, each of which must choose a channel selection and channel access pair $a_{j,c,t}$ =($x_{j,c,t}$,$\alpha_{j,c,t}$) $\in \mathcal{A}_j$ at $t = 0$ for each t of the next time window T. The objective of each SBS j is to maximize its total throughput over the selected channels:

$$u_j(\boldsymbol{a}_j, \boldsymbol{a}_{-j}) = \sum_{t=1}^{T} \sum_{c=1}^{C} \alpha_{j,c,t} \xi_{j,c,t} r_{j,c,t}, \tag{3.6}$$

where $\boldsymbol{a}_j = [(a_{j,1,1}, \cdots, a_{j,1,T}), \cdots, (a_{j,C,1}, \cdots, a_{j,C,T})]$ and $\boldsymbol{a}_{-j}$ correspond, respectively to the action vector of SBS j and all other SBSs, over all the channels $\mathcal{C}$ during T. Note that the utility function (3.6) of SBS j depends on its actions as well as those of other SBSs which makes the formulation of a game model suitable for this problem. This is mainly due to the interference from other SBSs transmitting on the same channel as SBS j as it was shown previously in the definition of the rate expression in (3.5). The goal of each SBS j is to maximize its own utility:

$$\max_{a_j \in \mathcal{A}_j} u_j(\boldsymbol{a}_j, \boldsymbol{a}_{-j}) \ \ \forall j \in \mathcal{J}, \tag{3.7}$$

$$\text{s.t.} \qquad \alpha_{j,c,t} \leq x_{j,c,t} \quad \forall c, t, \tag{3.8}$$

$$\sum_{c=1}^{C} x_{j,c,t} \leq \min(M_c, C) \quad \forall t, \tag{3.9}$$

$$\sum_{t_T=1}^{t}\sum_{c=1}^{C}\alpha_{j,c,t_T}B_c \leq \sum_{t_T=1}^{t} f(L_{j,t_T}) \quad \forall t, \tag{3.10}$$

$$\alpha_{w,c,t} + \alpha_{j,c,t} + \sum_{i=1,i\neq j}^{J} \alpha_{i,c,t} \leq t_{\max} \quad \forall c,t, \tag{3.11}$$

$$x_{j,c,t} \in \{0,1\}, \quad \alpha_{j,c,t} \in [0,1] \quad \forall c,t, \tag{3.12}$$

where M_c is the total number of unlicensed channels which an SBS can aggregate. (3.8) allows the allocation of a channel access proportion for SBS j on channel c during t only if SBS j transmits on channel c at time t. (3.9) guarantees that each SBS can aggregate a maximum of M_c channels at a given time t. (3.10) limits the amount of allocated BW to the required demand where $f(L_{j,t})$ captures the relationship between bandwidth requirement and offered load. (3.11) captures coupling constraints which limit the proportion of time used by SBSs and WLAN on a given unlicensed band to the maximum fraction of time an unlicensed channel can be used, $t_{\max}$[1]. (3.12) represents the feasibility constraints.

Given the fact that different operators and technologies have equal priorities on the unlicensed spectrum, we incorporate the Homo Egualis (HE) anthropological model, an inequity-averse based fairness model, into the strategy design of the agents [144].

Definition 1. *Inequity aversion* is the preference for fairness and resistance to incidental inequalities. In other words, it refers to the willingness of giving up some material payoff in order to move in the direction of more equitable outcomes.

In an HE society, agents focus not only on maximizing their own payoffs, but also become aware of how their payoffs are compared to other agents' payoffs [144, 145]. Therefore, their utility function is influenced not only by their own reward, but also by envy and altruism. An agent is altruistic to others if its payoff is above an equitable benchmark and is envious of the others if its payoff exceeds that benchmark and therefore, an unfair distribution of resources among agents results in disutility for inequity-averse agents. The HE concept comes from the anthropological literature in which Homo sapiens evolved in small hunter-gatherer groups without a centralized governance [144].

[1] $t_{\max}$ depends on the channel access method in the unlicensed band and should be strictly less than 1 in the case of LBT, otherwise, the channel will always be sensed busy and devices would not be able to access it.

In fact, we incorporate the notion of airtime fairness in the modeling of our HE agents. The average airtime per radio system is considered as one of the most important fairness metrics in the unlicensed band and is the focus of this work [146]. Our motivation for considering a time-fair channel allocation scheme is to overcome the rate anomaly problem that arises when different nodes use distinct data rates, which leads to the slowest link limiting the system performance [134, 146], and [141]. Therefore, to model our players as HE agents, we consider the following two coupling constraints for the allocated airtime fraction on each channel c for each SBS j:

$$\frac{1}{w_{j,c}}\frac{1}{T}\frac{\sum_{t=1}^{T}\alpha_{j,c,t}}{\sum_{t=1}^{T}\bar{L}_{j,t}}=\frac{1}{w_{i,c}}\frac{1}{T}\frac{\sum_{t=1}^{T}\alpha_{i,c,t}}{\sum_{t=1}^{T}\bar{L}_{i,t}}\ \forall c\in\widehat{\mathcal{C}}_j, i\in\widehat{\mathcal{S}}_{j,c}(i\neq j), \tag{3.13}$$

$$\frac{1}{T}\frac{\sum_{t=1}^{T}\sum_{n\in\mathcal{S}_{c,t}}\alpha_{n,c,t}}{P_{\text{LTE}}\sum_{t=1}^{T}\sum_{n\in\mathcal{S}_{c,t}}\bar{L}_{n,t}}=\frac{1}{T}\frac{\sum_{t=1}^{T}\alpha_{w,c,t}}{P_{\text{WiFi}}\sum_{t=1}^{T}L_{w,c,t}}\ \ \forall c\in\widehat{\mathcal{C}}_j, \tag{3.14}$$

where $\widehat{\mathcal{C}}_j$ is the subset of channels used by SBS j during T. $\mathcal{S}_{c,t}$ is the subset of SBSs that are transmitting over channel c, $c\in\widehat{\mathcal{C}}_j$, during time t and $\widehat{\mathcal{S}}_{j,c}$ is the subset of other neighbouring SBSs, $i\neq j$, that are using the same channel $c\in\widehat{\mathcal{C}}_j$ as SBS j during T. $\bar{L}_{j,t}=L_{j,t}-\sum_{c'}f(\alpha_{j,c',t})$ corresponds to the remaining traffic that needs to be served by SBSs j with $L_{j,t}$ being the *total* aggregate traffic demand of SBS j on channel c during time epoch t. $f(.)$ corresponds to the served traffic load as a function of the airtime allocation. c' represents all the set of channels except channel c. $\alpha_{w,c,t}=\min(f(L_{w,c,t}),\ t_{\max}-\alpha_{j,c,t}-\sum_{i\in\mathcal{S}_{j,c,t}}\alpha_{i,c,t})$ is the airtime allocated for WLAN transmissions over channel c during time t. $w_{j,c}=\sum_{t=1}^{T}x_{j,c,t}$ is the weight of SBS j on channel c during T and thus different SBSs are assigned different weights on each channel c based on the number of time epochs t a given SBS j uses that particular channel. P_{WiFi} and P_{LTE} correspond to the priority metric defined for each technology when operating on the unlicensed band. These parameters allow adaptation of the level of fairness between LTE-LAA and WLAN.

Constraint (3.13) represents inter-operator fairness which guarantees an equal weighted airtime allocation among SBSs belonging to different operators on a given channel c. The adopted notion of fairness is based on a long-term weighted equality over T, as opposed to instantaneous weighted equality. (3.14) defines an inter-technology fairness metric to guarantee a long-term equal weighted airtime

allocation over T for both LTE-LAA and WiFi. Therefore, (3.13) and (3.14) reflect the inequity aversion property of the SBSs.

In fact, the optimal value of T, which corresponds to the time window size that allows the maximum achievable throughput for LTE-LAA as compared to the reactive approach, is dataset dependent. Next, we characterize the optimal value of T under a uniform traffic distribution.

Proposition 1. For a uniform traffic distribution, the optimal value of T is equal to 1.

Proof. Under a uniform demand model, the traffic load for each of SBS j and WAP w is an independent and identically distributed sequence of random variables which implies that all requests of the same UE are statistically indistinguishable over time. In our model, WAPs are passive in that their channel selection action is fixed and, thus, the activity on a given channel is characterized by the level of activity of WAPs operating on that channel. In that case, the WLAN traffic load on each channel also follows a uniform distribution. At the convergence point, (3.8)-(3.14) are satisfied and, hence, the average airtime allocated to the LTE-LAA network on channel c over the time window T will be:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{j\in\mathcal{S}_{c,t}}\alpha_{j,c,t} = \frac{P_{\text{LTE}}}{P_{\text{WiFi}}}\frac{\sum_{t=1}^{T}\sum_{j\in\mathcal{S}_{c,t}}\bar{L}_{j,t}}{\sum_{t=1}^{T}L_{w,c,t}}\frac{1}{T}\sum_{t=1}^{T}\alpha_{w,c,t} \quad \forall c\in\mathcal{C}, \tag{3.15}$$

However, for the case of uniform traffic demand, the channel selection vector over T is the same for each SBS because the network state is the same for every t in T. Moreover, if an SBS aggregates multiple channels, then its load on each channel is the same for each t in T. This implies that $\bar{L}_{j,t}$ for each SBS j is uniform over T and thus $\frac{\sum_{t=1}^{T}\sum_{j\in\mathcal{S}_{c,t}}\bar{L}_{j,t}}{\sum_{t=1}^{T}L_{w,c,t}} = \frac{\sum_{j\in\mathcal{S}_{c,t}}\bar{L}_{j,t}}{L_{w,c,t}}$. Consequently, (3.15) can be written as:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{j\in\mathcal{S}_{c,t}}\alpha_{j,c,t} = \frac{P_{\text{LTE}}}{P_{\text{WiFi}}}\frac{\sum_{j\in\mathcal{S}_{c,t}}\bar{L}_{j,t}}{L_{w,c,t}}\frac{1}{T}\sum_{t=1}^{T}\alpha_{w,c,t} \quad \forall c\in\mathcal{C}, \tag{3.16}$$

When $T = 1$, the airtime allocated to the LTE-LAA network on channel c will be:

$$\sum_{j\in\mathcal{S}_{c,t}}\alpha_{j,c,t} = \frac{P_{\text{LTE}}}{P_{\text{WiFi}}}\frac{\sum_{j\in\mathcal{S}_{c,t}}\bar{L}_{j,t}}{L_{w,c,t}}\alpha_{w,c,t} \quad \forall t, c\in\mathcal{C}, \tag{3.17}$$

Over a fixed time window T, the average airtime allocated to the LTE-LAA network on channel c can be written as:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{j\in\mathcal{S}_{c,t}}\alpha_{j,c,t} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{P_{\text{LTE}}}{P_{\text{WiFi}}}\frac{\sum_{j\in\mathcal{S}_{c,t}}\bar{L}_{j,t}}{L_{w,c,t}}\alpha_{w,c,t}\right) \tag{3.18}$$

$$= \frac{P_{\text{LTE}}}{P_{\text{WiFi}}}\frac{\sum_{j\in\mathcal{S}_{c,t}}\bar{L}_{j,t}}{L_{w,c,t}}\frac{1}{T}\sum_{t=1}^{T}\alpha_{w,c,t} \;\; \forall c\in\mathcal{C}. \tag{3.19}$$

(3.19) is equivalent to (3.16) and, hence, our proposed framework does not offer any gain for the LTE-LAA network when considering a time window T larger than 1 in the case of a uniform traffic pattern. This completes the proof. □

From Proposition 1, we conclude that the gain of our proposed long-term fairness notion is evident in the case of traffic fluctuations. Under a uniform traffic distribution, the SBSs cannot make use of future off-peak times to shift part of their traffic forward in time and, hence, the gain is limited to predicting the network state for the next time epoch only. It is also worth noting that the gain of the proactive scheduling approach decreases in the case of a highly congested WLAN network. This is mainly due to the fact that the system becomes more congested with incoming requests, thereby restricting the opportunities of shifting part of the LTE-LAA load in the future.

3.3.2 Equilibrium Analysis

Our game $\mathcal{G}$ is a generalized Nash equilibrium problem (GNEP) in which both the objective functions and the action spaces are coupled. To solve the GNEP, we incorporate the Lagrangian penalty method into the utility functions thus reducing it to a simpler Nash equilibrium problem. The resulting penalized utility function will be, $\forall(j\in\mathcal{J})$:

$$\begin{aligned}\widehat{u}_j(\boldsymbol{a}_j,\boldsymbol{a}_{-j}) = \sum_{t=1}^{T}\sum_{c=1}^{C}\alpha_{j,c,t}r_{j,c,t} &- \rho_{1,j}\sum_{c=1}^{C}\sum_{t=1}^{T}\left(\min(0, t_{\max}-\alpha_{w,c,t}-\alpha_{j,c,t}-\sum_{i=1,i\neq j}^{J}\alpha_{i,c,t})\right)^2 \\ &- \rho_{2,j}\sum_{c\in\widehat{\mathcal{C}}_j}\sum_{i\in\widehat{\mathcal{S}}_{j,c}(i\neq j)}\frac{1}{T^2}\left(\frac{1}{w_{j,c}}\frac{\sum_{t=1}^{T}\alpha_{j,c,t}}{\sum_{t=1}^{T}\bar{L}_{j,t}}-\frac{1}{w_{i,c}}\frac{\sum_{t=1}^{T}\alpha_{i,c,t}}{\sum_{t=1}^{T}\bar{L}_{i,t}}\right)^2 \\ &- \rho_{3,j}\sum_{c\in\widehat{\mathcal{C}}_j}\frac{1}{T^2}\left(\frac{\sum_{t=1}^{T}\sum_{n\in\mathcal{S}_{c,t}}\alpha_{n,c,t}}{P_{\text{LTE}}\sum_{t=1}^{T}\sum_{n\in\mathcal{S}_{c,t}}\bar{L}_{n,t}}-\frac{\sum_{t=1}^{T}\alpha_{w,c,t}}{P_{\text{WiFi}}\sum_{t=1}^{T}L_{w,c,t}}\right)^2,\end{aligned} \tag{3.20}$$

where $\rho_{1,j}$, $\rho_{2,j}$ and $\rho_{3,j}$ are positive penalty coefficients corresponding to constraints (3.11), (3.13), and (3.14), respectively. Here, we consider equal penalty coefficients for all players for each coupled constraint, $\rho_{1,j} = \rho_1$, $\rho_{2,j} = \rho_2$ and $\rho_{3,j} = \rho_3$. This allows all SBSs to have equal incentives to give up some payoff in order to satisfy the coupled constraints. To determine the values of ρ_1, ρ_2 and ρ_3, we adopt the incremental penalty algorithm in [147] that guarantees the existence of penalty parameters $\boldsymbol{\rho}_l^* = [\rho_1^*, \rho_2^*, \rho_3^*]$ that satisfy the coupled constraints.

In our game $\mathcal{G}$, $\alpha_{j,c,t}$ is a continuous variable bounded between 0 and 1, however, for a particular network state, we are interested only in a certain region of the continuous space where the optimal actions are expected to be. Therefore, we will propose a sampling-based approach to discretize $\alpha_{j,c,t}$ in Section 3.4. Under such a discretization of the action space, we turn our attention to mixed strategies in which players choose their strategies probabilistically. Such a mixed-strategy approach enables us to analyze the frequency with which players choose different channels and channel access combinations. In fact, the optimal policy is often stochastic and therefore requires the selection of different actions with specific probabilities [148]. This, in turn, validates our choice of adopting a mixed strategy approach as opposed to a pure strategy one that is oriented towards finding deterministic policies. A player can possibly choose different possible actions with different probabilities which enables it to play a combination of strategies over time. Moreover, unlike pure strategies that might not exist for a particular game, there always exists at least one equilibrium in mixed strategies [149].

Let $\Delta(\mathcal{A})$ be the set of all probability distributions over the action space $\mathcal{A}$ and $\boldsymbol{p}_j = [p_{j,\boldsymbol{a}_1} \cdots, p_{j,\boldsymbol{a}_{|\mathcal{A}_j|}}]$ be a probability distribution with which SBS j selects a particular action from $\mathcal{A}_j$. Therefore, our objective is to maximize the expected value of the utility function, $\overline{u}_j(\boldsymbol{p}_j, \boldsymbol{p}_{-j}) = \mathbb{E}_{\boldsymbol{p}_j}\left[\widehat{u}_j\left(\boldsymbol{a}_j, \boldsymbol{a}_{-j}\right)\right] = \sum_{\boldsymbol{a}\in\mathcal{A}} \widehat{u}_j(\boldsymbol{a}_j, \boldsymbol{a}_{-j}) \prod_{j=1}^{J} p_{j,\boldsymbol{a}_j}$, where $\widehat{u}_j(\boldsymbol{a}_j, \boldsymbol{a}_{-j})$ is given in (3.20).

Definition 2. A mixed strategy $\boldsymbol{p}^*=(\boldsymbol{p}_1^*, \cdots, \boldsymbol{p}_J^*)=(\boldsymbol{p}_j^*, \boldsymbol{p}_{-j}^*)$ constitutes a *mixed-strategy Nash equilibrium* if, $\forall j \in \mathcal{J}$ and $\forall \boldsymbol{p}_j \in \Delta(\mathcal{A}_j)$, $\overline{u}_j(\boldsymbol{p}_j^*, \boldsymbol{p}_{-j}^*) \geq \overline{u}_j(\boldsymbol{p}_j, \boldsymbol{p}_{-j}^*)$.

Here, we note that any finite noncooperative game will admit at least one mixed-strategy NE [149]. To solve for the mixed-strategy NE of our game $\mathcal{G}$, we first consider the simpler scenario in which the number of SBSs is less than the number of unlicensed channels. Then, we develop a learning algorithm to handle the more realistic scenario in which the number of SBSs is much larger than the

number of unlicensed channels.

Remark 1. If the number of SBSs is less than the number of available unlicensed channels (i.e., $J \leq C$), then the mixed-strategy NE solution will simply reduce to a pure strategy that is reached when all SBSs occupy disjoint channels during each time epoch of the time window T.

To show this, we consider two cases depending on whether or not CA is enabled. Let $M_c = 1$. Consider the state in which each SBS is operating on a different unlicensed channel. If SBS j changes its channel from c to c' on which SBS i is transmitting, then it would have to share channel c' with SBS i in an equal weighted manner (based on the inter-operator fairness constraint). This leads to a decrease in the reward function of SBS i on channel c' (and potentially for SBS j), which makes SBS i deviate to another channel that is less occupied (e.g., c). Therefore, a given strategy cannot be a BR strategy for SBS i in case it results in its transmission on the same channel as SBS j. Therefore, all strategies that result in more than one SBS occupying the same channel are dominated by the alternative where different SBSs transmit on disjoint channels and hence cannot correspond to BR strategies. Consequently, at the NE point, all SBSs play their BR strategies that would result in each SBS occupying a disjoint channel. Similarly for $M_c > 1$. If SBS j transmits on multiple channels, then aggregating a channel that is already occupied by SBS i would make SBS i change its operating channel to a less congested one. This implies that an SBS would not aggregate more channels unless they are not occupied by other SBSs.

Therefore, we can conclude that our proposed scheme results in having fewer SBSs on each of the unlicensed bands. This leads to a lower collision probability on each channel and a better coexistence with WLAN. Moreover, enabling CA does not necessarily allow LTE to offload more traffic to the unlicensed band. On the other hand, our proposed scheme can avoid causing performance degradation to WLAN in case a large number of LTE operators deploy LTE-LAA in the unlicensed bands.

Now, when $J > C$, multiple SBSs will then potentially have to share the same channel. In this case, the mixed-strategy NE is challenging to characterize, and therefore, next, we propose a learning-based approach for solving our game $\mathcal{G}$. Given the fact that each SBS needs to learn a *sequence* of actions over the time window T at $t = 0$ based on a *sequence* of previous network states, the

proposed learning algorithm must be capable of generating data that is sequential in nature. This necessitates the knowledge of historical traffic values as well as future network states for all the time epochs of the following time window T. Moreover, in order to satisfy the long-term fairness constraints (3.13) and (3.14), future actions cannot be assumed to be independent due to the long-term temporal dependence among these actions. Conventional RL algorithms such as Q-learning and multi-armed bandit take as an input the current state of the network and enable the prediction of the next state only and therefore do not account for the interdependence of future actions [150]. To learn several steps ahead in time, recursive learning can be adopted. However, such an approach uses values already predicted, instead of measured past values which produces an accumulation of errors that may grow very fast. In contrast, deep learning techniques, such as time series prediction algorithms, are capable of learning long-term temporal dependence sequences based on input sequences [19]. This is viable due to their adaptive memory that allows them to store necessary previous state information to predict future events. Therefore, next, we develop a novel time series prediction algorithm based on deep learning techniques for solving the mixed-strategy NE of our game.

3.4 RL-LSTM for Self-organizing Resource Allocation

The proposed game requires each SBS to learn a sequence of actions over the prediction time window T, at $t = 0$, without any knowledge of future network states. This necessitates a learning approach with memory for storing previous states whenever needed while being able to learn a sequence of future network states. Employing LSTMs is therefore an obvious choice for learning as they are capable of generating data that is sequential in nature [19]. Consequently, we propose a novel sequence level training algorithm based on RL-LSTM that allows SBSs to learn a sequence of future actions at operation time based on a sequence of historic traffic load thus maximizing the sum of their future rewards.

LSTMs are a special kind of "deep" RNN capable of storing information for long periods of time to learn the long-term dependency within a sequence [151]. LSTMs process a variable-length sequence $\boldsymbol{y} = (y_1, y_2, ..., y_m)$ by incrementally

adding new content into a single memory slot, with gates controlling the extent to which new content should be memorized, old content should be erased, and current content should be exposed. Unlike conventional one-step RL techniques (e.g., Q-learning), LSTM networks are capable of predicting a sequence of future actions [19]. Predictions at a given time step are influenced by the network activations at previous time steps thus making LSTMs suitable for our application. The total number of parameters W in a standard LSTM network with one cell in each memory block is given by:

$$W = n_c \times n_c \times 4 + n_i \times n_c \times 4 + n_c \times n_o + n_c \times 3 \tag{3.21}$$

where n_c is the number of memory cells, n_i is the number of input units, and n_o is the number of output units. The computational complexity of learning LSTM models per weight and time step is linear i.e., $O(1)$. Therefore, the learning computational complexity per time step is $O(W)$ [152].

Consequently, we consider an end-to-end RL-LSTM based approach to train the network to find a mixed-strategy NE of the game $\mathcal{G}$. LSTMs have three types of layers, one input and one output layer as well as a varying number of hidden layers depending on the dataset under study. For our dataset, adding more hidden layers does not improve performance and thus one layer is sufficient. Moreover, in order to allow a sequence to sequence mapping, we consider an encoder-decoder model. The encoder network takes the input sequence and maps it to a vector of a fixed dimensionality. The encoded representation is then used by the decoder network to decode the target sequence from the vector. Fig. 3.4 summarizes the proposed approach. The traffic encoder takes as an input the historical traffic loads and learns a vector representation of the input time-series. The multi-layer perceptron (MLP) summarizes the input vectors into one vector. In our scheme, an MLP is required to encode all the vectors together since a particular action at time t depends on the values of all other input vectors (i.e., traffic values of all SBSs and WLAN on all the unlicensed channels). The action decoder takes as an input the summarized vector to reconstruct the predicted action sequence. All SBSs have the same input vector for the traffic encoders and thus they share the same traffic encoders. On the other hand, SBSs learn different action sequences and thus different SBSs use different action decoders.

In the first step, we need to train the neural networks in order to learn the parameters of the algorithm that would maximize the proposed utility function.

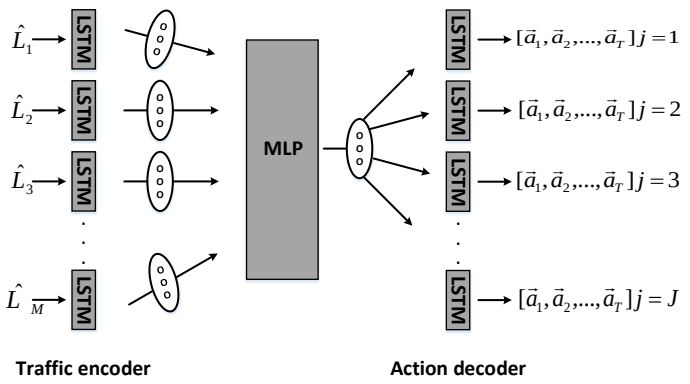

Figure 3.4. Proposed framework.

Therefore, the proposed algorithm is divided into *two phases, the training phase followed by the testing phase*. In the former, SBSs are trained offline before they become active in the network using the architecture given in Fig. 3.4. The input dataset represents the WiFi traffic load distribution on the unlicensed channels as well as the SBSs traffic load collected over several days. On the other hand, the testing phase corresponds to the actual execution of the algorithm after which the parameters have been optimized and is implemented on each SBS for execution during run time.

For the training phase, we train the weights of our neural network using a policy gradient approach that aims at maximizing the expected return of a policy. This is achieved by representing the policy by its own function approximator and updating it according to the gradient of the expected reward with respect to the policy parameters [148]. Consider the set $\mathcal{M}$ of M history traffic sequences corresponding to either an SBS or WiFi on each unlicensed channel, where $M = J + C$. Let $\boldsymbol{h}_{m,t} \in \mathbb{R}^n$ and $\boldsymbol{h}_{j,t} \in \mathbb{R}^n$ be, respectively, the hidden vectors of the traffic encoder m and action decoder of SBS j at time t. $\boldsymbol{h}_{m,t}$ and $\boldsymbol{h}_{j,t}$ are then computed by:

$$\boldsymbol{h}_{m,t}=\phi\left(\boldsymbol{v}_{m,t}, \boldsymbol{h}_{m,t-1}\right), \quad \boldsymbol{h}_{j,t}=\phi\left(\boldsymbol{v}_{j,t}, \boldsymbol{h}_{j,t-1}\right), \tag{3.22}$$

where ϕ refers to the LSTM cell function [151] being used, and $\boldsymbol{v}_{m,t}$ is the input vector. For the encoder, $\boldsymbol{v}_{m,t} = \left[\widehat{L}_{m,t}\right]$ is the history traffic value. For the decoder, $\boldsymbol{v}_{j,t} = [\boldsymbol{W}_d \boldsymbol{e}(\boldsymbol{x}_{j,t-1})||\alpha_{j,c,t-1}]$ is the vector of the previous predicted action where $\boldsymbol{e}()$ maps discrete value to a one-hot vector, $\boldsymbol{W}_d \in \mathbb{R}^{n \times N_x}$ is a matrix that is used to transform the discrete actions of each of the unlicensed channels into a vector, and N_x is the number of discrete actions. In our approach, we learn the channel selection vector for all the channels simultaneously and thus $\boldsymbol{x}_{j,t} =$

$[x_{j,1,t}, \cdots, x_{j,C,t}]$.

To learn the mixed strategy of our proposed game, we need to initialize the action space with a subset of the continuous action space of $\alpha_{j,c,t}$. A naive approach for working with continuous action spaces is to discretize the action space; however, this approach would lead to combinatorial explosion and thus the well known problem of "curse of dimensionality" when highly discretizing our space and a loss in the accuracy of the predicted action when considering less discretized values. Therefore, we consider a sampling-based approach where we first define a probability distribution for the continuous variable $\alpha_{j,c,t}$ and for the discrete variable $x_{j,c,t}$ in order to deal with the large discrete action space as T increases. We use a softmax classifier to predict the distribution for the discrete variable $\boldsymbol{x}_{j,t}$ and a Gaussian policy for the distribution of the continuous variable $\alpha_{j,c,t}$. For the Gaussian policy, the probability of an action is proportional to a Gaussian distribution with a parameterized mean and a fixed value for the variance in our implementation. The variance of the Gaussian distribution defines the area around the mean from which we explore the action space. For our implementation, the initial value of the variance is set to 0.06 in order to increase exploration and then is decreased linearly towards 0.02. Therefore, defining probability distributions for our variables allows the initialization of the action space $\mathcal{A}_j$ by sampling Z actions from the proposed distributions. This enables the SBSs to learn more accurate transmission probabilities for $\alpha_{j,c,t}$, as opposed to fixed discretization, thus satisfying the fairness constraints. The hidden vector $\boldsymbol{h}_{j,t}$ in the decoder is used to predict the t-th output actions $\boldsymbol{x}_{j,t}$ and $\alpha_{j,c,t}$. The probability vector over $\boldsymbol{x}_{j,t}$ and $\alpha_{j,c,t}$ can be defined, respectively, as:

$$\boldsymbol{x}_{j,t}|\boldsymbol{x}_{j,<t}, \alpha_{j,c,<t}, \widehat{\boldsymbol{L}}_t \sim \sigma\left(\boldsymbol{W}_x \boldsymbol{h}_{j,t}\right), \tag{3.23}$$

$$\mu_{j,c,t} = S\left(\boldsymbol{W}_\mu \boldsymbol{h}_{j,t}\right), \;\; \alpha_{j,c,t} \sim \mathcal{N}(\mu_{j,c,t}, \mathrm{Var}(\alpha_{j,c,t})), \tag{3.24}$$

where $\mu_{j,c,t}$ and $\mathrm{Var}(\alpha_{j,c,t})$ correspond to the mean value and variance of the Gaussian policy respectively, $\boldsymbol{W}_x \in \mathbb{R}^{|V_a| \times n}, \boldsymbol{W}_\mu \in \mathbb{R}^n$ are parameters, $\sigma(.)$ is the softmax function $\sigma(\boldsymbol{b})_q = \frac{e^{b_q}}{\sum_{o=1}^{O} e^{b_o}}$ for $q = 1, \cdots, O$, and $S(.)$ is the sigmoid function where $S(b) = \frac{1}{1+e^{-b}}$ and is used to normalize the value to $(0, 1)$. $\alpha_{j,c,t}$ is computed only when $x_{j,c,t} = 1$. The probability of the whole action sequence for SBS j, given a historic traffic sequence $\widehat{\boldsymbol{L}}$, $p_{j,\boldsymbol{a}_j|\widehat{\boldsymbol{L}}}$, is given by:

$$p_{j,\boldsymbol{a}_j|\widehat{\boldsymbol{L}}} = \prod_{t=1}^{T} p\left((\boldsymbol{x}_{j,t}, \alpha_{j,c,t})|\boldsymbol{x}_{j,<t}, \alpha_{j,c,<t}, \widehat{\boldsymbol{L}}_t\right), \tag{3.25}$$

where $\widehat{\boldsymbol{L}}_t$=$(\widehat{L}_{1,t}, \cdots, \widehat{L}_{M,t})$, $\boldsymbol{x}_{j,<t}$=$[\boldsymbol{x}_{j,1}, \cdots, \boldsymbol{x}_{j,t-1}]$, and $\mu_{j,c,<t}$=$[\mu_{j,c,1}, \cdots, \mu_{j,c,t-1}]$.

Our goal is to maximize the exact expectation of the reward $\widehat{u}_j(\boldsymbol{a}_j, \boldsymbol{a}_{-j})$ over the action space for the training dataset. Therefore, the objective function can be defined as:

$$\max_{a_j \in \mathcal{A}_j} \sum_{\mathcal{D}} \overline{u}_j(\boldsymbol{p}_j, \boldsymbol{p}_{-j}), \tag{3.26}$$

where $\mathcal{D}$ is the training dataset. For this objective function, the REINFORCE algorithm [153] can be used to compute the gradient of the expected reward with respect to the policy parameters, and then standard gradient descent optimization algorithms [148] can be adopted to allow the model to generate optimal action sequences for input history traffic values. Specifically, Monte Carlo sampling is adopted to compute the expectation.

In particular, we adopt the RMSprop gradient descent optimization algorithm for the update rule [117]. The learning rate of a particular weight is divided by a running average of the magnitudes of recent gradients for that weight. The RMSprop update rule is given by:

$$\mathbb{E}[g^2]_t = \gamma \mathbb{E}[g^2]_{t-1} + (1-\gamma) g_t^2, \tag{3.27}$$

$$\theta_{t+1} = \theta_t - \frac{\lambda}{\sqrt{\mathbb{E}[g^2]_t + \epsilon}} g_t, \tag{3.28}$$

where θ_t corresponds to the model parameters at time t, g_t is the gradient of the objective function with respect to the parameter θ at time step t, $\mathbb{E}[g^2]_t$ is the expected value of the magnitudes of recent gradients, γ is the discount factor, λ is the learning rate and ϵ is a smoothing parameter. It is important to note here that the more complex the network is, the longer training phase. In particular, as the number of SBSs increases, the time required for training the proposed architecture increases. For a network size less than 15 SBSs, the training time of the proposed framework is in the order of hours.

Meanwhile, the testing phase corresponds to the actual execution of the algorithm on each SBS. Based on historical traffic values, each SBS learns the future sequence of actions based on the learned parameters from the training phase. For practicality, we assume knowledge of historical measurements of the WiFi activity on each of the unlicensed channels using simple network management protocol

Algorithm 1: Training phase of the proposed approach.

Input: $\mathcal{J}; \mathcal{W}; \mathcal{C}; \widehat{L}_{j,t} \forall j \in \mathcal{J}, t; \widehat{L}_{w,c,t} \forall c \in \mathcal{C}, t$.
***Initialization*:** The weights of all LSTMs are initialized following a uniform distribution with arbitrarily small values.
***Training*:** Each SBS j is modelled as an LSTM network.
while Any of the coupled constraints is not satisfied **do**
  for Number of training epochs **do**
    for Size of the training dataset **do**
      Step 1. Run Algorithm 2 to compute the best actions for all SBSs.
      for j=1:J **do**
        Step 2. Sample actions for SBS j based on the best expected actions of other SBSs.
        Step 3. Use REINFORCE [153] to update rule and compute the gradient of the expected value of the reward function.
        Step 4. Update model parameters with back-propagation algorithm [154].
      end for
    end for
  end for
  Step 5. Using the incremental penalty algorithm, check the feasibility of the coupled constraints and update the values of $\boldsymbol{\rho}_l$ accordingly.
end while

Algorithm 2: Testing phase of the proposed approach.

Input: $\mathcal{J}; \mathcal{W}; \mathcal{C}; \widehat{L}_{j,t} \forall j \in \mathcal{J}, t; \widehat{L}_{w,c,t} \forall c \in \mathcal{C}, t$.
for For each SBS j **do**
  Step 1. ***Traffic history encoding*:** The history traffic of each SBS and WLAN activity on each channel is fed into each of the M LSTM traffic encoders.
  Step 2. ***Vector summarization*:** The encoded vectors are transformed to initialize action decoders.
  Step 3. ***Action decoding*:** Action sequence is decoded for each SBS j.
end for

statistics with accurate calibration [155] and of other SBSs by exchanging past traffic information via the X2 interface as done in [25] and [15]. For our proposed scheme, the SBSs are trained over a large training dataset taking into account the traffic load over multiple days. The likeliness that an error occurs at the same time over multiple days is thus very rare. Moreover, our proposed scheme takes into account a *sequence* of history traffic values. Therefore, in case of non-ideal information, the impact of this error can be considered to be negligible. The proposed approach can also be combined with online machine learning [156] to accommodate changes in the traffic model, by properly re-training the developed learning mechanism. Consequently, the proposed algorithm offers a practical solution that is amenable to implementation. Here, we note that one practical challenge for deploying this algorithm in a real-world network is synchronization between SBSs and WAPs. In essence, such synchronization can be achieved by inter-operator cooperation, using mechanisms such as in [40]. The training and

the testing phases are given in Algorithms 1 and 2 respectively.

Note that guaranteeing the convergence of the proposed algorithm is challenging as it is highly dependent on the hyperparameters used during the training phase. It has been shown in [99] that the learning rate and the hidden layer size are the two most important hyperparameters for the convergence of LSTMs. For instance, using too few neurons in the hidden layers results in underfitting which could make it hard for the neural network to detect the signals in a complicated data set. On the other hand, using too many neurons in the hidden layers can result in either overfitting [157] or an increase in the training time. Therefore, in this work, we limit our contribution to providing simulation results (see Section 3.5) to show that, under a reasonable choice of the hyperparameters, convergence is observed for our proposed game, as per the following theorem:

Theorem 1. If Algorithm 1 converges, then the convergence strategy profile corresponds to a mixed-strategy NE of game $\mathcal{G}$.

Proof. In order to prove this theorem, we first need to show that the solution of the adopted multi-agent learning algorithm converges to an equilibrium point. In fact, every strict NE is a local optimum for a gradient descent learning approach but the reverse is not always true (Theorems 2 and 3 in [158]). Therefore, to show that a gradient-based learning method guarantees convergence of our proposed game to an equilibrium point, we define the following lemma.

Lemma 1. *The square of a linear function is convex. It follows that the payoff function of player j defined in (3.20) is an affine combination of convex functions, and hence is convex. Therefore, a gradient-based learning algorithm for our game $\mathcal{G}$ allows the convergence to an equilibrium point of that game.*

Lemma 1 is the consequence of the convexity of the players' payoffs where it has been shown in [159] that under certain convexity assumptions about the shape of payoff functions, the gradient-descent process converges to an equilibrium point. However, convergence is only guaranteed under a decreasing step-size sequence [160]. Therefore, given the fact that we employ an adaptive learning rate method satisfying the Robbins-Monro conditions ($\lambda > 0, \sum_{t=0}^{\infty} \lambda(t) = +\infty, \sum_{t=0}^{\infty} \lambda^2(t) < +\infty$), one can guarantee that under suitable initial conditions, our proposed algorithm converges to an equilibrium point.

Moreover, following the penalized reformulation of our game $\mathcal{G}$, one can easily show that a strategy that violates the coupled constraints cannot be a BR

Table 3.2 System parameters.

Parameters	**Values**	**Parameters**	**Values**
Transmit power (P_t)	20 dBm	BW (channel)	20 MHz
CCA threshold	-80 dBm	Noise variance	92 dBm/Hz
Path loss	$15.3 + 50\log_{10}(m)$	SIFS	16 μs
Hidden size (encoder)	70	DIFS	34 μs
Hidden size (decoder)	70	$\text{CW}_{\min}$	15 slots
time epoch (t)	5 min	$\text{CW}_{\max}$	1023 slots
Action sampling (Z)	100 samples	ACK	256 bits
History traffic size	7 time epochs	$P_{\text{LTE}}, P_{\text{WiFi}}$	1, 1
Learning rate (λ)	0.01	LSTM layers	1
Learning rate decay (γ)	0.95	$t_{\max}$	0.9

strategy. From [147], there exists $\boldsymbol{\rho}_l^*$ such that the incremental penalty algorithm terminates. Therefore, there exists a mixed strategy for which the coupled constraints are satisfied at $\boldsymbol{\rho}_l^*$. In that case, there is no incentive for an SBS to violate any of the coupled constraints, otherwise, its reward function would be penalized by the corresponding penalty function. Hence, all strategies that violate the coupled constraints are dominated by the alternative of complying with these constraints. Since in the proposed algorithm, the optimal strategy profile results in maximizing $\mathbb{E}_{\boldsymbol{p}_j}\left[\widehat{u}_j\left(\boldsymbol{a}_j, \boldsymbol{a}_{-j}\right)\right]$, we can conclude that the converged mixed-strategy NE is guaranteed not to violate the coupled constraints and hence it corresponds to a mixed-strategy NE for the game $\mathcal{G}$. Therefore, our proposed learning algorithm learns a mixed strategy of the game $\mathcal{G}$, by using a deep neural network function approximator to represent strategies, and by averaging those strategies via gradient descent machine learning techniques. □

3.5 Simulation Results and Analysis

For our simulations, we consider a 300 m × 300 m square area in which we randomly deploy a number of SBSs and 7 WAPs that share 7 unlicensed channels. We use real data for traffic loads from the dataset provided in [161] and divide it as 80% for training and 20% for testing. During the training phase, we randomly shuffle examples in the training dataset in order to prevent cycles when approximating the reward function. Table 3.2 summarizes the main simulation parameters. All statistical results are averaged over a large number of independent runs.

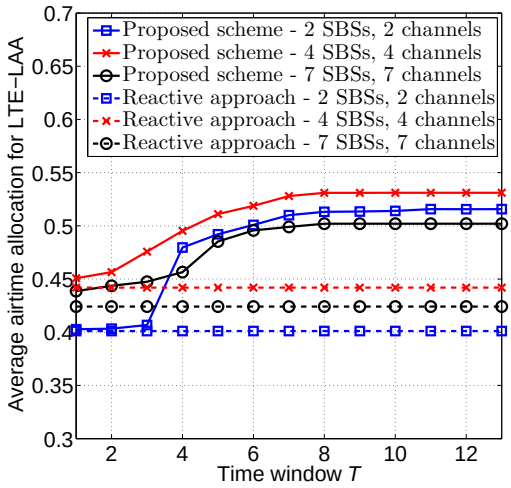

Figure 3.5. The average throughput gain for LTE-LAA upon applying a proactive approach (with varying T) as compared to a reactive approach.

Fig. 3.5 shows the average throughput gain, compared to a reactive approach, achieved by the proposed approach for different values of T under three different network scenarios. Here, we note that, in Fig. 3.5, the case in which $T = 1$ corresponds to other proactive schemes such as exponential smoothing and conventional RL algorithms (e.g., Q-learning and multi-armed bandit) [150]. Intuitively, a larger T provides the framework additional opportunities to benefit over the reactive approach, which does not account for future traffic loads. First, evidently, for very small time windows, our proposed approach does not yield any significant gains. However, as T increases, LTE-LAA network utilizes statistical predictions for allocating resources and thus the gains start to become more pronounced as compared to the reactive approach as well as to other proactive approaches at $T = 1$. For example, from Fig. 3.5, we can see that, for 4 SBSs and 4 channels, our proposed scheme achieves an increase of 17% and 20% in the average airtime allocation for LTE-LAA as compared to other proactive schemes and the reactive approach, respectively. Eventually, as T grows, the gain of our proposed framework remains almost constant at the maximum achievable value. This corresponds to the minimum value of T required to allow the LTE-LAA network smooth out its load over time and thus achieve maximum gain while guaranteeing fairness to WLAN.

Fig. 3.6 shows the proportion of LTE-LAA served load for different values of T. Clearly, as T increases, the proportion of LTE-LAA served traffic increases. For example, the proportion of served load increases from 82% to 97% for the case of 4

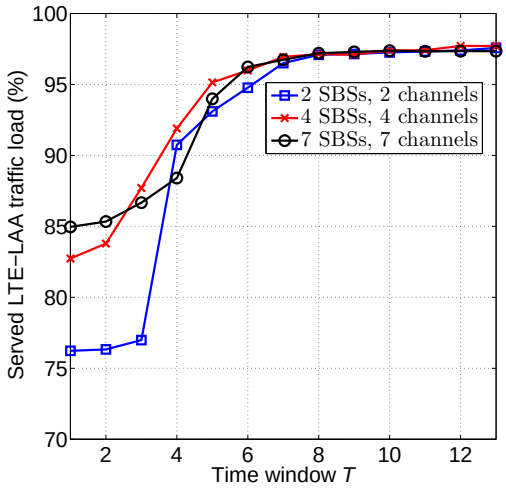

Figure 3.6. The proportion of load served over LTE-LAA as a function of T.

SBSs and 4 channels. The gain of the LTE-LAA network stems from the flexibility of choosing actions over a large time horizon T. In contrast to the myopic reactive approach, our proposed proactive scheme takes into account future predictions of the network state along with the current state. Therefore, the optimal policy will balance the instantaneous reward and the available information for future use and thus maximizing the total load served over time.

Fig. 3.7 shows the (a) average airtime allocated for the LTE-LAA network, (b) average airtime allocated for WLAN, (c) proportion of served total network traffic load, and (d) Jain's fairness index as a function of T resulting from our proposed scheme as well as from a centralized solution considering a proportional fairness utility function that is widely used for resource allocation [162], subject to constraints (3.8)-(3.12) with $T = 1$. Here, we compute the Jain's index based on the proportion of served traffic load for each network using $\mathcal{J}(l_o) = \frac{(\sum_{o=1}^{O} l_o)^2}{O \cdot \sum_{o=1}^{O} l_o^2}$, where l_o is the proportion of served traffic load for network o and O is number of networks [163]. The centralized solution of the PF resource allocation is obtained using the branch-and-bound algorithm in [164]. From Fig. 3.7 (a), we can see that for small values of T, the PF allocation offers higher airtime allocation for the LTE-LAA network. For example, for the scenario of 4 SBSs and 4 channels, PF offers airtime gains of 7% and 5% as compared to our proposed approach for $T = 1$ and 2 respectively. However, as T increases, our proposed scheme achieves more transmission opportunities for the LTE-LAA network as compared to the PF solution. For instance, for the scenario of 2 SBSs and 2 channels, our proposed scheme achieves an increase of 11% in the transmission opportunities for $T \geq 8$. This gain stems from the proactive resource allocation approach that allows more

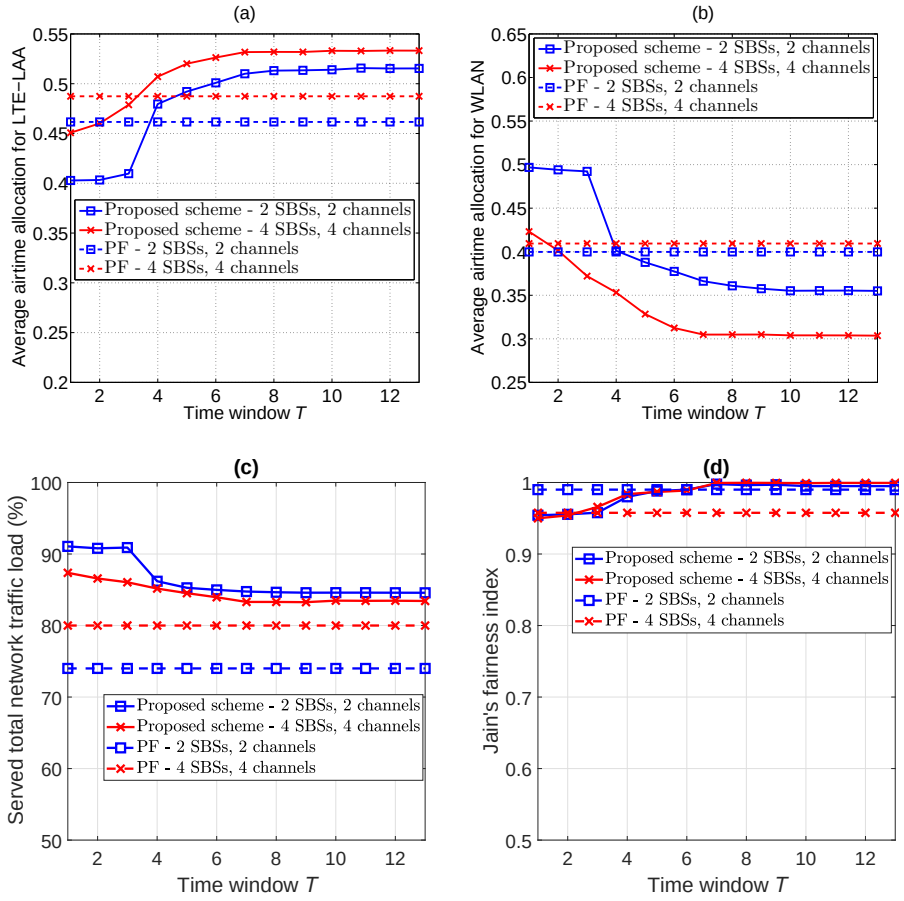

Figure 3.7. The (a) average airtime allocated for LTE-LAA, (b) average airtime allocated for WLAN, (c) proportion of served total network traffic load, and (d) Jain's fairness index resulting from our proposed scheme as well as from a centralized proportional fairness utility maximization scheme (with varying T).

flexibility in spectrum allocation as T increases. From Figs. 3.7 (b) and (c), we can see that, although the average airtime allocation for WiFi resulting from our proposed scheme is less than that of the PF scheme for $T > 4$, the proportion of the total network traffic load served by our proposed scheme is higher than that of the PF scheme for all values of T (e.g., 84% for our proposed scheme as compared to 74% for PF for the case of 2 SBSs and 2 channels and for $T > 6$). Moreover, from Fig. 3.7 (d), we can conclude that, as T increases, our proposed scheme achieves similar fairness performance as that of PF. This is due to the fact that, for our proposed scheme, as T increases, the proportion of LTE served traffic load increases while that of WiFi decreases eventually, converging to a constant value for $T > 7$. In particular, a relatively large time window allows SBSs to exploit future off-peak hours on the unlicensed band and thus increasing their transmission opportunities. Therefore, at the convergence point, the proportion

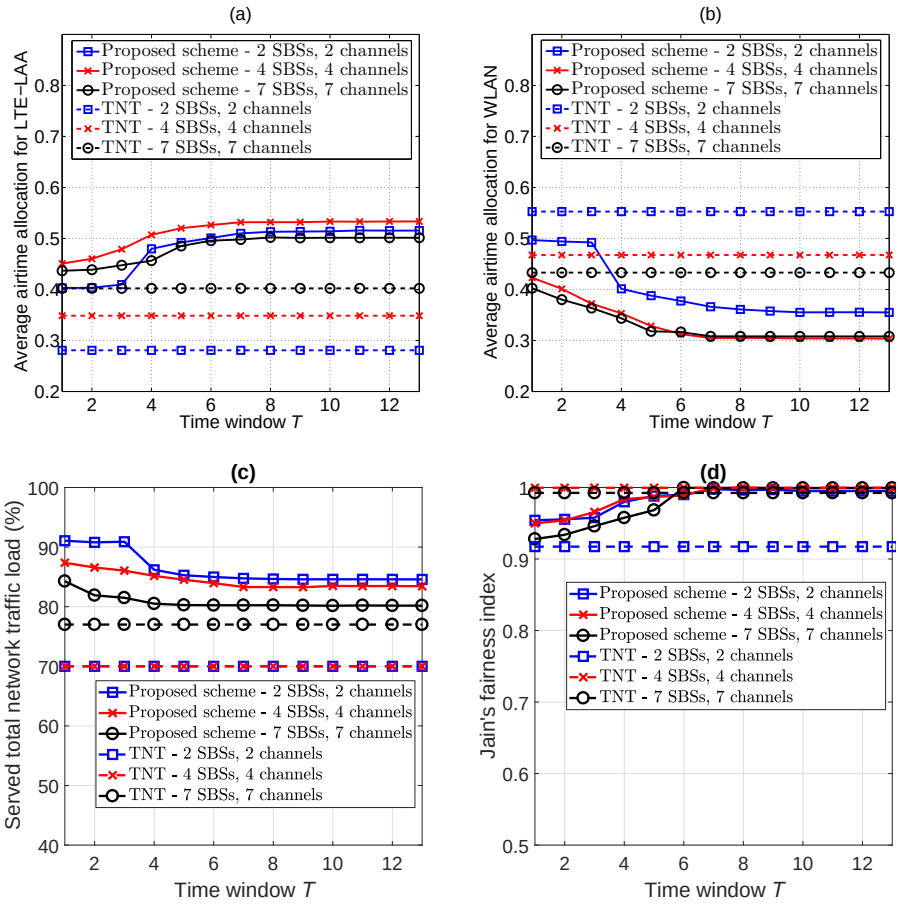

Figure 3.8. The (a) average airtime allocated for LTE-LAA, (b) average airtime allocated for WLAN, (c) proportion of served total network traffic load, and (d) Jain's fairness index resulting from our proposed scheme as well as from a centralized total network throughput utility maximization scheme (with varying T).

of served traffic load of both technologies is almost the same. In summary, our proposed scheme allows more transmission opportunities for LTE-LAA, increases the proportion of the total network served load while also preserving fairness with WiFi. It offers better tradeoff in terms of efficiency and fairness compared to the centralized PF allocation scheme. Note that the resulting problem for the PF solution is a mixed integer nonlinear optimization problem (MINLP) and therefore, finding its solution becomes challenging for larger network scenarios due to the polynomial computational complexity.

Fig. 3.8 shows the (a) average airtime allocated for the LTE-LAA network, (b) average airtime allocated for WLAN, (c) proportion of served total network traffic load, and (d) Jain's fairness index as a function of T resulting from our proposed scheme as well as a centralized solution considering a total network throughput (TNT) utility function subject to constraints (3.8)-(3.12) with $T = 1$. From

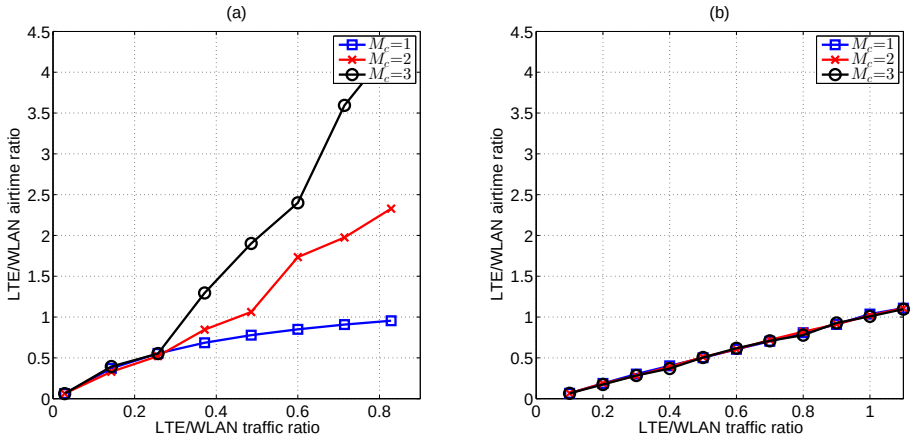

Figure 3.9. LTE/WLAN airtime ratio as a function of the LTE/WLAN traffic ratio for 3 different values of M_c (M_c = 1, 2 and 3). The LTE and WLAN airtime fraction correspond to the average airtime allocated per SBS and per WAP, respectively. Moreover, the number of unlicensed channels is fixed to 7 and the number of SBSs is equal to 2 and 7 in (a) and (b) respectively.

Fig. 3.8 (a), we can see that our proposed resource allocation scheme offers higher transmission opportunities for LTE-LAA for all values of T as compared to the centralized solution considering a TNT utility function. For example, for the case of 4 SBSs and 4 channels, the gain for our proposed approach can reach up to 52% for $T \geq 8$. Similarly, from Figs. 3.8 (b) and (c), we can observe that, although the average airtime allocation for WLAN for our proposed scheme is less than that of the TNT scheme for all T, the proportion of the total served network traffic load for our proposed scheme is higher than that of the TNT scheme. From Fig. 3.8 (d), we can also conclude that, as T increases, our scheme achieves similar fairness to TNT due to the fact that, as T increases, the proportion of LTE served traffic load increases while that of WiFi decreases for our proposed scheme, converging to a constant value for $T > 7$. At this convergence point, the proportion of served traffic load of both technologies is almost the same. In summary, our proposed scheme offers a better tradeoff in terms of efficiency and fairness as compared to the centralized TNT allocation scheme.

Fig. 3.9 shows the value of the LTE/WLAN airtime ratio under varying LTE/WLAN traffic ratio and for different values of M_c. Note here that the LTE and WLAN airtime fraction correspond to the average airtime allocated per SBS and per WAP, respectively. We consider two different scenarios with varying number of SBSs (2 and 7 SBSs for scenarios (a) and (b) respectively), while the

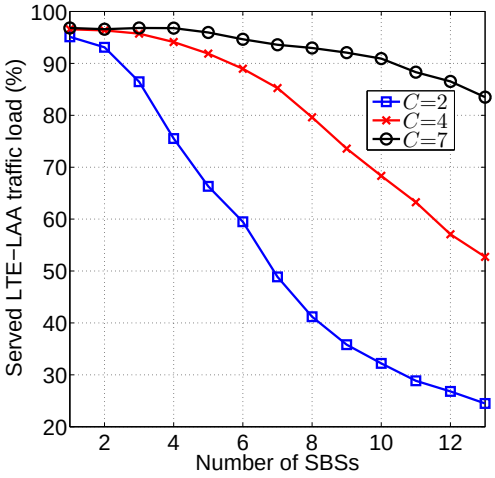

Figure 3.10. The proportion of LTE-LAA served traffic load as a function of the number of SBSs and for different number of unlicensed channels ($C = 2$, 4, and 7).

number of unlicensed channels is fixed to 7. Fig. 3.9 shows that inter-technology fairness is satisfied. This can be clearly seen in scenario (b) for the case of $M_c = 1$. For instance, when the traffic ratio is 1, LTE/WLAN airtime ratio is 1 and thus equal weighted airtime is allocated for each technology (given that $P_{\text{LTE}} = 1$ and $P_{\text{WiFi}} = 1$). From Fig. 3.9, we can also see that enabling carrier aggregation impacts the resource allocation outcome. In fact, we can see that a considerable gain in terms of spectrum access time can be achieved with CA. For instance, in the case of 2 SBSs, the LTE/WLAN airtime ratio increases from 0.84 for $M_c = 1$ to 1.7 and 2.4 for $M_c = 2$ and 3 respectively for the value of 0.6 for LTE/WLAN traffic ratio. On the other hand, this gain decreases as more SBSs are deployed and for a densely deployed LTE-LAA network, there is no need to aggregate more channels. This can be seen from (b) where the LTE-LAA network gets the same airtime share for $M_c = 1$, 2 and 3 (as also shown in Remark 1). Moreover, Fig. 3.9 shows that deploying more SBSs does not necessarily allow more airtime for the LTE-LAA network. For example, LTE/WLAN airtime ratio of scenarios (a) and (b) corresponding to 0.6 LTE/WLAN traffic ratio is equal to 0.84 and 0.6 respectively for $M_c = 1$. Consequently, the proposed scheme can avoid causing performance degradation to WLAN in the case LTE operators selfishly deploy a high number of SBSs.

Fig. 3.10 investigates the proportion of served LTE-LAA traffic for different network parameters. From Fig. 3.10, we can see that, as the number of SBSs increases, the proportion of LTE-LAA served traffic, relative to its corresponding offered load decreases. Moreover, reducing the number of unlicensed channels

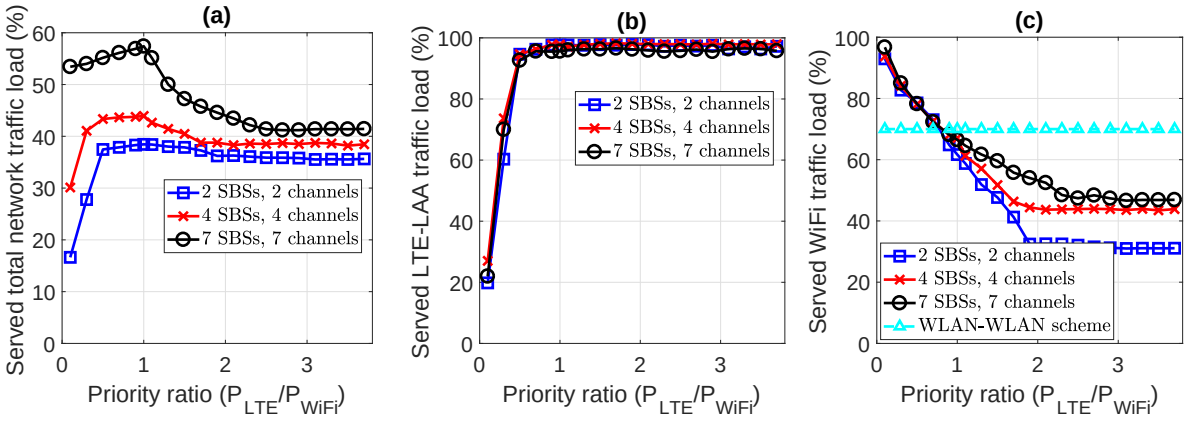

Figure 3.11. The proportion of the (a) total network served traffic load (b) LTE-LAA served traffic load and (c) WiFi served traffic load as a function of the priority fairness ratio on the unlicensed band, (P_{LTE}/P_{WiFi}). The straight line in (c) represents the proportion of WiFi served traffic load for the case when the LTE network is replaced by an equivalent WiFi network.

leads to a decrease in the proportion of LTE-LAA served traffic. Although the number of available unlicensed channels are not players in the game, they affect spectrum allocation action selection for each SBS. As the number of channels increases, the action space for the channel selection vector increases, thus giving more opportunities for an SBS to serve more of its offered load.

Fig. 3.11 shows the total network served traffic load as well as that of LTE-LAA and WiFi as a function of the priority fairness ratio on the unlicensed band ($P_{\text{LTE}}/P_{\text{WiFi}}$) for three different network scenarios considering $T = 6$. From Figs. 3.11 (b) and (c), we can see that more LTE-LAA and less WiFi traffic load is served as $P_{\text{LTE}}/P_{\text{WiFi}}$ increases and thus the priority fairness parameters P_{LTE} and P_{WiFi} can be regarded as network design parameters that can be adjusted in a way that would avoid LTE-LAA from aggressively offloading traffic to the unlicensed bands. Moreover, from Fig. 3.11 (a), we can see that the served total network traffic load is maximized at $P_{\text{LTE}}/P_{\text{WiFi}} = 1$ thus allowing an efficient utilization of the unlicensed spectrum. On the other hand, from Fig. 3.11 (c), we can see that the served WiFi traffic load for our proposed scheme is greater than or equal to the served WiFi traffic load for the case in which LTE-LAA is replaced by an equivalent WiFi network for values of $P_{\text{LTE}}/P_{\text{WiFi}}$ less than 0.8. From Fig. 3.11 (c), we can conclude that the WiFi performance for our proposed spectrum sharing scheme, when considering equal weighted airtime share (i.e., $P_{\text{LTE}}/P_{\text{WiFi}} = 1$), achieves very close performance to the case when only WLAN

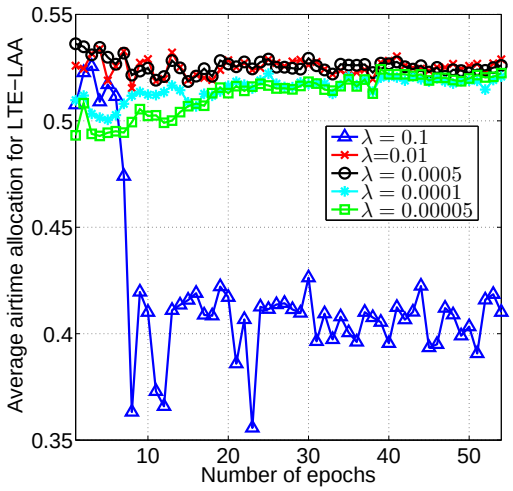

Figure 3.12. The average airtime allocated for LTE-LAA as a function of the number of epochs for different values of the learning rate.

is operating over the unlicensed spectrum. For instance, the proportion of WiFi served traffic load corresponds to 68% for the WiFi-LTE scenario as opposed to 70% for the WiFi-WiFi scenario in the case of 4 SBS and 4 channels. This slight decrease is mainly due to the differences in the MAC layers of both technologies. For instance, LTE adopts a more efficient scheduling mechanism and has less overhead as compared to WiFi. In particular, the DCF protocol of WiFi results in the channel being unused for some period of time and, thus, WiFi should be given a slightly larger priority in that case. In summary, we can deduce that the values of P_{LTE} and P_{WiFi} can be regarded as tuning parameters that allow the network operator to achieve a tradeoff between efficiency and fairness.

Fig. 3.12 shows the average value of airtime allocated to the LTE-LAA network as a function of the number of epochs required for the network to converge while considering different values for the learning rate. The learning rate determines the step size the algorithm takes to reach the minimizer and thus has an impact on the convergence rate of our proposed framework. Moreover, an epoch, which consists of multiple iterations, is a single pass through the entire training set, followed by testing of the verification set. From Fig. 3.12, we can see that for $\lambda = 0.1$, our proposed algorithm requires more than 50 epochs to approximate the reward function, while, for $\lambda = 0.01$, it only needs 20 epochs. In fact, for $\lambda = 0.1$, we can see that our proposed algorithm fluctuates around a different region of the optimization space. Clearly, a learning rate that is too large can cause the algorithm to diverge from the optimal solution. This is because too large initial learning rates will decay the loss function faster and thus make the

model get stuck at a particular region of the optimization space instead of better exploring it. On the other hand, a learning rate that is too small results in a low speed of convergence. For instance, for $\lambda = 0.0001$ and $\lambda = 0.00005$, our proposed algorithm requires ~ 40 epochs to converge. Therefore, although we use an adaptive learning rate approach, the optimization algorithm relies heavily on a good choice of an initial learning rate [165]. In other words, the initial value of the learning rate should be within a particular range in order to have good performance. Choosing a proper learning rate is an important key aspect that has an impact on the solution as well as the convergence speed. The optimal value of the initial learning rate is dependent on the dataset under study, where for each dataset, there exists an interval of good learning rates at which the performance does not vary much [99]. This in turn necessitates the need for experimental studies in order to search for good problem-specific learning rates [165]. A typical range of the learning rate for the dataset under study falls approximately between 0.0005 and 0.01, requiring ~ 20 epochs.

3.6 Summary

In this chapter, we have proposed a proactive resource allocation framework for the coexistence of LTE-LAA and WiFi in the unlicensed band. We have formulated a game model where each SBS seeks to maximize its rate over a given time horizon while achieving long-term equal weighted fairness with WLAN and other LTE-LAA operators transmitting on the same channel. To solve this problem, we have developed a novel deep learning algorithm based on LSTMs. The proposed algorithm enables each SBS to decide on its spectrum allocation scheme autonomously with limited information on the network state. Simulation results have shown that the proposed approach yields significant performance gains in terms of rate compared to conventional approaches that considers only instantaneous network parameters such as instantaneous equal weighted fairness, proportional fairness and total network throughput maximization. Results have also shown that our proposed scheme prevents disruption to WLAN operation in the case large number of LTE operators selfishly deploy LTE-LAA in the unlicensed spectrum.

Chapter 4

Holistic Small Cell Traffic Balancing across Licensed and Unlicensed Bands

4.1 Introduction

As we have discussed in Chapter 2, LTE-LAA has emerged as an effective solution to overcome the scarcity of the radio spectrum [131]. When operating over the licensed and the unlicensed bands simultaneously, the problem of SBS traffic balancing arises which essentially impacts the coexistence of LTE and WiFi over the unlicensed band. Therefore, the main scope of this chapter is to provide a holistic approach for LTE-LAA small cell traffic balancing by jointly optimizing the use of the licensed and unlicensed bands. In particular, we pose this traffic balancing as an optimization problem that seeks proportional fair coexistence of WiFi, small cell, and macro cell users by adapting the transmission probability of the LTE-LAA small cell in the licensed and unlicensed bands. The motivation for this formulation is for the LTE-LAA small cell to autonomously switch between or aggregate licensed and unlicensed bands depending on the interference/traffic level and the number of active users in each band. The work presented in this chapter has appeared in the Proceedings of ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM) [25]. Table 4.1 provides a summary for the description of the main notations used in this chapter. Next, we give an overview on the related literature and then we present our problem statement and contribution.

Table 4.1 Variables and notations.

Notation	Description
N_m	Number of macro UEs
N_f	Number of small cell UEs
N_w	Number of WiFi STAs
α	Fraction of time SBS is muted on the unlicensed channel
β	Fraction of time SBS is transmitting over the licensed band
$\Gamma^l_{F,f}$	DL SINR at SUE f over the licensed band
$\Gamma^u_{F,f}$	DL SINR at SUE f over the unlicensed band
$P_{F,f}$	Received signal power for SUE f
σ^2	Thermal noise power
$P_{M,m}$	Received signal power for MUE m
$P_{w,succ}$	Probability of successful transmission
$P_{w,idle}$	Probability of an idle slot
$P_{w,busy}$	Probability of a busy slot
τ_w	Stationary Probability of STA w
s_m	Throughput of MUEs
s_f	Throughput of SUEs
s_w	Throughput of WiFi STAs
s_m^{noABS}	Throughput of MUE m when SBS is transmitting over the licensed band
s_m^{ABS}	Throughput of MUE m during the ABS period of SBS
s_f^l	Throughput of SUE f when SBS is transmitting over the licensed band
s_f^u	Throughput of SUE f when SBS is transmitting over the unlicensed band
$\overline{R}_w$	Normalized offered load of WiFi STAs

4.1.1 Related Work

LTE use of unlicensed bands has been receiving growing amount of attention within the research community in recent years. The authors in [166] provide an overview of LTE-LAA as well as the benefits and challenges it brings. Several papers have looked at the performance impact of LTE operating in unlicensed bands on WiFi. In a recent paper [167], the authors conduct an experimental evaluation for characterizing the interference impact of LTE-LAA on WiFi under various network conditions; it is shown that the impact of LTE-LAA on WiFi throughput depends on the channel bandwidth, center frequency and MIMO and can be heavily degraded for some scenarios. Concerning mechanisms for LTE-WiFi coexistence, most of the previous work uses muting (adaptive duty cycling) [15,17,22–24]. More crucially, much of the existing work does not consider the operation of LTE-LAA SBS in the licensed band while optimizing its use in

the unlicensed bands alongside WiFi. This can however lead to a suboptimal resource allocation when seen globally. For instance, it can result in an over-utilization of the unlicensed band by LTE-LAA SBS and a decrease in WLAN performance, as it will be shown later in Section 4.5.

LTE-LAA small cells enable efficient and flexible use of the unlicensed spectrum, leveraging the LTE-Advanced carrier aggregation feature. Nevertheless, early work on traffic balancing across licensed and unlicensed bands (e.g., [168, 169]) focused on dual-access small cells (with both LTE and WiFi air interfaces) and thus lacking these benefits. To the best of our knowledge, [1] is the only notable traffic balancing work in the literature that applies to LTE-LAA small cells. The proposed traffic balancing technique in [1] is based on adjusting the power level in the licensed spectrum and the number of muted subframes in the unlicensed bands. We identify three aspects of the work in [1] discussed below, which together result in a lower WLAN performance and a degradation in the overall network performance compared to our proposed scheme, as shown later in Section 4.5.

1. *Use of power control in the licensed band.* In the context of ICIC management in heterogeneous networks (HetNets), 3GPP Release 10 introduced ABS as an efficient way to enhance the network performance. In [170], the authors evaluate the 3GPP enhanced ICIC (eICIC) techniques through realistic system-level simulations where it is shown that the ABS eICIC time method provides the best macrocell UE (MUE) protection as compared to other eICIC power methods. There is other work (e.g., [171]) which also shows that ABS muting achieves better macro-layer performance at less degradation of the SBS layer performance as compared to power adaptation. Therefore, the use of power control on the licensed band in [1] leads to a sub-optimal performance on both the licensed and the unlicensed bands given the fact that the coexistence mechanism in the licensed spectrum directly influences the optimization process in the unlicensed band.

2. *Considering a fixed level of performance for MBS.* The use of a fixed and predefined interference threshold value for MBS in [1] results in prioritizing the MBS performance irrespective of the degradation level caused to the SBS layer. This uncoordinated optimization approach on the licensed band would result in an unfair share of that band which in turn could lead to an

over-utilization of the unlicensed band by the SBS and thus a degradation in the WLAN performance.

3. *Sequential approach to optimizing the licensed band first then the unlicensed band.* The authors in [1] consider a sequential approach for optimizing both bands i.e., the output of the power allocation sub-problem in the licensed spectrum serves as an input to the muting sub-problem for the unlicensed bands. This results in prioritizing the licensed band and potentially over-utilizing the unlicensed band by SBS as well as degrading the total network performance.

4.1.2 Problem Statement and Contribution

The main objective of this chapter is to propose a holistic small cell traffic balancing scheme across the licensed and the unlicensed bands. In essence, LTE-WiFi coexistence depends on the extent to which LTE-LAA small cells (operating in both licensed and unlicensed bands) rely on unlicensed spectrum to meet their traffic demand, and this in turn is dependent on the nature of inter-tier interference in the licensed spectrum shared by a macro cell and small cells in its coverage area. This link between LTE small cell operation in the unlicensed band and inter-tier/inter-cell interference in the licensed spectrum is essentially the traffic balancing problem[1]. The transmission of the SBS on the unlicensed band can disrupt WiFi transmissions as the latter relies on a contention-based channel access and hence starvation may occur when coexisting with LTE. On the other hand, LTE-LAA SBS transmission on the licensed band can cause inter-tier/inter-cell interference to the macro cell and other small cell users, potentially degrading their throughput. Thus addressing the traffic balancing problem is challenging as it entails a LTE-LAA small cell base station to adaptively decide on how to steer its traffic between the licensed and unlicensed bands while optimizing the overall network performance and achieving fair coexistence among the technologies operating on both bands. Though the above discussion highlights the importance of traffic balancing for optimizing the performance of co-located networks based on different technologies (LTE and WiFi) sharing same unlicensed bands, and for more effective LTE-WiFi coexistence, this problem has till date received little at-

[1] Traffic balancing can be seen as addressing LTE-WiFi coexistence and LTE traffic offloading challenges together.

tention in the research literature with [1] as the only notable work. Nevertheless, the work in [1] leads to an inefficient utilization of the available radio resources due to the inefficient coexistence mechanism on the licensed band as well as the sequential adaptation approach for optimizing both bands, as discussed in the previous subsection.

The main contribution of the work presented in this chapter is a holistic framework for LTE-LAA small cell traffic balancing across licensed and unlicensed bands. In other words, we aim to jointly address the LTE-LAA small cell operation in licensed and unlicensed bands by determining its transmission behavior on both bands in a coordinated fashion depending on the interference/traffic levels on each of the bands. Specifically, we make the following key contributions:

- We present a formulation of the optimization problem for holistic traffic balancing that seeks PF coexistence of WiFi, small cell and macro cells by deciding on the transmission probability of LTE-LAA small cell in the licensed and unlicensed bands. The intention behind this formulation is for the LTE-LAA SBS to switch between or aggregate licensed and unlicensed bands depending on the interference/traffic level and number of active UEs in each cell. We derive a closed form solution for the aforementioned optimization problem. An attractive aspect of our solution is that it can be applied online by each LTE-LAA SBS, adapting its transmission behavior in each of the bands, and without explicit communication with WiFi nodes.

- We also propose a transmission mechanism for the operation of SBS on the licensed and unlicensed bands. Our mechanism leverages the above mentioned traffic balancing solution and aims at avoiding the disruption to on-going WiFi transmissions while adhering to the LTE frame structure.

- We provide extensive numerical and simulation results using several scenarios to highlight the main capabilities of our proposed scheme. Results show that LTE-LAA SBS, aided by our scheme, would adaptively steer its traffic from one band to another or transmit on both bands simultaneously depending on the interference/traffic levels and number of active UEs on each of the bands. Simulation results additionally demonstrate the effectiveness of our proposed scheme in comparison with [1] and other approaches, representing the state-of-the-art. They reveal that approaches focusing on coexistence in one band while ignoring the other cause load imbalance and

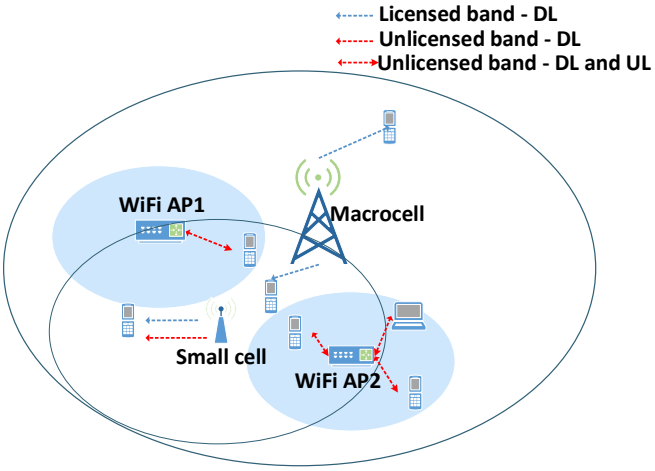

Figure 4.1. Illustration of the system model.

a decrease in the total network throughput and/or fairness. On the other hand, our approach, aided by its holistic nature, results in improved network performance as it achieves a better tradeoff between maximizing the total network throughput and attaining fairness among all network flows while also providing better LTE-WiFi coexistence.

The rest of this chapter is organized as follows. Section 4.2 details the system model for the coexistence of an LTE-LAA SBS with an LTE MBS and WLANs in the overlapping coverage area. In Section 4.3, we present an optimization problem for balancing traffic of the LTE-LAA SBS on the licensed and unlicensed bands and also derive a closed form solution for the problem. Section 4.4 describes our proposed transmission mechanism for the operation of the LTE-LAA SBS on the licensed and unlicensed bands. Section 4.5 presents numerical and simulation results for the proposed algorithm and compares its with other approaches. Finally, a summary is provided in Section 4.6.

4.2 System Model

We consider a system model (depicted in Fig. 4.1) similar to that in [1,24] consisting of a macrocell base station, a small cell and multiple independently operated WiFi networks. We assume a dual band small cell that transmits on both licensed and unlicensed bands via the LTE CA feature. The licensed band is shared between MBS and SBS where smaller portions of the spectrum, referred

to as RBs, are allocated to UEs. On the other hand, SBS and WiFi networks share an unlicensed channel in the time domain and hence at a particular time, the unlicensed channel is occupied by either SBS or WiFi. This represents a dense WiFi deployment scenario where SBS and WiFi may need to time share the same channel.

Let N_m, N_f and N_w, respectively, denote the number of macro-cell UEs, small cell UEs (SUEs) and WiFi stations (STAs) in a given time period T. We assume the supplemental downlink mode for the transmission of the small cell in the unlicensed band. On the other hand, traffic for WiFi STAs can be in either DL or UL directions. A full-buffer traffic model is assumed for the SBS, consistent with the motivation for SBS to use both licensed and unlicensed bands to meet its traffic demand.

In order to coexist with MBS on the licensed band and WLAN on the unlicensed band, we adopt in our model a holistic traffic balancing approach where SBS adjusts the proportion of time it transmits on both licensed and unlicensed bands. Therefore, at a particular time, the small cell would adaptively choose to transmit on the licensed, unlicensed or both bands depending on the interference level and traffic load of MUEs and WiFi nodes. The proposed scheme can be implemented at the MAC layer and hence the traffic assignment would be transparent to applications on the UEs. SBS would defer from transmission on the unlicensed band in order to allow WiFi transmission opportunities and on the licensed band in order to avoid inter-tier interference. Therefore, to decide on the proportion of time the small cell transmits on the licensed and unlicensed bands, the following decision variables are defined:

- $\alpha \epsilon [0, 1]$: the fraction of time SBS is *muted* on the unlicensed channel.

- $\beta \epsilon [0, 1]$: the fraction of time SBS is *transmitting* on the licensed band.

Note that upon muting on the licensed band, SBS would defer from sending data on the physical channels, however, would still send control and reference signals, an approach known as ABS [170]. On the other hand, the use of unlicensed band by the small cell is limited to data plane traffic while control and reference signals are transmitted by the SBS on a licensed carrier, which is essentially the license assisted access aspect of LTE-LAA. It is important to note here, that the work in [172] shows that conceptually both LBT and adaptive duty cycling

(muting) provide the same level of fairness to WiFi transmissions when properly configured.

4.2.1 Throughput Modeling

In order to assess the network performance for the coexistence of LTE MBS, LTE-LAA small cell and WiFi, we define the throughput for each of the MUEs, SUEs and WiFi STAs.

Upon the transmission on the licensed band, SBS would share the frequency band with MBS. In LTE, the DL RB allocation among UEs is via OFDMA, implying no intra-cell interference. However, frequency reuse in LTE can be one where macro and adjacent small cells may transmit on the same frequency leading to inter-cell interference. On the other hand, when SBS is transmitting on the unlicensed channel, it shares the channel with WLAN. Therefore, the DL SINR at SUE f, served by SBS F, in our model assuming a single MBS and SBS, during the transmission of SBS on the licensed and unlicensed channels respectively, can be expressed as follows:

$$\Gamma^{l}_{F,f} = \frac{P_{F,f}}{\sigma^2 + I_{M,f}} \qquad \text{and} \qquad \Gamma^{u}_{F,f} = \frac{P_{F,f}}{\sigma^2 + I_{W,f}} \tag{4.1}$$

where $P_{F,f}$ denotes the received signal power for SUE f from its serving SBS F, σ^2 is the thermal noise power, $I_{M,f}$ represents the interference power from MBS M on SUE f and $I_{W,f}$ corresponds to the aggregate interference power from neighbouring WLAN APs/STAs on SUE f. Note that upon the transmission of SBS on the unlicensed channel, WLAN would defer from transmission since WiFi STAs sense the carrier, i.e. listen to the channel before transmissions, and transmit only if the channel is idle. Therefore, $I_{W,f}$ corresponds to the interference power due to WLAN hidden terminals.

Similarly, the DL SINR at MUE m, served by MBS M, during the non-ABS and ABS periods of SBS on the licensed band respectively, can be expressed as follows:

$$\Gamma^{\text{noABS}}_{M,m} = \frac{P_{M,m}}{\sigma^2 + I_{F,m}} \qquad \text{and} \qquad \Gamma^{\text{ABS}}_{M,m} = \frac{P_{M,m}}{\sigma^2} \tag{4.2}$$

where $P_{M,m}$ denotes the received signal power for MUE m from its serving MBS M, and $I_{F,m}$ represents the interference power from SBS F on MUE m.

We denote by s_k the total throughput attained by an LTE UE k (where k is m or f). An upper bound for the DL UE throughput, based on Shannon's capacity,

is computed as follows:

$$s_k(\text{bps}) = \text{BW}_k \cdot \log_2(1 + \Gamma_k) \tag{4.3}$$

where BW_k is the channel bandwidth allocated to UE k and Γ_k is the SINR value of UE k.

To derive the throughput attained by a WiFi STA w when using the unlicensed band exclusively, we consider a slotted channel, as per the IEEE 802.11 modus operandi [173]. Let τ_w denote the stationary probability that station w is attempting transmission in a randomly chosen slot time. The total throughput $\hat{s}_w$ attained by a WiFi STA w when using the channel *exclusively* is:

$$\hat{s}_w(\text{bps}) = \frac{P_{w,succ} \cdot E[D_w]}{P_{w,idle} \cdot \sigma + P_{w,busy} \cdot T_b}, \tag{4.4}$$

where $E[D_w]$ is the expected payload size for station w, $P_{w,succ}$ is the probability of a successful transmission and can be expressed as $P_{w,succ} = \tau_w \prod_{i=1,i\neq w}^{N_w}(1 - \tau_i)$, $P_{w,idle}$ is the probability of an idle slot and can be expressed as $P_{w,idle} = \prod_{w=1}^{N_w}(1 - \tau_w)$ and $P_{w,busy}$ is the probability of a busy slot, regardless of whether it corresponds to a collision or a successful transmission and can be expressed as $P_{w,busy} = 1 - \prod_{w=1}^{N_w}(1 - \tau_w)$ [174]. σ and T_b correspond to the average durations of an idle and a busy slot respectively and thus the denominator corresponds to the mean duration of a WiFi MAC slot.

Therefore, during an epoch T, the throughput attained by a macro, small cell and WiFi UE respectively can be expressed as follows:

$$s_m = \beta s_m^{\text{noABS}} + (1 - \beta) s_m^{\text{ABS}} \tag{4.5}$$

$$s_f = \beta s_f^{\text{l}} + (1 - \alpha) s_f^{\text{u}} \tag{4.6}$$

and

$$s_w = \alpha \hat{s}_w \tag{4.7}$$

where s_m, s_f and s_w are the achieved throughputs of MUEs, SUEs and WiFi STAs respectively during a given period of time T. s_m^{noABS} and s_m^{ABS} correspond to the throughput achieved by MUE m during the transmission of the SBS on the licensed band and during the ABS period of SBS, respectively. s_f^{l} and s_f^{u} correspond to the throughput of SUE f during the transmission of SBS on the licensed band and an unlicensed channel, respectively.

4.3 Holistic Traffic Balancing

In order to maximize the total network throughput while coexisting fairly with other LTE and WiFi cells, we aim in this section at proposing a traffic balancing approach that aims at providing a proportional fair coexistence of WiFi STAs, SUEs and MUEs. The rationale behind this approach is to allow SBS to either switch between or aggregate the unlicensed and licensed bands based on the interference level on each band. This will allow higher throughput for MUEs that are in the vicinity of the SBS when SBS is not transmitting on the licensed band, and similarly, more transmission opportunities for WiFi nodes when SBS is not transmitting on the unlicensed band. Therefore, the utility function can be expressed as the product of the throughputs obtained by SUEs, MUEs and WiFi STAs:

$$\mathcal{U} = \prod_{m=1}^{N_m} s_m \prod_{f=1}^{N_f} s_f \prod_{w=1}^{N_w} s_w \tag{4.8}$$

$\mathcal{U}$ in turn can be expressed as the summation of the logarithmic function of the achieved rates as given below:

$$\begin{aligned} \mathcal{U}_{\log} = & \sum_{m=1}^{N_m} \log(s_m) + \sum_{f=1}^{N_f} \log(s_f) + \sum_{w=1}^{N_w} \log(s_w) \\ = & \sum_{m=1}^{N_m} \log\left[\beta s_m^{\text{noABS}} + (1-\beta) s_m^{\text{ABS}}\right] \\ & + \sum_{f=1}^{N_f} \log\left[\beta s_f^{\text{l}} + (1-\alpha) s_f^{\text{u}}\right] \\ & + \sum_{w=1}^{N_w} \log\left[\alpha s_w\right] \end{aligned} \tag{4.9}$$

The proposed utility function $\mathcal{U}_{\log}$ corresponds to a PF coexistence of MUEs, SUEs and WiFi STAs. The PF scheduling algorithm has been an attractive allocation criterion in wireless networks since it maintains a balance between maximizing the total network throughput while achieving good fairness among network flows [162]. Therefore, our optimization problem is formulated as follows:

$$\max_{\alpha,\beta} \ \mathcal{U}_{\log} \tag{4.10}$$

subject to

$$\alpha \leq \overline{R}_w \tag{4.11}$$

$$\alpha \leq \beta \tag{4.12}$$

$$0 \leq \alpha \leq 1, 0 \leq \beta \leq 1 \tag{4.13}$$

where $\overline{R}_w(\leq 1)$ corresponds to the normalized offered load across all WiFi stations; it can be obtained via long-term channel sensing where SBS would monitor the WLAN activity on the unlicensed band and estimate the average WLAN traffic load. In the above formulation, constraint (4.11) limits the fraction of time SBS is muted on the unlicensed band to the time it is busy due to WiFi activity. In other words, it is to make sure that the unlicensed band is not underutilized. The purpose of constraint (4.12) is to ensure that SBS transmits on either the licensed or the unlicensed channel at any given point in time. Constraints (4.13) limit the range of values variables α and β can take.

Lemma 2. $\log(x)$ *is concave. It follows that the utility function* $\mathcal{U}_{\log}$ *is an affine combination of concave functions, and hence is concave. Therefore, the optimization problem defined by (4.10)-(4.13) is concave since the objective function and the feasible region defined by the constraints are concave and hence a closed form solution can be obtained using the Karush-Kuhn-Tucker (KKT) conditions at optimality [175].*

Based on the above lemma, we now aim to derive a closed form solution for the optimization problem (4.10)-(4.13) using the KKT conditions at optimality. The KKT conditions are necessary and sufficient for convex optimization problems and consist of the stationarity, primal and dual feasibility, and complementary slackness conditions [175]. Therefore, the Lagrangian of the optimization problem (4.10)-(4.13) can be written as follows:

$$\begin{aligned}\mathcal{L}(\alpha, \beta, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) = -\mathcal{U}_{\text{total}} + \lambda_1(\alpha - \overline{R}_w) + \lambda_2(\alpha - \beta) - \lambda_3\alpha \\ + \lambda_4(\alpha - 1) - \lambda_5\beta + \lambda_6(\beta - 1)\end{aligned} \tag{4.14}$$

where λ_1, λ_2, λ_3, λ_4, λ_5 and λ_6 correspond to the Lagrangian multipliers of constraints (4.11)-(4.13).

In the *first step*, we compute the candidates for an optimal solution pair (α^*, β^*) from the possible combinations of feasible solutions satisfying the stationarity and complementary slackness conditions. Note that the total number of possible combinations for the Lagrangian multipliers is 64 (i.e., 2^6) where a given multiplier could be either zero or non-zero (NZ) at an optimal solution. However,

for our optimization formulation, only 6 combinations are possible candidates for an optimal solution due to some infeasible and redundant combinations. For instance, the combinations that have λ_4 and λ_5 as NZ can be omitted since their corresponding solution is (α^*, β^*) = (1,0), however, this will lead to the violation of constraint (4.12). Similarly, if a constraint has finite values for both lower and upper bounds, one would need to consider the possible combinations when at most one of the Lagrange multipliers for that constraint is NZ. This is due to the fact that one or the other, or both, of the multipliers will always be equal to zero since only one of the bounds can be active at a time. Therefore, the combinations that have both λ_3 and λ_4 or λ_5 and λ_6 as NZ can be omitted. Moreover, we impose a non-zero muting period on the unlicensed band (i.e., restrict α to be greater than 0) in order to allow the small cell to sense WiFi activity and number of stations and thus we omit the combinations having λ_3 as NZ. More details on the possible combinations for solution candidates is given in Appendix A.1. Based on the above, the 6 candidate solutions for α^*, β^* and $(\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*, \lambda_5^*, \lambda_6^*)$ are as follows:

Candidate solution 1: λ=(NZ,0,0,0,0,NZ)

$$\alpha_1 = \overline{R}_w \quad \text{and} \quad \beta_1 = 1$$

$$\lambda_1 = -\sum_{f=1}^{N_f} \frac{s_f^u}{\beta_1 s_f^l + (1-\alpha_1) s_f^u} + \frac{N_w}{\alpha_1}$$

$$\lambda_6 = \sum_{m=1}^{N_m} \frac{(s_m^{\text{noABS}} - s_m^{\text{ABS}})}{\beta_1 s_m^{\text{noABS}} + (1-\beta_1) s_m^{\text{ABS}}} + \sum_{f=1}^{N_f} \frac{s_f^l}{\beta_1 s_f^l + (1-\alpha_1) s_f^u}$$

Candidate solution 2: λ=(0,0,0,0,0,NZ)

α_2 corresponds to the solution of the following equation:

$$\sum_{f=1}^{N_f} \frac{s_f^u}{s_f^l + (1-\alpha_2) s_f^u} - \frac{N_w}{\alpha_2} = 0$$

$$\beta_2 = 1$$

$$\lambda_6 = \sum_{m=1}^{N_m} \frac{(s_m^{\text{noABS}} - s_m^{\text{ABS}})}{\beta_2 s_m^{\text{noABS}} + (1-\beta_2) s_m^{\text{ABS}}} + \sum_{f=1}^{N_f} \frac{s_f^l}{\beta_2 s_f^l + (1-\alpha_2) s_f^u}$$

Candidate solution 3: λ=(NZ,NZ,0,0,0,0)

$$\alpha_3 = \overline{R}_w \quad \text{and} \quad \beta_3 = \overline{R}_w$$

$$\lambda_2 = -\sum_{m=1}^{N_m} \frac{(s_m^{\text{noABS}} - s_m^{\text{ABS}})}{\beta_3 s_m^{\text{noABS}} + (1-\beta_3) s_m^{\text{ABS}}} - \sum_{f=1}^{N_f} \frac{s_f^l}{\beta_3 s_f^l + (1-\alpha_3) s_f^u}$$

$$\lambda_1 = -\sum_{f=1}^{N_f} \frac{s_f^u}{\beta_3 s_f^l + (1-\alpha_3) s_f^u} + \frac{N_w}{\alpha_3} - \lambda_2$$

Candidate solution 4: λ=(NZ,0,0,0,0,0)

$$\alpha_4 = \overline{R}_w$$

β_4 corresponds to the solution of the following equation:

$$-\sum_{m=1}^{N_m} \frac{(s_m^{\text{noABS}} - s_m^{\text{ABS}})}{\beta_4 s_m^{\text{noABS}} + (1-\beta_4) s_m^{\text{ABS}}} - \sum_{f=1}^{N_f} \frac{s_f^l}{\beta_4 s_f^l + (1-\alpha_4) s_f^u} = 0$$

$$\lambda_1 = -\sum_{f=1}^{N_f} \frac{s_f^u}{\beta_4 s_f^l + (1-\alpha_4) s_f^u} + \frac{N_w}{\alpha_4}$$

Candidate solution 5: λ=(0,NZ,0,0,0,0)

α_5 is equal to β_5 and their corresponding value is the solution of the following equation:

$$-\sum_{m=1}^{N_m} \frac{(s_m^{\text{noABS}} - s_m^{\text{ABS}})}{\alpha_5 s_m^{\text{noABS}} + (1-\alpha_5) s_m^{\text{ABS}}} - \sum_{f=1}^{N_f} \frac{s_f^l}{\alpha_5 s_f^l + (1-\alpha_5) s_f^u} + \sum_{f=1}^{N_f} \frac{s_f^u}{\alpha_5 s_f^l + (1-\alpha_5) s_f^u} - \frac{N_w}{\alpha_5} = 0$$

$$\lambda_2 = -\sum_{f=1}^{N_f} \frac{s_f^u}{\beta_5 s_f^l + (1-\alpha_5) s_f^u} + \frac{N_w}{\alpha_5}$$

Candidate solution 6: λ=(0,0,0,0,0,0)

α_6 and β_6 correspond to the solution of the following two equations:

$$\sum_{f=1}^{N_f} \frac{s_f^u}{\beta_6 s_f^l + (1-\alpha_6) s_f^u} - \frac{N_w}{\alpha_6} = 0$$

$$-\sum_{m=1}^{N_m} \frac{(s_m^{\text{noABS}} - s_m^{\text{ABS}})}{\beta_6 s_m^{\text{noABS}} + (1-\beta_6) s_m^{\text{ABS}}} - \sum_{f=1}^{N_f} \frac{s_f^l}{\beta_6 s_f^l + (1-\alpha_6) s_f^u} = 0$$

Note that two more candidate solutions exist for λ= (NZ,NZ,0,NZ,0,NZ) and λ= (0,NZ,0,NZ,0,NZ) where α and β are both equal to 1. However, we can avoid checking these two candidate solutions as they exist only in the case when $\overline{R}_w$=1 and hence their solution matches with that of candidate solution 1.

In the *second step*, we check the primal and dual feasibility conditions for each of the 6 candidate solution pairs and the pair satisfying these conditions is the optimal solution.

Note that all the candidate solutions are independent of the WiFi throughput s_w and hence the SBS needs to know only the normalized WiFi offered load as well as the number of active WiFi STAs; the SBS can learn the number of active WiFi STAs based on their corresponding MAC addresses during the sensing period [1]. The number of MUEs and their throughput can be conveyed to the SBS through the X2 interface. Using this information, SBS can determine the optimal values for α and β *locally* when needed.

4.4 A Transmission Mechanism for LTE-LAA SBS Operation

LTE is designed for the exclusive use of the spectrum and hence when operating on the unlicensed band, a new channel access scheme is needed to coexist with other devices having different air interfaces. Therefore, in this section, we propose a transmission mechanism for the operation of an LTE-LAA small cell on the licensed and unlicensed bands. This mechanism builds upon the problem formulation from Section 4.3 and incorporates a channel access scheme on the unlicensed channel that would allow LTE-LAA SBS to transmit on the unlicensed band in a way that would not disrupt any ongoing WiFi transmissions.

For our proposed mechanism, we divide the time domain into T epochs, where in each epoch we aim at finding the optimal values of α and β using the results of Section 4.3. Taking into account that LTE transmits only at the beginning of a subframe, our proposed transmission mechanism is aligned with LTE frame structure where $(1-\alpha)T$ and βT are rounded to an integer multiple of an LTE

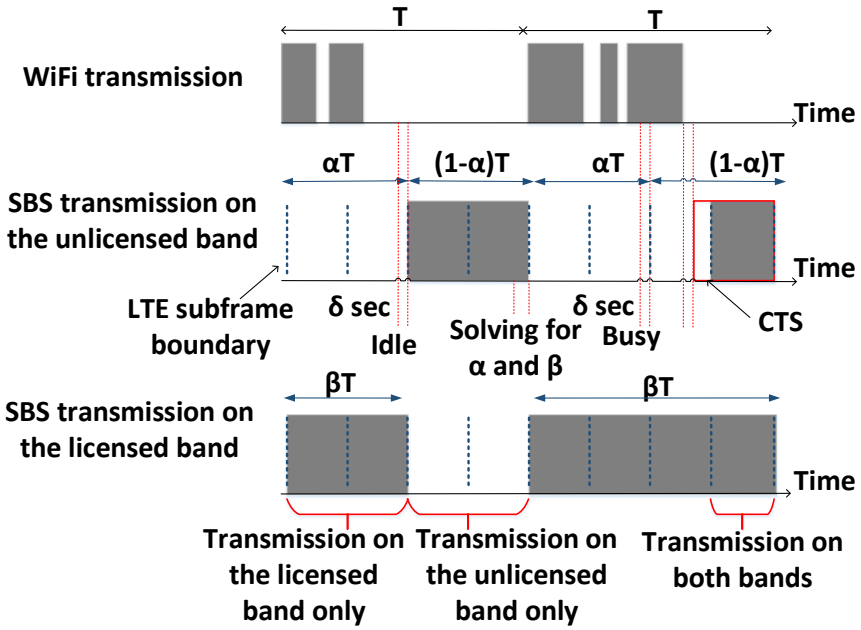

Figure 4.2. Illustration of the proposed SBS transmission mechanism on the licensed and unlicensed bands. The two possible states upon sensing the unlicensed channel (idle and busy) are demonstrated. SBS will remain in a sensing state when it encounters a busy channel. The three states of SBS (i.e., transmission on the licensed, unlicensed and both bands) are also shown.

subframe duration (1 msec). Moreover, we define δ as the duration of time the SBS would sense the unlicensed channel before attempting to transmit. Let δ be such that SIFS $< \delta <$ DIFS, and hence this will guarantee that the ACK of any previous WiFi transmission is received at the sender and that SBS would get access to the unlicensed channel before any other WiFi STA that would be sensing the channel at the same time. The proposed LTE-LAA transmission mechanism is illustrated in Fig. 4.2 where the two possible states upon sensing the channel (idle and busy) are demonstrated. Moreover, the steps of the proposed mechanism are summarized as follows:

1. SBS calculates the values of α and β before the beginning of a T period based on the throughput values and number of active nodes of the previous T period and using the results of Section 4.3.

2. At the beginning of a T period, SBS remains silent for the period αT on the unlicensed band and transmits for the period βT on the licensed band.

3. SBS senses the unlicensed channel for δ sec before αT expires in order to

detect any ongoing WiFi transmissions and guarantee alignment with LTE frame structure.

4. If the channel is idle, SBS transmits for a period of $(1-\alpha)T$.

5. If the channel is busy, SBS keeps on listening to the channel until it detects a silent period for a duration of δ sec in order to avoid the disruption to any ongoing WiFi transmission. After detecting a silent period of δ sec, SBS sends a CTS with the duration of the remaining time of the $(1-\alpha)T$ period to reserve the channel for SBS transmission on the unlicensed band. It is important to note that the maximum channel occupancy time is limited to 10 msec after which the unlicensed channel must be released and the LBT process is repeated. Therefore, for the cases where $(1-\alpha)T$ is less than 10 msec, there is a risk that the SBS will not be able to get access to the unlicensed band when the WLAN burst is larger than $(1-\alpha)T$. For such scenarios, the WLAN transmission period for the next T period is shortened accordingly to maintain the average time allocated for LTE-LAA and WLAN.

4.5 Simulation Results and Analysis

In this section, we examine the behavior of our proposed holistic traffic balancing scheme in various scenarios using a combination of numerical and simulation results. We also conduct a comparative study of our holistic traffic balancing approach with respect to [1] and other alternative approaches, representing other proposed techniques from the literature.

In simulations, for WiFi we consider the 802.11 DCF medium access mechanism based on CSMA/CA. We assume randomly located STAs that transmit and receive packets according to an independent Poisson process. For simplicity, we consider that all WiFi STAs use the same physical layer parameters, 64-QAM modulation with a 5/6 coding rate when using a 20 MHz channel, which provides a 65 Mbps MAC layer throughput. The simulation parameters for the 802.11 network are the same as those used in [24].

For the LTE and LTE-LAA networks, we assume the same channel conditions for all RBs on both bands and hence the same modulation and coding scheme i.e., 64 QAM with 5/6 coding rate, is applied to all RBs of the given 20 MHz

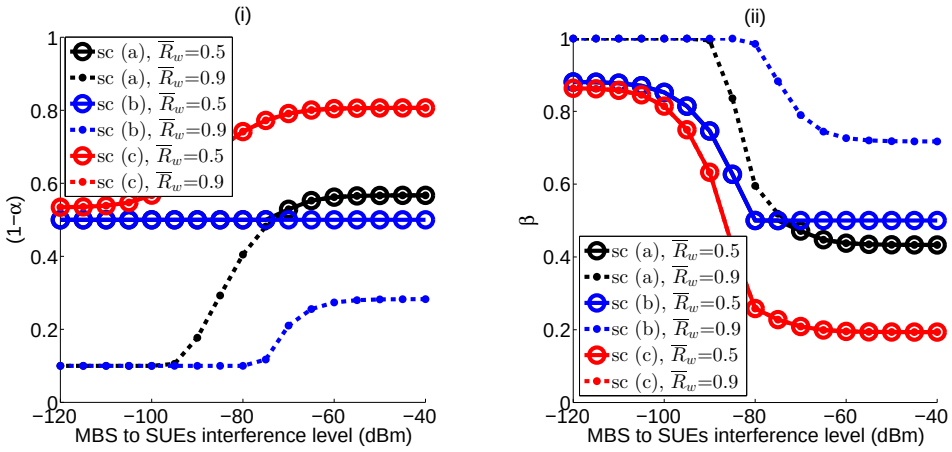

Figure 4.3. Numerical results for the optimal values of (i) $(1-\alpha)$ and (ii) β for varying levels of MBS to SUEs interference in three different scenarios; sc (a) considers an equal number of MUEs, SUEs and WiFi STAs, sc (b) considers the number of WiFi STAs to be three times that of each of MUEs and SUEs and sc (c) considers the number of each of MUEs and SUEs to be three times that of WiFi STAs. For the studied scenarios, we consider medium and high WiFi offered load i.e., $\overline{R}_w$=0.5 and 0.9 respectively, as well as a fixed value for SBS to MUEs interference level (-85 dBm).

channel. Maximum MAC layer throughput for LTE with the above settings is 75 Mbps. These simulation parameters are similar to the ones used in [1]. We assume a round robin scheduler and equal transmit power for all OFDM symbols in a transmission time interval due to the fact that all RBs have the same modulation and coding scheme and thus equal number of bits are allocated to each subcarrier. The maximum transmit power for MBS and SBS is 43 dBm and 23 dBm, respectively. We consider an urban area characterized by the path loss model (for outdoor and indoor locations of the base station and UEs) as given in [176]. A constant payload size of 1500 bytes is assumed for MUEs, SUEs and WiFi STAs. Simulation results are provided for the average of 1000 runs with a 95% confidence interval.

4.5.1 Behavior of α and β in different scenarios

In this subsection, we study the effect of the variation of the traffic arrival rate as well as the number of active UEs on the values of α and β by conducting numerical and simulation results for different practical deployment scenarios.

For the numerical results, we consider three different scenarios with different

number of MUEs, SUEs and WiFi STAs. Fig. 4.3 shows the optimal values of $(1-\alpha)$ and β as a function of the MBS to SUE interference level on the licensed band, for a fixed value of the SBS to MUE interference level (-85 dBm) and two different WLAN traffic loads ($\overline{R}_w$=0.5 and 0.9). Note that the MBS to SUE interference and the SBS to MUE interference levels are relevant during the non-ABS period only.

For the simulation results (shown in Fig. 4.4), we consider only scenario (a) of Fig. 4.3. Fig. 4.4 shows the variation of the proportion of time SBS is transmitting on the licensed and unlicensed bands during the period T as a function of the WLAN traffic arrival rate (λ_{WLAN} (packets/sec)) and for a low and high MUEs traffic arrival rates, i.e., $\lambda_{\text{MUE}} = 0.5$ and 2 (packets/sec) respectively. Note that λ_{WLAN} and λ_{MUE} correlate to $\overline{R}_w$ and inter-tier interference level respectively of Fig. 4.3. Each data point in the simulation results is obtained from 1000 runs, each of length 200 msec and with T set to 20 msec.

We can make the following observations from Figs. 4.3 and 4.4. First, comparing the three considered scenarios of Fig. 4.3, we conclude that *our proposed traffic balancing scheme provides per node airtime fairness among each of the MUEs, SUEs and WiFi STAs.* For example, consider -60 dBm for the value of MBS to SUEs interference level and $\overline{R}_w$=0.5 for the WLAN load, we observe that in scenario (c), SBS transmits more on the unlicensed band (80%) and less on the licensed band (20%) as compared to scenario (b) where SBS transmits 50% on the unlicensed band and 50% on the licensed band. This is because the number of each MUEs and SUEs is larger than that of WiFi STAs in scenario (c) while in scenario (b) the number of WiFi STAs is larger than each of the number of MUEs and SUEs.

Second, *our proposed scheme copes with the interference level on both bands by adapting the values of α and β.* This can be observed for high values of inter-tier interference in Fig. 4.3 or high values of λ_{MUE} in Fig. 4.4. In those scenarios, WLAN shares the unlicensed band with SBS for a proportion of time larger than its idle period, i.e., larger than (1-$\overline{R}_w$), in order to decrease the effect of inter-tier interference on the UEs throughput on the licensed band. For example, in Fig. 4.3, for scenario (a) and $\overline{R}_w$=0.9, SBS transmits for 55% of the time on the unlicensed band when the MBS to SUEs interference level is -60 dBm as compared to 10% when the MBS to SUEs interference level is -95 dBm. This can also be noted from Fig. 4.4 where (1-α) is equal to 20% for λ_{MUE}= 0.5 (packets/sec) but

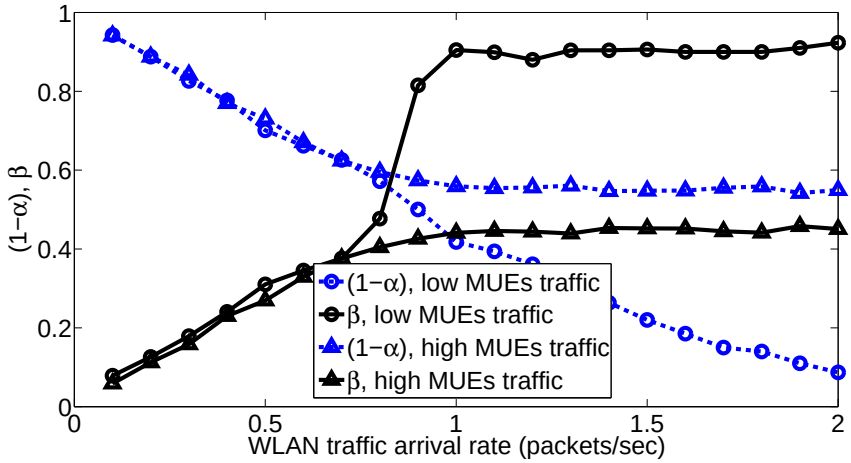

Figure 4.4. Simulation results for the variation of the proportion of time the SBS transmits on the licensed (β) and unlicensed bands ($1-\alpha$) as a function of the WLAN traffic arrival rate (λ_{WLAN}) and for a low and high MUEs traffic arrival rates i.e., λ_{MUE}= 0.5 and 2 (packets/sec) respectively, for a scenario of equal number of MUEs, SUEs and WLAN STAs.

increases to 55% for λ_{MUE}= 2 (packets/sec), for λ_{WLAN}=1.5.

Third, *our proposed traffic balancing scheme allows SBS to transmit on either one of the two bands or aggregate both bands through CA and thus increasing its capacity.* Given that SBS is muted for the period of α and (1-β) on the unlicensed and licensed bands respectively, we can deduce that it transmits on both bands simultaneously for a period of $(\beta-\alpha)T$ sec, and on one of the two bands for the remaining duration of the T period i.e., for a period of $(1-(\beta-\alpha))T$ sec, as per our proposed transmission mechanism of Section 4.4. For example, in Fig. 4.3, for scenario (b), $\overline{R}_w$=0.5 and MBS to SUEs interference level of -90 dBm, α=0.5 and β=0.75 and thus SBS transmits on both bands simultaneously for 25% of the T period. This can also be shown in Fig. 4.4 where α=0.6 and β=0.9 for λ_{WLAN}=1 and λ_{MUE}=0.5 and hence SBS transmits on both bands simultaneously for 30% of the T period.

Fourth, for all the considered scenarios of Figs. 4.3 and 4.4, we notice that *the unlicensed band is always utilized by either WLAN or SBS and hence this avoids its underutilization.* In other words, SBS is always transmitting on the unlicensed band for at least the portion of time that it is not utilized by WLAN i.e., (1-α) is always greater than or equal to (1-$\overline{R}_w$), consistent with constraint (4.11) in the optimization problem, irrespective of the value of inter-tier interference on the licensed band. For example, for $\overline{R}_w$=0.5 and 0.9, (1-α) is always greater than or equal to 0.5 and 0.1 respectively.

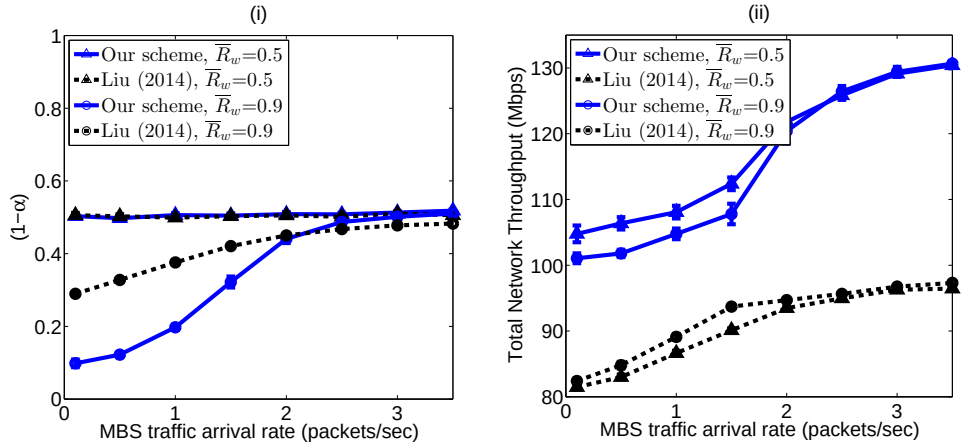

Figure 4.5. Simulation results for (i) the optimal value of the transmission ratio of SBS on the unlicensed band i.e., (1-α) and (ii) the total achieved network throughput as a function of the MBS traffic arrival rate (λ_{MUE}) for our proposed traffic balancing scheme (Our scheme) and the scheme in [1] (Liu (2014)). For the comparative study, we consider moderate and high WLAN offered load i.e., $\overline{R}_w$=0.5 and 0.9 respectively.

Fifth, for all the studied scenarios, *there exists an upper limit for the value of (1-α) which corresponds to the maximum proportion of time that WLAN would share its unlicensed band with LTE.* This can be observed in the cases of high inter-cell interference on the licensed band where a minimum airtime portion for WLAN, that is a function of the number of active UEs and WLAN activity, is guaranteed and thus allowing a fair LTE-WiFi coexistence. For example, in Fig. 4.3, for an equal number of SBS and WLAN UEs (i.e. scenario (a)), the upper limit for (1-α) is approximately 0.5.

Overall, the results demonstrate that our traffic balancing scheme performs as per expectations by steering SBS traffic from one band to another or using both bands simultaneously depending on the level of inter-tier interference on the licensed band, WiFi offered load and number of UEs in each band.

4.5.2 Comparison with existing traffic balancing scheme [1]

In this subsection, we compare the performance of our proposed scheme with that of [1] which also studies the problem of SBS traffic balancing across licensed and unlicensed bands. Unlike our scheme that jointly optimizes the muting pattern on both bands, the work in [1] takes a sequential approach adapting the power level in the licensed band first followed by adjusting the muting pattern on the

unlicensed channel. Fig. 4.5 shows simulation results for (i) the value of (1-α) and (ii) the total network throughput for the two schemes as a function of the MBS traffic arrival rate for two different values of the WLAN traffic load ($\overline{R}_w$=0.5 and $\overline{R}_w$=0.9). We can make the following high-level observations from Fig. 4.5:

Observation 1: *Overall, our proposed traffic balancing scheme achieves better LTE-WiFi coexistence.*

Observation 2: *For all the studied network scenarios, our proposed traffic balancing scheme achieves higher total network throughput.*

In what follows, we examine the reasons behind these observations. First, for scenarios of high WLAN load and when MBS is not in a full buffer state (i.e. $\lambda_{\text{MUE}} < 2.5$ (packets/sec)), corresponding to candidate solutions 2 or 6, our proposed scheme provides better LTE-WiFi coexistence while also achieving higher total network throughput as compared to [1]. This gain is due to the use of subframe muting instead of power adaptation, optimizing the MBS and SBS in a coordinated fashion instead of having a fixed level of performance for MBS, and optimizing the licensed and unlicensed bands in a holistic (joint) manner instead of adopting a sequential approach (see all aspects for [1] discussed in Subsection 4.1.1). The gain for solving the problem holistically as compared to sequentially is characterized separately in Section 4.5.3 where we consider a variant of our scheme that adopts an independent muting strategy on both bands. On the other hand, the gain due to the other two differences between our scheme and that of [1] can be clearly seen from the value of α for candidate solutions 2 or 6 with N_f=1:

$$\alpha = \frac{N_w(T_f^l + s_f^u)}{s_f^u(N_w + 1)} \tag{4.15}$$

where T_f^l is the throughput achieved by SBS on the licensed band and corresponds to $\beta \cdot s_f^L$ for our proposed scheme and $s_f^L(P_f^*)$ (i.e., a function of the optimal allocated power) for the proposed algorithm of [1]. Therefore, from Equation (4.15), we can note that higher values of T_f^l result in higher values for α and thus less utilization of the unlicensed band. Given that ABS muting achieves better macro-layer performance at less degradation of the SBS layer performance as compared to power adaptation, for a specified level of performance for MUEs (e.g., minimum outage level, minimum interference level from SBSs to MUEs), ABS muting causes less degradation in the performance of the SBS layer as compared to power control, i.e., $\beta \cdot s_f^L > s_f^L(P_f^*)$. Following Equation (4.15), our

proposed scheme results in less utilization of the unlicensed band and thus allows more WLAN transmission opportunities as compared to [1] while maximizing the total network performance.

On the other hand, in the case of a full-buffer MBS (i.e. $\lambda_{\text{MUE}} \geq 2.5$ (packets/sec)) and at high WLAN load, corresponding to candidate solution 5, we can notice that the value of (1-α) for our proposed scheme (0.51) is slightly higher than that of [1] (0.49). This is due to the high interference level on the licensed band and thus the need to steer more traffic on the unlicensed band in order to guarantee that the SBS is transmitting on at least one of the two bands at a given time (see constraint (4.12) in the optimization problem). Note, however, that (1-α) would converge to its upper limit (i.e., ~ 0.5 for the studied scenarios) and thus allowing a fair LTE-WiFi coexistence.

Second, our proposed scheme achieves similar performance on the *unlicensed* band as that of [1] for the case of moderate WLAN load ($\overline{R}_w = 0.5$) but it results in a higher total network throughput. For these scenarios, the value of α is limited by $\overline{R}_w$ (corresponding to candidate solutions 1, 3 or 4) and thus the increase in the total network throughput is due to the improvement in the performance on the licensed band i.e., due to the use of subframe muting instead of power adaptation and optimizing the MBS and SBS in a coordinated fashion instead of having a fixed level of performance for MBS (i.e., see aspects (1) and (2) of [1] discussed in Subsection 4.1.1).

In summary, our proposed scheme achieves better utilization of the available radio resources compared to [1] (an increase of 28.3% in the total network throughput for the studied scenarios) while increasing the transmission opportunities for WiFi on the unlicensed band.

4.5.3 Comparison with alternative approaches

In this subsection, we compare the performance of our proposed traffic balancing approach with a broad spectrum of alternative approaches. As performance metrics, we consider throughput and fairness obtained using each of the various different approaches. Denote by $\eta(s_i)$ the efficiency of a resource allocation scheme where $\eta(s_i)$ is defined as the sum of all the UEs throughput i.e., $\eta(s_i)= \sum_{i=1}^{N} s_i$ (where i is m, f, or w and $N=N_m+N_f+N_w$), and its fairness is given by

the Jain's index defined below [163]:

$$\mathcal{J}(s_i) = \frac{\left(\sum_{i=1}^{N} s_i\right)^2}{N \cdot \sum_{i=1}^{N} s_i^2} \tag{4.16}$$

The value of the Jain's fairness index lies in $[\frac{1}{N}, 1]$ where the value of $(\frac{1}{N})$ corresponds to the least fair allocation in which only one UE attains a non-zero throughput and the value of (1) corresponds to the most fair allocation in which all UEs achieve equal rates. Therefore, an efficient allocation of the radio resources seeks to provide a tradeoff between $\eta(s_i)$ and $\mathcal{J}(s_i)$ [163].

We compare the throughput and fairness of our proposed scheme with the following set of approaches:

- *Case 1 - No Muting on Licensed*: SBS operates on both bands, however, considering a PF muting strategy on the unlicensed band only and hence providing a coexistence technique with WLAN only. On the licensed band, MBS and SBS transmit simultaneously, and hence inter-tier interference is not eliminated.

- *Case 2 - No Muting on Unlicensed*: SBS operates on both bands, however, considering a PF muting strategy on the licensed band only and hence providing a coexistence technique with MBS only. On the unlicensed band, SBS is transmitting all the time, and hence excluding any opportunity for WiFi transmissions.

- *Case 3 - No Transmission on Licensed*: SBS operates on the unlicensed band only and shares the spectrum with WLAN by muting adaptively. This corresponds to previously suggested approaches such as the work proposed in [15, 17, 23, 24]. For this case, we specifically consider a muting pattern based on PF coexistence of SBS and WLAN on the unlicensed band which is similar to [24].

- *Case 4 - No Transmission on Unlicensed*: SBS operates on the licensed band only and shares the spectrum with MBS by muting adaptively. This corresponds to previously suggested approaches in the area of ICIC such as the work proposed in [170] based on ABS muting. For this case, we specifically consider a muting pattern based on PF coexistence of MBS and SBS on the licensed band.

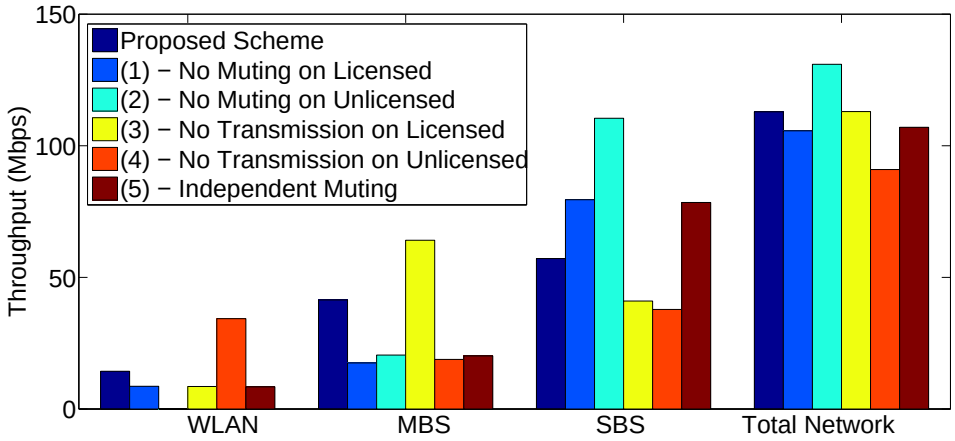

Figure 4.6. The aggregate throughput of the WLAN, MBS, SBS and total network for our proposed traffic balancing scheme in comparison with other approaches.

- *Case 5 - Independent Muting*: SBS operates on both bands, however, an independent mechanism is applied on each band for its coexistence with LTE and WLAN i.e., the coexistence of SBS and MBS on the licensed band and the coexistence of SBS and WLAN on the unlicensed band are solved separately. To realize this case, we consider two independent PF coexistence formulations for the muting of SBS on each of the licensed and unlicensed bands. In other words, when solving for α, we consider the WLAN and SBS throughput on the unlicensed band only, and when solving for β, we consider the MBS and SBS throughput on the licensed band only.

Note that cases 1 and 2, respectively, do not consider coexistence mechanisms on the licensed and unlicensed bands and thus are not practical solutions; however, we include them in our study for the sake of completeness.

Fig. 4.6 shows the throughput achieved by WLAN, MBS, SBS and the total network for our proposed scheme as well as the other five studied approaches; the corresponding Jain's fairness index $\mathcal{J}(s_i)$ values are given in Table 4.2. We can make the following observations from these results. First, the WLAN throughput can be improved when coexisting with LTE-LAA small cells on the unlicensed band by taking into account the transmission of LTE-LAA small cells on the licensed and unlicensed bands and considering a holistic approach for the allocation of the radio resources on both bands i.e., optimizing both bands jointly. This can be observed from Fig. 4.6 by comparing the total achieved throughput of WLAN for our proposed scheme with that of cases 1, 2, 3 and 5. Similarly, MBS throughput is higher with our proposed scheme compared to cases 1, 2, 4

and 5. Note that the WLAN and MBS throughputs will be, respectively, maximum when they exclusively use the unlicensed (case 4) and licensed bands (case 3), due to the absence of inter-technology interference in the former and lack of inter-tier interference in the latter. However, the total network throughput is the lowest for case 4; and case 3 results in a relatively unfair sharing of the radio resources as compared to our proposed scheme.

Second, considering an independent muting mechanism on the licensed and unlicensed bands (case 5) leads to performance degradation in terms of throughput and fairness, indicating that the effectiveness of our proposed traffic balancing scheme stems from its holistic nature. This is validated from Fig. 4.6 and Table 4.2 by comparing the total network throughput and Jain's fairness index of our approach to that of case 5 i.e., $\mathcal{J}(s_i)$=0.82 and 0.57 respectively and 5.5% improvement in the total network throughput. As another observation, the independent muting approach provides very close performance for MBS to case 4 due to the fact that α=1 and hence the optimization problem would be a function of the variable β only and would correspond to the sub-problem of the coexistence on the licensed band of case 5. Similar argument applies for the WLAN throughput of case 5 which is similar to that of case 3 (where β=0).

Third, our proposed traffic balancing scheme utilizes the radio resources in the most efficient way compared to the other studied schemes as it provides a better tradeoff between efficiency (throughput) $\eta(s_i)$ and fairness $\mathcal{J}(s_i)$. In terms of efficiency, case 2 provides the maximum total network throughput since SBS will be transmitting on both bands simultaneously, however, WLAN would not be given opportunities for transmission and hence this would result in the least value of $\mathcal{J}(s_i)$ (0.45) as the radio resources are not shared fairly among the different technologies. Note also that our proposed scheme provides similar throughput as case 3; the major contribution to overall throughput in case 3 comes from MBS throughput which is maximum due to its exclusive use of the licensed band. However, comparing Jain's index fairness of our approach to that of case 3, we observe that our scheme allocates the radio resources in a more fair way unlike case 3 that causes a degradation in the WLAN and SBS throughputs. In terms of fairness, case 4 provides the most fair allocation of the licensed and the unlicensed bands as $\mathcal{J}(s_i)$ is the closest to 1 but it comes at the expense of throughput efficiency; total network throughput is the lowest with case 4. The reason for this high value of $\mathcal{J}(s_i)$ is because WLAN would have more transmission opportunities

Table 4.2 Jain's fairness index for the UEs achieved throughput of our proposed scheme and the other five cases.

Cases	**Our scheme**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$\mathcal{J}(s_i)$	0.82	0.55	0.45	0.73	0.92	0.57

and hence its throughput would increase when using the channel exclusively as compared to sharing it with LTE-LAA SBS. On the other hand, the decrease in the value of $\eta(s_i)$ is due to the difference in the MAC layer throughputs with WiFi and LTE (65 Mbps and 75 Mbps respectively in our simulation setup) and the inter-tier interference level on the licensed band which results in the degradation of the SBS and MBS throughput.

4.6 Summary

In this chapter, we have presented a formulation of the holistic LTE-LAA SBS traffic balancing across the licensed and unlicensed bands as an optimization problem that seeks to achieve a proportional fair coexistence of WiFi STAs, SUEs and MUEs. We have derived a closed form solution for the aforementioned optimization problem and proposed a transmission mechanism for the operation of the LTE-LAA SBS on both bands. Results show that LTE-LAA SBS aided by our solution would switch between or aggregate the licensed and unlicensed bands based on the interference/traffic level and number of active UEs in each band. It also provides a better performance for WLAN when coexisting with LTE and an efficient utilization of the radio resources compared to alternative approaches from the literature as it allows a better tradeoff between maximizing the total network throughput and achieving fairness among all network flows. In what follows, we briefly discuss a couple of issues that warrant detailed exploration in future work:

- **Multiple channels:** Although we consider a single unlicensed channel, the proposed traffic balancing scheme can be extended to multiple unlicensed channels, each with a different muting variable $\{\alpha_1, ..., \alpha_c\}$, provided that the WiFi networks occupy disjoint channels (non-overlapping channels). Note that in such scenarios, the computational complexity increases due

to the increase in the number of variables and thus would make it hard to obtain an online solution. Therefore, one could potentially combine the proposed approach in this chapter with the channel selection learning scheme proposed in chapter 3 in a joint framework.

- **Hidden terminals:** Hidden and exposed terminals are a major problem in wireless networks and can result in a dramatic throughput degradation, if not managed. In essence, LTE use of unlicensed bands in the SDL mode gives rise to hidden terminal situations that need to be handled. In WLAN, this issue is addressed via the request-to-send/clear-to-send (RTS/CTS) messages; however, this method cannot be used for LTE-LAA since only DL transmissions are supported and hence SUEs are not able to transmit the CTS on the unlicensed spectrum. Therefore, to solve the hidden node problem, device-assisted enhancements need to be considered along with other existing mechanisms of the LTE system such as the periodic transmission of UE CSI/interference measurement over the licensed band. On the unlicensed band, a hidden terminal can be detected if SBS senses a good channel while the CSI report from the SUE shows a high interference value. This allows SBS to perform scheduling changes prior and during its operation on the unlicensed channel i.e., exclude the victim SUE for scheduling until its channel becomes idle and schedule other SUEs meanwhile. Alternatively, SBS may select another unlicensed channel to operate on [177].

Chapter 5

Interference Management for Cellular-Connected UAVs

5.1 Introduction

Cellular-connected UAVs will be an integral component of future wireless networks as evidenced by recent interest from academia, industry, and 3GPP standardizations [26–28,178–180]. As discussed earlier in Chapter 2, providing cellular connectivity to the UAVs offers several advantages to other current short-range wireless connections such as WiFi, bluetooth, and radio waves. Consequently, cellular-connected UAV-UEs will pave the way to new UAV-UE applications such as real-time video streaming and surveillance. Nevertheless, the ability of UAV-UEs to establish LoS connectivity to cellular base stations is both a blessing and a curse. On the one hand, it enables high-speed data access for the UAV-UEs. On the other hand, it can lead to substantial inter-cell mutual interference among the UAVs and to the ground users. As such, a wide-scale deployment of UAV-UEs is only possible if interference management challenges are addressed [26–28]. To this end, in this chapter, we propose an interference-aware path planning scheme for a network of cellular-connected UAVs. In particular, each UAV aims at achieving a tradeoff between maximizing energy efficiency and minimizing both wireless latency and the interference level caused on the ground network along its path. This in essence is realized by allowing each UAV to learn its optimal path, transmission power level, and cell association vector at different locations along its path. The work presented in this chapter has appeared in the Proceedings of the IEEE International Conference on Communications [36] and an

Table 5.1 Variables and notations.

Notation	Description
S	Number of ground BSs
Q	Number of ground UEs
J	Number of cellular-connected UAVs
K_s	Number of ground UEs served by BS s
N_s	Number of cellular-connected UAVs served by BS s
$C_{j,s}$	Number of RBs allocated to UAV j
$C_{q,s}$	Number of RBs allocated to UE q
A	Number of discretized unit areas
B_c	Bandwidth of resource block c
N_0	Noise level over a resource block
$\xi_{j,s,a}$	Path loss between UAV j at location a and BS s
$\xi_{q,s}$	Path loss between UE q and BS s
$\widehat{P}_{j,s,a}$	Total transmit power of UAV j at location a to BS s
$P_{j,s,c,a}$	Transmit power of UAV j at location a to BS s over RB c
$h_{j,s,c,a}$	Channel gain between UAV j and BS s on RB c at location a
$\lambda_{j,s}$	Average packet arrival rate traversing link (j,s)
$P_{q,s,c}$	Transmit power of UE q to its serving BS s on RB c
$h_{q,s,c}$	Channel gain between UE q and BS s on RB c
$\overline{P}_j$	Maximum transmission power of UAV j
L	Number of interferers
$\alpha_{j,a,b}$	Indicates whether or not a directed link is formed from a towards b for UAV j
$\beta_{j,s,a}$	Indicates whether or not UAV j is associated with BS s at location a
$\tau_{j,s,a}$	Latency over the UAV-BS wireless link (j,s)
$\Gamma_{j,s,c,a}$	SINR of UAV j to BS s on RB c at location a
$\overline{\Gamma}_j$	SINR threshold value for UAV j
$\boldsymbol{z}_j(t)$	Action taken by UAV j at t
$\boldsymbol{v}_j(t)$	Observed network state by UAV j at t
$\delta_{j,l,a}(t)$	Euclidean distance from UAV j at location a to BS l at t
$\theta_{j,l,a}(t)$	Orientation angle in the xy-plane from UAV j at location a to BS l at t
$\theta_{j,d_j,a}(t)$	Orientation angle in the xy-plane from UAV j at location a to its destination d_j at t
$\boldsymbol{\pi}_j(\boldsymbol{v}_j(t))$	Behavioral strategy of UAV j at state $\boldsymbol{v}_j(t)$
γ	Discount factor for delayed rewards
$\overline{u}$	Expected long term reward
$\boldsymbol{W}_{j,\text{in}}^{(n)}$	Input-to-reservoir matrix of UAV j at layer n at t
$\boldsymbol{W}_j^{(n)}$	Recurrent reservoir weight matrix for UAV j at layer n
$\boldsymbol{W}_{j,\text{out}}$	Reservoir-to-output matrix of UAV j for layer n
$N_{j,R}^{(n)}$	Number of internal units of the reservoir of UAV j at layer n
$N_{j,U}$	External input dimension of UAV j
$N_{j,L}$	Number of layers in the stack for UAV j
$\rho_j^{(n)}$	Spectral radius of UAV j at layer n
λ_j	Learning rate at UAV j
$\omega_j^{(n)}$	Leaking parameter at layer n
$y_j(t)$	Output of the deep ESN at t
$e(\boldsymbol{v}_j(t))$	Error function

extended journal version has been submitted to IEEE Transactions on Wireless Communications [37]. Table 5.1 provides a summary for the description of the main notations used in this chapter. Next, we give an overview on the related literature and then we present our problem statement and contribution.

5.1.1 Related Work

While some literature has recently studied the use of UAVs as mobile BSs [30–32, 46, 181–183], the performance analysis of cellular-connected UAV-UEs (*shorthanded hereinafter as UAVs*) remains relatively scarce [26–29]. For instance, in [26], the authors study the impact of UAVs on the UL performance of a ground LTE network. Meanwhile, the work in [27] uses measurements and ray tracing simulations to study the airborne connectivity requirements and propagation characteristics of UAVs. The authors in [28] analyze the coverage probability of the DL of a cellular network that serves both aerial and ground users. In [29], the authors consider a network consisting of both ground and aerial UEs and derive closed-form expressions for the coverage probability of the ground and drone UEs. Nevertheless, this prior art is limited to studying the impact that cellular-connected UAVs have on the ground network. Indeed, the existing literature [26–29] does not provide any concrete solution for optimizing the performance of a cellular network that serves both aerial and ground UEs in order to overcome the interference challenge that arises in this context. UAV trajectory optimization is essential in such scenarios. An online path planning that accounts for wireless metrics is vital and would, in essence, assist in addressing the aforementioned interference challenges along with new improvements in the design of the network, such as 3D frequency reuse. Such a path planning scheme allows the UAVs to adapt their movement based on the rate requirements of both aerial UAV-UEs and ground UEs, thus improving the overall network performance. The problem of UAV path planning has been studied mainly for non-UAV-UE applications [30–33] with [34] being the only work considering a cellular-connected UAV-UE scenario. In [30], the authors propose a distributed path planning algorithm for multiple UAVs to deliver delay-sensitive information to different ad-hoc nodes. The authors in [31] optimize a UAV's trajectory in an energy-efficient manner. The authors in [32] propose a mobility model that combines area coverage, network connectivity, and UAV energy constraints for path

planning. In [33], the authors propose a fog-networking-based system architecture to coordinate a network of UAVs for video services in sports events. However, despite being interesting, the body of work in [30–32] and [33] is restricted to UAVs as BSs and does not account for UAV-UEs and their associated interference challenges. Hence, the approaches proposed therein cannot readily be used for cellular-connected UAVs. On the other hand, the authors in [34] propose a path planning scheme for minimizing the time required by a cellular-connected UAV to reach its destination. Nevertheless, this work is limited to one UAV and does not account for the interference that cellular-connected UAVs cause on the ground network during their mission. Moreover, the work in [34] relies on offline optimization techniques that cannot adapt to the uncertainty and dynamics of a cellular network.

5.1.2 Problem Statement and Contribution

The main objective of this chapter is to develop an interference-aware path planning scheme for a network of cellular-connected UAVs. In essence, providing wireless cellular connectivity for UAV-UEs is contingent upon proper management of their resulting interference which mainly results from the fact that most UAV communication links are LoS dominated. Consequently, a challenging aspect in the implementation of the communication links of cellular-connected UAVs is to maintain high data rate, low delay and reliable network connection while minimizing the interference level caused on the ground network. As such, an interference-aware online path planning scheme allows the UAVs to adapt their paths based on the dynamics of the network instead of seeking their shortest paths to reach their corresponding destinations. Such a scheme improves the total network performance and allows both the UAV-UEs and the ground UEs to achieve higher data rates.

The main contribution of this chapter is therefore a novel deep RL framework based on ESN cells for optimizing the trajectories of multiple cellular-connected UAVs in an online manner. This framework will allow cellular-connected UAVs to minimize the interference they cause on the ground network as well as their wireless transmission latency. To realize this, we propose a dynamic noncooperative game in which the players are the UAVs and the objective of each UAV is to *autonomously* and *jointly* learn its path, transmit power level, and associa-

tion vector. For our proposed game, the UAV's cell association vector, trajectory optimization, and transmit power level are closely coupled with each other and their optimal values vary based on the dynamics of the network. Therefore, a major challenge in this game is the need for each UAV to have full knowledge of the ground network topology, ground UEs service requirements, and other UAVs' locations. Consequently, to solve this game, we propose a deep RL ESN-based algorithm, using which the UAVs can predict the dynamics of the network and subsequently determine their optimal paths as well as the allocation of their radio resources along their paths. Unlike previous studies which are either centralized or rely on the coordination among UAVs, our approach is based on a self-organizing path planning and resource allocation scheme. In essence, two important features of our proposed algorithm are *adaptation* and *generalization*. Indeed, UAVs can take decisions for *unseen* network states, based on the reward they got from previous states. This is mainly due to the use of ESN cells which enable the UAVs to retain their previous memory states. We have shown that the proposed algorithm reaches a SPNE upon convergence. Moreover, upper and lower bounds on the UAVs' altitudes, that guarantee a maximum interference level on the ground network and a maximum wireless transmission delay for the UAV, have been derived. Simulation results show that the proposed approach improves the tradeoff between energy efficiency, wireless latency, and the interference level caused on the ground network. Results also show that each UAV's altitude is a function of the ground network density and the UAV's objective function and is an important factor in achieving the UAV's target.

The rest of this chapter is organized as follows. Section 5.2 presents the system model. Section 5.3 describes the proposed noncooperative game model. The deep RL ESN-based algorithm is proposed in Section 5.4. In Section 5.5, simulation results are analyzed. Finally, a summary is provided in Section 5.6.

5.2 System Model

Consider the UL of a wireless cellular network composed of a set $\mathcal{S}$ of S ground BSs, a set $\mathcal{Q}$ of Q ground UEs, and a set $\mathcal{J}$ of J cellular-connected UAVs. The UL is defined as the link from UE q or UAV j to BS s. Each BS $s \in \mathcal{S}$ serves a set $\mathcal{K}_s \subseteq \mathcal{Q}$ of K_s UEs and a set $\mathcal{N}_s \subseteq \mathcal{J}$ of N_s cellular-connected UAVs. The total system bandwidth, B, is divided into a set $\mathcal{C}$ of C RBs. Each UAV $j \in \mathcal{N}_s$ is allocated a set $\mathcal{C}_{j,s} \subseteq \mathcal{C}$ of $C_{j,s}$ RBs and each UE $q \in \mathcal{K}_s$ is allocated a set

$\mathcal{C}_{q,s} \subseteq \mathcal{C}$ of $C_{q,s}$ RBs by its serving BS s. At each BS s, a particular RB $c \in \mathcal{C}$ is allocated to *at most* one UAV $j \in \mathcal{N}_s$, or UE $q \in \mathcal{K}_s$.

An airborne IoT is considered in which the UAVs are equipped with different IoT devices, such as cameras, sensors, and GPS that can be used for various applications such as surveillance, monitoring, delivery, and real-time video streaming. The 3D coordinates of each UAV $j \in \mathcal{J}$ and each ground user $q \in \mathcal{Q}$ are, respectively, (x_j, y_j, h_j) and $(x_q, y_q, 0)$. All UAVs are assumed to fly at a fixed altitude h_j above the ground (as done in [31,34,184,185]) while the horizonal coordinates (x_j, y_j) of each UAV j vary in time. Each UAV j needs to move from an initial location o_j to a final destination d_j while transmitting *online* its mission-related data such as sensor recordings, video streams, and location updates. We assume that the initial and final locations of each UAV are pre-determined based on its mission objectives.

For ease of exposition, we consider a virtual grid for the mobility of the UAVs. We discretize the space into a set $\mathcal{A}$ of A equally sized unit areas. The UAVs move along the center of the areas $c_a = (x_a, y_a, z_a)$, which yields a finite set of possible paths $\boldsymbol{p}_j$ for each UAV j. The path $\boldsymbol{p}_j$ of each UAV j is defined as a sequence of area units $\boldsymbol{p}_j = (a_1, a_2, \cdots, a_l)$ such that $a_1 = o_j$ and $a_l = d_j$. The area size of the discretized area units $(a_1, a_2, \cdots, a_A) \in \mathcal{A}$ is chosen to be sufficiently small such that the UAVs' locations can be assumed to be approximately constant within each area even at the maximum UAV's speed, as commonly done in the literature [184]. We assume a constant speed $0 < V_j \leq \widehat{V}_j$ for each UAV where $\widehat{V}_j$ is the maximum speed of UAV j. Therefore, the time required by each UAV to travel between any two unit areas is constant.

5.2.1 Channel Models

We consider the sub-6 GHz band and the free-space path loss model for the UAV-BS data link. The path loss between UAV j at location a and BS s, $\xi_{j,s,a}$, is given by [78]:

$$\xi_{j,s,a}(\text{dB}) = 20 \log_{10}(d_{j,s,a}) + 20 \log_{10}(\hat{f}) - 147.55, \tag{5.1}$$

where $\hat{f}$ is the system center frequency and $d_{j,s,a}$ is the Euclidean distance between UAV j at location a and BS s. We consider a Rician distribution for modeling the small-scale fading between UAV j and ground BS s thus accounting for the LoS and multipath scatterers that can be experienced at the BS. In particular,

adopting the Rician channel model for the UAV-BS link is validated by the fact that the channel between a given UAV and a ground BS is mainly dominated by a LoS link [31]. We assume that the Doppler spread due to the mobility of the UAVs is compensated for based on existing techniques such as frequency synchronization using a phase-locked loop [186] as done in [31] and [184].

For the terrestrial UE-BS links, we consider a Rayleigh fading channel. For a carrier frequency, $\hat{f}$, of 2 GHz, the path loss between UE q and BS s is given by [187]:

$$\zeta_{q,s}(\text{dB}) = 15.3 + 37.6 \log_{10}(d_{q,s}), \tag{5.2}$$

where $d_{q,s}$ is the Euclidean distance between UE q and BS s.

The average SINR, $\Gamma_{j,s,c,a}$, of the UAV-BS link between UAV j at location a ($a \in \mathcal{A}$) and BS s over RB c will be:

$$\Gamma_{j,s,c,a} = \frac{P_{j,s,c,a} h_{j,s,c,a}}{I_{j,s,c} + B_c N_0}, \tag{5.3}$$

where $P_{j,s,c,a} = \widehat{P}_{j,s,a}/C_{j,s}$ is the transmit power of UAV j at location a to BS s over RB c and $\widehat{P}_{j,s,a}$ is the total transmit power of UAV j to BS s at location a. Here, the total transmit power of UAV j is assumed to be distributed uniformly among all of its associated RBs. $h_{j,s,c,a} = g_{j,s,c,a} 10^{-\xi_{j,s,a}/10}$ is the channel gain between UAV j and BS s on RB c at location a where $g_{j,s,c,a}$ is the Rician fading parameter. N_0 is the noise power spectral density and B_c is the bandwidth of an RB c. $I_{j,s,c} = \sum_{r=1,r\neq s}^{S} (\sum_{k=1}^{K_r} P_{k,r,c} h_{k,s,c} + \sum_{n=1}^{N_r} P_{n,r,c,a'} h_{n,s,c,a'})$ is the total interference power on UAV j at BS s when transmitting over RB c, where $\sum_{r=1,r\neq s}^{S} \sum_{k=1}^{K_r} P_{k,r,c} h_{k,s,c}$ and $\sum_{r=1,r\neq s}^{S} \sum_{n=1}^{N_r} P_{n,r,c,a'} h_{n,s,c,a'}$ correspond, respectively, to the interference from the K_r UEs and the N_r UAVs (at their respective locations a') connected to neighbouring BSs r and transmitting using the same RB c as UAV j. $h_{k,s,c} = m_{k,s,c} 10^{-\zeta_{k,s}/10}$ is the channel gain between UE k and BS s on RB c where $m_{k,s,c}$ is the Rayleigh fading parameter. Therefore, the achievable data rate of UAV j at location a associated with BS s can be defined as $R_{j,s,a} = \sum_{c=1}^{C_{j,s}} B_c \log_2(1 + \Gamma_{j,s,c,a})$.

Given the achievable data rate of UAV j, and assuming that each UAV is an M/D/1 queueing system (a queue with Poisson arrivals, constant service times and 1 server), the corresponding latency over the UAV-BS wireless link is given by [188]:

$$\tau_{j,s,a} = \frac{\lambda_{j,s}}{2\mu_{j,s,a}(\mu_{j,s,a} - \lambda_{j,s})} + \frac{1}{\mu_{j,s,a}}, \tag{5.4}$$

where $\lambda_{j,s}$ is the average packet arrival rate (packets/s) traversing link (j, s) and originating from UAV j. $\mu_{j,s,a} = R_{j,s,a}/\nu$ is the service rate over link (j, s) at location a where ν is the packet size. Here, it is important to note that, given that it is difficult to find the delay under very rapid fading changes, it is reasonable to assume that the fading is essentially an average fading of the channel that an arbitrary packet will experience. This provides a tractable way to model the delay. This assumption (and the delay expression in 6.11) has been commonly adopted in the literature, such as in [189] Equation 27 and [190] Equation 6. Moreover, it is important to note that, UAV communication links are mainly LoS dominated. As such, one can assume that the channel variations are not too rapid between packets.

On the other hand, the achievable data rate for a ground UE q served by BS s is given by:

$$R_{q,s} = \sum_{c=1}^{C_{q,s}} B_c \log_2\left(1 + \frac{P_{q,s,c} h_{q,s,c}}{I_{q,s,c} + B_c N_0}\right), \tag{5.5}$$

where $h_{q,s,c} = m_{q,s,c} 10^{-\zeta_{q,s}/10}$ is the channel gain between UE q and BS s on RB c and $m_{q,s,c}$ is the Rayleigh fading parameter. $P_{q,s,c} = \widehat{P}_{q,s}/C_{q,s}$ is the transmit power of UE q to its serving BS s on RB c and $\widehat{P}_{q,s}$ is the total transmit power of UE q. Here, we also consider equal power allocation among the allocated RBs for the ground UEs. $I_{q,s,c} = \sum_{r=1,r\neq s}^{S}(\sum_{k=1}^{K_r} P_{k,r,c} h_{k,s,c} + \sum_{n=1}^{N_r} P_{n,r,c,a'} h_{n,s,c,a'})$ is the total interference power experienced by UE q at BS s on RB c where $\sum_{r=1,r\neq s}^{S} \sum_{k=1}^{K_r} P_{k,r,c} h_{k,s,c}$ and $\sum_{r=1,r\neq s}^{S} \sum_{n=1}^{N_r} P_{n,r,c,a'} h_{n,s,c,a'}$ correspond, respectively, to the interference from the K_r UEs and the N_r UAVs (at their respective locations a') associated with the neighbouring BSs r and transmitting using the same RB c as UE q.

5.2.2 Problem Formulation

Our objective is to find the optimal path for each UAV j based on its mission objectives as well as its interference on the ground network. Thus, we seek to minimize: a) the interference level that each UAV causes on the ground UEs and other UAVs, b) the transmission delay over the wireless link, and c) the time needed to reach the destination. To realize this, we optimize the paths of the UAVs jointly with the cell association vector and power control at each location $a \in \mathcal{A}$ along each UAV's path. We consider a directed graph $G_j = (\mathcal{V}, \mathcal{E}_j)$ for

each UAV j where $\mathcal{V}$ is the set of vertices corresponding to the centers of the unit areas $a \in \mathcal{A}$ and $\mathcal{E}_j$ is the set of edges formed along the path of UAV j. We let $\widehat{\boldsymbol{P}}$ be the transmission power vector with each element $\widehat{P}_{j,s,a} \in [0, \overline{P}_j]$ being the transmission power level of UAV j to its serving BS s at location a where $\overline{P}_j$ is the maximum transmission power of UAV j. $\boldsymbol{\alpha}$ is the path formation vector with each element $\alpha_{j,a,b} \in \{0,1\}$ indicating whether or not a directed link is formed from area a towards area b for UAV j, i.e., if UAV j moves from a to b along its path. $\boldsymbol{\beta}$ is the UAV-BS association vector with each element $\beta_{j,s,a} \in \{0,1\}$ denoting whether or not UAV j is associated with BS s at location a. Next, we present our optimization problem whose goal is to determine the path of each UAV along with its cell association vector and its transmit power level at each location a along its path $\boldsymbol{p}_j$:

$$\min_{\widehat{\boldsymbol{P}},\boldsymbol{\alpha},\boldsymbol{\beta}} \vartheta \sum_{j=1}^{J}\sum_{s=1}^{S}\sum_{c=1}^{C_{j,s}}\sum_{a=1}^{A}\sum_{r=1,r\neq s}^{S}\frac{\widehat{P}_{j,s,a}h_{j,r,c,a}}{C_{j,s}} + \varpi\sum_{j=1}^{J}\sum_{a=1}^{A}\sum_{b=1,b\neq a}^{A}\alpha_{j,a,b} + \phi\sum_{j=1}^{J}\sum_{s=1}^{S}\sum_{a=1}^{A}\beta_{j,s,a}\tau_{j,s,a}, \quad (5.6)$$

$$\sum_{b=1,b\neq a}^{A}\alpha_{j,b,a} \leq 1 \;\; \forall j \in \mathcal{J}, a \in \mathcal{A}, \quad (5.7)$$

$$\sum_{a=1,a\neq o_j}^{A}\alpha_{j,o_j,a}=1 \;\; \forall j \in \mathcal{J}, \sum_{a=1,a\neq d_j}^{A}\alpha_{j,a,d_j}=1 \;\; \forall j \in \mathcal{J}, \quad (5.8)$$

$$\sum_{a=1,a\neq b}^{A}\alpha_{j,a,b} - \sum_{f=1,f\neq b}^{A}\alpha_{j,b,f}=0 \; \forall j \in \mathcal{J}, b \in \mathcal{A} \; (b \neq o_j, b \neq d_j), \quad (5.9)$$

$$\widehat{P}_{j,s,a} \geq \sum_{b=1,b\neq a}^{A}\alpha_{j,b,a} \;\; \forall j \in \mathcal{J}, s \in \mathcal{S}, a \in \mathcal{A}, \quad (5.10)$$

$$\widehat{P}_{j,s,a} \geq \beta_{j,s,a} \;\; \forall j \in \mathcal{J}, s \in \mathcal{S}, a \in \mathcal{A}, \quad (5.11)$$

$$\sum_{s=1}^{S}\beta_{j,s,a} - \sum_{b=1,b\neq a}^{A}\alpha_{j,b,a} = 0 \quad \forall j \in \mathcal{J}, a \in A, \quad (5.12)$$

$$\sum_{c=1}^{C_{j,s}}\Gamma_{j,s,c,a} \geq \beta_{j,s,a}\overline{\Gamma}_j \quad \forall j \in \mathcal{J}, s \in \mathcal{S}, a \in \mathcal{A}, \quad (5.13)$$

$$0 \leq \widehat{P}_{j,s,a} \leq \overline{P}_j \ \ \forall j \in \mathcal{J}, s \in \mathcal{S}, a \in \mathcal{A}, \tag{5.14}$$

$$\alpha_{j,a,b} \in \{0,1\}, \beta_{j,s,a} \in \{0,1\} \ \ \forall j \in \mathcal{J}, s \in \mathcal{S}, \ a, b \in \mathcal{A}. \tag{5.15}$$

The objective function in (5.6) captures the total interference level that the UAVs cause on neighbouring BSs along their paths, the length of the paths of the UAVs, and their wireless transmission delay. ϑ, ϖ and ϕ are multi-objective weights used to control the tradeoff between the three considered metrics. These weights can be adjusted to meet the requirements of each UAV's mission. For instance, the time to reach the destination is critical in search and rescue applications while the latency is important for online video streaming applications. (5.7) guarantees that each area a is visited by UAV j at most once along its path $\boldsymbol{p}_j$. (5.8) guarantees that the trajectory of each UAV j starts at its initial location o_j and ends at its final destination d_j. (5.9) guarantees that if UAV j visits area b, it should also leave from area b ($b \neq o_j, b \neq d_j$). (5.10) and (5.11) guarantee that UAV j transmits to BS s at area a with power $\widehat{P}_{j,s,a} > 0$ only if UAV j visits area a, i.e., $a \in \boldsymbol{p}_j$ and such that j is associated with BS s at location a. (5.12) guarantees that each UAV j is associated with one BS s at each location a along its path $\boldsymbol{p}_j$. (5.13) guarantees a lower limit, $\overline{\Gamma}_j$, for the SINR value $\Gamma_{j,s,c,a}$ of the transmission link from UAV j to BS s on RB c at each location a, $a \in \boldsymbol{p}_j$. This, in turn, ensures successful decoding of the transmitted packets at the serving BS. The value of $\overline{\Gamma}_j$ is application and mission specific. Note that the SINR check at each location a is valid for our problem since we consider small-sized area units. (5.14) and (5.15) are the feasibility constraints. The formulated optimization problem is MINLP, which is computationally complex to solve for large networks.

To address this challenge, we adopt a distributed approach in which each UAV decides autonomously on its next path location along with its corresponding transmit power and association vector. In fact, a centralized approach requires control signals to be transmitted to the UAVs at all time. This might incur high round-trip latencies that are not desirable for real-time applications such as online video streaming. Further, a centralized approach requires a central entity to have full knowledge of the current state of the network and the ability to communicate with all UAVs at all time. However, this might not be feasible in case the UAVs belong to different operators or in scenarios in which the environment changes dynamically. Therefore, we next propose a distributed approach for each UAV j to learn its path $\boldsymbol{p}_j$ along with its transmission power level and association vector at each location a along its path in an autonomous and online manner.

5.3 Towards a Self-Organizing Network of an Airborne Internet of Things

5.3.1 Game-Theoretic Formulation

Our objective is to develop a distributed approach that allows each UAV to take actions in an autonomous and online manner. For this purpose, we model the multi-agent path planning problem as a finite dynamic noncooperative game model $\mathcal{G}$ with perfect information [127]. Formally, we define the game as $\mathcal{G} = (\mathcal{J}, \mathcal{T}, \mathcal{Z}_j, \mathcal{V}_j, \Pi_j, u_j)$ with the set $\mathcal{J}$ of UAVs being the agents. $\mathcal{T}$ is a finite set of stages which correspond to the steps required for all UAVs to reach their sought destinations. $\mathcal{Z}_j$ is the set of actions that can be taken by UAV j at each $t \in \mathcal{T}$, $\mathcal{V}_j$ is the set of all observed network states by UAV j up to stage T, Π_j is a set of probability distributions defined over all $z_j \in \mathcal{Z}_j$, and u_j is the payoff function of UAV j. At each stage $t \in \mathcal{T}$, the UAVs take actions simultaneously. In particular, each UAV j aims at determining its path $\boldsymbol{p}_j$ to its destination along with its optimal transmission power and cell association vector for each location $a \in \mathcal{A}$ along its path $\boldsymbol{p}_j$. Therefore, at each t, UAV j chooses an action $z_j(t) \in \mathcal{Z}_j$ composed of the tuple $\boldsymbol{z}_j(t) = (\boldsymbol{a}_j(t), \widehat{P}_{j,s,a}(t), \boldsymbol{\beta}_{j,s,a}(t))$, where $\boldsymbol{a}_j(t)$={left, right, forward, backward, no movement} corresponds to a fixed step size, $\widetilde{a}_j$, in a given direction. $\widehat{P}_{j,s,a}(t) = [\widehat{P}_1, \widehat{P}_2, \cdots, \widehat{P}_O]$ corresponds to O different maximum transmit power levels for each UAV j and $\boldsymbol{\beta}_{j,s,a}(t)$ is the UAV-BS association vector.

For each UAV j, let $\mathcal{L}_j$ be the set of its L_j nearest BSs. The observed network state by UAV j at stage t, $\boldsymbol{v}_j(t) \in \mathcal{V}_j$, is:

$$\boldsymbol{v}_j(t) = \left[\{\delta_{j,l,a}(t), \theta_{j,l,a}(t)\}_{l=1}^{L_j}, \theta_{j,d_j,a}(t), \{x_j(t), y_j(t)\}_{j \in \mathcal{J}} \right], \tag{5.16}$$

where $\delta_{j,l,a}(t)$ is the Euclidean distance from UAV j at location a to BS l at stage t, $\theta_{j,l,a}$ is the orientation angle in the xy-plane from UAV j at location a to BS l defined as $\tan^{-1}(\Delta y_{j,l}/\Delta x_{j,l})$ [191] where $\Delta y_{j,l}$ and $\Delta x_{j,l}$ correspond to the difference in the x and y coordinates of UAV j and BS l, $\theta_{j,d_j,a}$ is the orientation angle in the xy-plane from UAV j at location a to its destination d_j defined as $\tan^{-1}(\Delta y_{j,d_j}/\Delta x_{j,d_j})$, and $\{x_j(t), y_j(t)\}_{j \in \mathcal{J}}$ are the horizonal coordinates of all UAVs at stage t. For our model, we consider different range intervals for mapping each of the orientation angle and distance values, respectively, into different

states.

Moreover, based on the optimization problem defined in (5.6)-(5.15) and by incorporating the Lagrangian penalty method into the utility function definition for the SINR constraint (5.13), the resulting utility function for UAV j at stage t, $u_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t), \boldsymbol{z}_{-j}(t))$, will be given by:

$$u_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t), \boldsymbol{z}_{-j}(t))= \begin{cases} \Phi(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))+C, \text{ if } \delta_{j,d_j,a}(t) < \delta_{j,d_j,a'}(t-1), \\ \Phi(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)), \text{ if } \delta_{j,d_j,a}(t) = \delta_{j,d_j,a'}(t-1), \\ \Phi(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))\text{-}C, \text{ if } \delta_{j,d_j,a}(t) > \delta_{j,d_j,a'}(t-1), \end{cases} \tag{5.17}$$

where $\Phi(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ is defined as:

$$\Phi(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))=-\vartheta' \sum_{c=1}^{C_{j,s}(t)} \sum_{r=1,r\neq s}^{S} \frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t))h_{j,r,c,a}(t)}{C_{j,s}(t)} - \phi' \tau_{j,s,a}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) - \varsigma(\min(0, \sum_{c=1}^{C_{j,s}(t)} \Gamma_{j,s,c,a}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) - \overline{\Gamma}_j))^2, \tag{5.18}$$

subject to (5.7)-(5.12), (5.14) and (5.15). ς is the penalty coefficient for (5.13) and C is a constant parameter. a' and a are the locations of UAV j at $(t-1)$ and t where $\delta_{j,d_j,a}$ is the distance between UAV j and its destination d_j. It is worth noting here that the action space of each UAV j and, thus, the complexity of the proposed game $\mathcal{G}$ increases exponentially when updating the 3D coordinates of the UAVs. Nevertheless, each UAV's altitude must be bounded in order to guarantee an SINR threshold for the UAV and a minimum achievable data rate for the ground UEs. Next, we derive an upper and lower bound for the optimal altitude of any given UAV j based on the proposed utility function in (5.17). In essence, such bounds are valid for all values of the multi-objective weights ϑ', ϕ', and ς.

Theorem 2. For all values of ϑ', ϕ', and ς, a given network state $\boldsymbol{v}_j(t)$, and a particular action $\boldsymbol{z}_j(t)$, the upper and lower bounds for the altitude of UAV j are, respectively, given by:

$$h_j^{\max}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) = \min(\psi, \max(\chi, \hat{h}_j^{\max}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)))), \tag{5.19}$$

$$h_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) = \min(\psi, \max(\chi, \hat{h}_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)))), \tag{5.20}$$

where χ and ψ correspond, respectively, to the minimum and maximum altitudes at which a UAV can fly. $\hat{h}_j^{\max}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ and $\hat{h}_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ are expressed as:

$$\hat{h}_j^{\max}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) = \sqrt{\frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t))}{C_{j,s}(t)\cdot\overline{\Gamma}_j\cdot\left(\frac{4\pi\hat{f}}{\hat{c}}\right)^2}\cdot\sum_{c=1}^{C_{j,s}(t)}\frac{g_{j,s,c,a}(t)}{I_{j,s,c}(t)+B_cN_0}-(x_j-x_s)^2-(y_j-y_s)^2}, \quad (5.21)$$

and

$$\hat{h}_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) = \max_r \hat{h}_{j,r}^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)), \quad (5.22)$$

where $\hat{h}_{j,r}^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ is the minimum altitude that UAV j should operate at with respect to a particular neighbouring BS r and is expressed as:

$$\hat{h}_{j,r}^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) = \sqrt{\frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t))\cdot\sum_{c=1}^{C_{j,s}(t)}g_{j,r,c,a}(t)}{C_{j,s}(t)\cdot\left(\frac{4\pi\hat{f}}{\hat{c}}\right)^2\cdot\sum_{c=1}^{C_{j,s}(t)}\bar{I}_{j,r,c,a}}-(x_j-x_r)^2-(y_j-y_r)^2}, \quad (5.23)$$

Proof. See Appendix A.2. □

From the above theorem, we can deduce that the optimal altitude of the UAVs is a function of their objective function, location of the ground BSs, network design parameters, and the interference level from other UEs and UAVs in the network. Therefore, at each time step t, UAV j would adjust its altitude level based on the values of $h_j^{\max}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)$ and $h_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)$ thus adapting to the dynamics of the network. In essence, the derived upper and lower bounds for the optimal altitude of the UAVs allows a reduction of the action space of game $\mathcal{G}$ thus simplifying the process needed for the UAVs to find a solution, i.e., equilibrium, of the game. Next, we analyze the equilibrium point of the proposed game $\mathcal{G}$.

5.3.2 Equilibrium Analysis

For our game $\mathcal{G}$, we are interested in studying the SPNE in behavioral strategies. An SPNE is a profile of strategies which induces a NE on every subgame of the original game. Moreover, a *behavioral strategy* allows each UAV to assign independent probabilities to the set of actions at each network state that is independent

across different network states. Here, note that there always exists at least one SPNE for any finite horizon extensive game with perfect information [Selten's Theorem] [192]. Let $\boldsymbol{\pi}_j(\boldsymbol{v}_j(t)) = (\pi_{j,z_1}(\boldsymbol{v}_j(t)), \pi_{j,z_2}(\boldsymbol{v}_j(t)), \cdots, \pi_{j,\boldsymbol{z}_{|Z_j|}}(\boldsymbol{v}_j(t))) \in \Pi_j$ be the behavioral strategy of UAV j at state $\boldsymbol{v}_j(t)$ and let $\Delta(\mathcal{Z})$ be the set of all probability distributions over the action space $\mathcal{Z}$. Next, we define the notion of an SPNE.

Definition 3. A behavioral strategy $(\boldsymbol{\pi}_1^*(\boldsymbol{v}_j(t)), \cdots, \boldsymbol{\pi}_J^*(\boldsymbol{v}_j(t))) = (\boldsymbol{\pi}_j^*(\boldsymbol{v}_j(t)), \boldsymbol{\pi}_{-j}^*(\boldsymbol{v}_j(t)))$ constitutes a *subgame perfect Nash equilibrium* if, $\forall j \in \mathcal{J}$, $\forall t \in \mathcal{T}$ and $\forall \boldsymbol{\pi}_j(\boldsymbol{v}_j(t)) \in \Delta(\mathcal{Z})$, $\overline{u}_j(\boldsymbol{\pi}_j^*(\boldsymbol{v}_j(t)), \boldsymbol{\pi}_{-j}^*(\boldsymbol{v}_j(t))) \geq \overline{u}_j(\boldsymbol{\pi}_j(\boldsymbol{v}_j(t)), \boldsymbol{\pi}_{-j}^*(\boldsymbol{v}_j(t)))$.

Therefore, at each state $\boldsymbol{v}_j(t)$ and stage t, the goal of each UAV j is to maximize its expected sum of discounted rewards, which is computed as the summation of the immediate reward for a given state along with the expected discounted utility of the next states:

$$\begin{aligned}
&\overline{u}(\boldsymbol{v}_j(t), \boldsymbol{\pi}_j(\boldsymbol{v}_j(t)), \boldsymbol{\pi}_{-j}(\boldsymbol{v}_j(t))) \\
&\quad = \mathbb{E}_{\boldsymbol{\pi}_j(t)} \left\{ \sum_{l=0}^{\infty} \gamma^l u_j(\boldsymbol{v}_j(t+l), \boldsymbol{z}_j(t+l), \boldsymbol{z}_{-j}(t+l)) | \boldsymbol{v}_{j,0} = \boldsymbol{v}_j \right\} \\
&\quad = \sum_{\boldsymbol{z} \in \mathcal{Z}} \sum_{l=0}^{\infty} \gamma^l u_j(\boldsymbol{v}_j(t+l), \boldsymbol{z}_j(t+l), \boldsymbol{z}_{-j}(t+l)) \prod_{j=1}^{J} \pi_{j,z_j}(\boldsymbol{v}_j(t+l)), \quad (5.24)
\end{aligned}$$

where $\gamma^l \in (0,1)$ is a discount factor for delayed rewards and $\mathbb{E}_{\boldsymbol{\pi}_j(\boldsymbol{v}_j(t))}$ denotes an expectation over trajectories of states and actions, in which actions are selected according to $\boldsymbol{\pi}_j(\boldsymbol{v}_j(t))$. Here, $\boldsymbol{u}_j$ is the short-term reward for being in state $\boldsymbol{v}_j$ and $\overline{\boldsymbol{u}}_j$ is the expected long-term total reward from state $\boldsymbol{v}_j$ onwards.

Here, note that the UAV's cell association vector, trajectory optimization, and transmit power level are closely coupled with each other and their corresponding optimal values vary based on the UAVs' objectives. In a multi-UAV network, each UAV must have full knowledge of the future reward functions at each information set and thus for all future network states in order to find the SPNE. This in turn necessitates knowledge of all possible future actions of all UAVs in the network and becomes challenging as the number of UAVs increases. To address this challenge, we rely on deep RNNs [3]. In essence, RNNs exhibit dynamic temporal behavior and are characterized by their adaptive memory that enables them to store necessary previous state information to predict future actions. On the other hand, deep neural networks are capable of dealing with large datasets.

Therefore, next, we develop a novel deep RL based on ESNs, a special kind of RNN, for solving the SPNE of our game $\mathcal{G}$.

5.4 Deep Reinforcement Learning for Online Path Planning and Resource Management

In this section, we first introduce a deep ESN-based architecture that allows the UAVs to store previous states whenever needed while being able to learn future network states. Then, we propose an RL algorithm based on the proposed deep ESN architecture to learn an SPNE for our proposed game.

5.4.1 Deep ESN Architecture

ESNs are a new type of RNNs with feedback connections that belong to the family of reservoir computing [3]. An ESN is composed of an input weight matrix $\boldsymbol{W}_{\text{in}}$, a recurrent matrix $\boldsymbol{W}$, and an output weight matrix $\boldsymbol{W}_{\text{out}}$. Because only the output weights are altered, ESN training is typically quick and computationally efficient compared to training other RNNs. Moreover, multiple non-linear reservoir layers can be stacked on top of each other resulting in a *deep ESN architecture*. Deep ESNs exploit the advantages of a hierarchical temporal feature representation at different levels of abstraction while preserving the RC training efficiency. They can learn data representations at different levels of abstraction, hence disentangling the difficulties in modeling complex tasks by representing them in terms of simpler ones hierarchically. Let $N_{j,R}^{(n)}$ be the number of internal units of the reservoir of UAV j at layer n, $N_{j,U}$ be the external input dimension of UAV j and $N_{j,L}$ be the number of layers in the stack for UAV j. Next, we define the following ESN components:

- $\boldsymbol{v}_j(t) \in \mathbb{R}^{N_{j,U}}$ the external input of UAV j at stage t which effectively corresponds to the current network state,
- $\boldsymbol{x}_j^{(n)}(t) \in \mathbb{R}^{N_{j,R}^{(n)}}$ as the state of the reservoir of UAV j at layer n at stage t,
- $\boldsymbol{W}_{j,\text{in}}^{(n)}$ as the input-to-reservoir matrix of UAV j at layer n, where $\boldsymbol{W}_{j,\text{in}}^{(n)} \in \mathbb{R}^{N_{j,R}^{(n)} \times N_{j,U}}$ for $n = 1$, and $\boldsymbol{W}_{j,\text{in}}^{(n)} \in \mathbb{R}^{N_{j,R}^{(n)} \times N_{j,R}^{(n-1)}}$ for $n > 1$,

- $\boldsymbol{W}_j^{(n)} \in \mathbb{R}^{N_{j,R}^{(n)} \times N_{j,R}^{(n)}}$ as the recurrent reservoir weight matrix for UAV j at layer n,
- $\boldsymbol{W}_{j,\text{out}} \in \mathbb{R}^{|\mathcal{Z}_j| \times (N_{j,U} + \sum_n N_{j,R}^{(n)})}$ as the reservoir-to-output matrix of UAV j for layer n only.

The objective of the deep ESN architecture is to approximate a function $\boldsymbol{F}_j = (F_j^1, F_j^2, \cdots, F_j^{N_{j,L}})$ for learning an SPNE for each UAV j at each stage t. For each $n = 1, 2, \cdots, N_{j,L}$, the function $F_j^{(n)}$ describes the evolution of the state of the reservoir at layer n, i.e., $\boldsymbol{x_j^{(n)}}(t) = F_j^{(n)}(\boldsymbol{v}_j(t), \boldsymbol{x}_j^{(n)}(t-1))$ for $n = 1$ and $\boldsymbol{x_j^{(n)}}(t) = F_j^{(n)}(\boldsymbol{x}_j^{(n-1)}(t), \boldsymbol{x}_j^{(n)}(t-1))$ for $n > 1$. $\boldsymbol{W}_{j,\text{out}}$ and $\boldsymbol{x}_j^{(n)}(t)$ are initialized to zero while $\boldsymbol{W}_{j,\text{in}}^{(n)}$ and $\boldsymbol{W}_j^{(n)}$ are randomly generated. Note that although the dynamic reservoir is initially generated randomly, it is combined later with the external input, $\boldsymbol{v}_j(t)$, in order to store the network states and with the trained output matrix, $\boldsymbol{W}_{j,\text{out}}$, so that it can approximate the reward function. Moreover, the spectral radius of $\boldsymbol{W}_j^{(n)}$ (i.e., the largest eigenvalue in absolute value), $\rho_j^{(n)}$, must be strictly smaller than 1 to guarantee the stability of the reservoir [125]. In fact, the value of $\rho_j^{(n)}$ is related to the variable memory length of the reservoir that enables the proposed deep ESN framework to store necessary previous state information, with larger values of $\rho_j^{(n)}$ resulting in longer memory length.

We next define the deep ESN components: the input and reward functions. For each deep ESN of UAV j, we distinguish between two types of inputs: external input, $\boldsymbol{v}_j(t)$, that is fed to the first layer of the deep ESN and corresponds to the current state of the network and input that is fed to all other layers for $n > 1$. For our proposed deep ESN, the input to any layer $n > 1$ at stage t corresponds to the state of the previous layer, $\boldsymbol{x}_j^{(n-1)}(t)$. Define $\widetilde{u}_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t), \boldsymbol{z}_{-j}(t)) = u_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t), \boldsymbol{z}_{-j}(t)) \prod_{j=1}^{J} \pi_{j,z_j}(\boldsymbol{v}_j(t))$ as the expected value of the instantaneous utility function $u_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t), \boldsymbol{z}_{-j}(t))$ in (5.17) for UAV j at stage t. Therefore, the reward that UAV j obtains from action $\boldsymbol{z}_j$ at a given network state $\boldsymbol{v}_j(t)$:

$$r_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t), \boldsymbol{z}_{-j}(t)) = \begin{cases} \widetilde{u}_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t), \boldsymbol{z}_{\text{-}j}(t)), \text{ if UAV } j \text{ reaches } d_j, \\ \widetilde{u}_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t), \boldsymbol{z}_{\text{-}j}(t)) + \gamma \max_{\boldsymbol{z}_j \in \mathcal{Z}_j} \boldsymbol{W}_{j,\text{out}}(\boldsymbol{z}_j(t+1), t+1) \\ \quad [\boldsymbol{v}'_j(t), \boldsymbol{x}_j'^{(1)}(t), \boldsymbol{x}_j'^{(2)}(t), \cdots, \boldsymbol{x}_j'^{(n)}(t)], \text{ otherwise.} \end{cases} \tag{5.25}$$

Here, $\boldsymbol{v}'_j(t+1)$ and $\boldsymbol{x}_j'^{(n)}(t)$, correspond, respectively, to the next network state and reservoir state of layer (n), at stage $(t+1)$, upon taking actions $\boldsymbol{z}_j(t)$ and

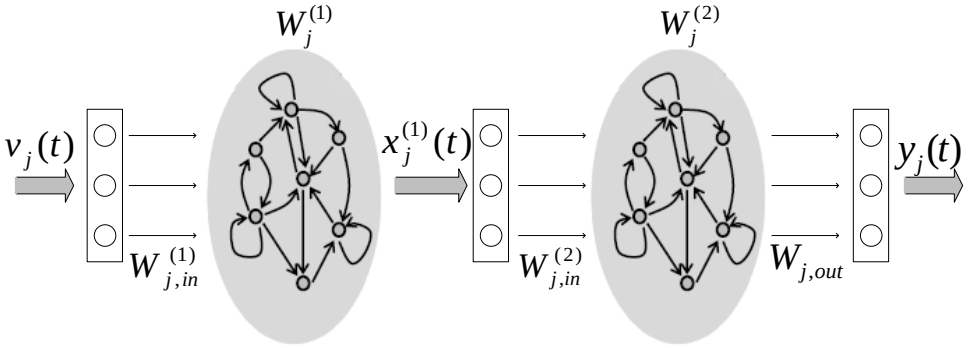

Figure 5.1. Proposed Deep ESN architecture.

$\boldsymbol{z}_{-j}(t)$ at stage t. Fig. 5.1 shows the proposed reservoir architecture of the deep ESN consisting of two layers.

5.4.2 Update Rule Based on Deep ESN

We now introduce the deep ESN's update phase that each UAV uses to store and estimate the reward function of each path and resource allocation scheme at a given stage t. In particular, we consider leaky integrator reservoir units [193] for updating the state transition functions $\boldsymbol{x}_j^{(n)}(t)$ at stage t. Therefore, the state transition function of the first layer $\boldsymbol{x}_j^{(1)}(t)$ will be:

$$\boldsymbol{x}_j^{(1)}(t) = (1-\omega_j^{(1)})\boldsymbol{x}_j^{(1)}(t-1) + \omega_j^{(1)}\tanh(\boldsymbol{W}_{j,\text{in}}^{(1)}\boldsymbol{v}_j(t) + \boldsymbol{W}_j^{(1)}\boldsymbol{x}_j^{(1)}(t-1)), \quad (5.26)$$

where $\omega_j^{(n)} \in [0,1]$ is the leaking parameter at layer n for UAV j which relates to the speed of the reservoir dynamics in response to the input, with larger values of $\omega_j^{(n)}$ resulting in a faster response of the corresponding n-th reservoir to the input. The state transition of UAV j, $\boldsymbol{x}_j^{(n)}(t)$, for $n > 1$ is given by:

$$\boldsymbol{x}_j^{(n)}(t) = (1-\omega_j^{(n)})\boldsymbol{x}_j^{(n)}(t-1) + \omega_j^{(n)}\tanh(\boldsymbol{W}_{j,\text{in}}^{(n)}\boldsymbol{x}_j^{(n-1)}(t) + \boldsymbol{W}_j^{(n)}\boldsymbol{x}_j^{(n)}(t-1)), \quad (5.27)$$

The output $y_j(t)$ of the deep ESN at stage t is used to estimate the reward of each UAV j based on the current adopted action $\boldsymbol{z}_j(t)$ and $\boldsymbol{z}_{-j}(t)$ of UAV j and other UAVs $(-j)$, respectively, for the current network state $\boldsymbol{v}_j(t)$ after training $\boldsymbol{W}_{j,\text{out}}$. It can be computed as:

$$y_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t)) = \boldsymbol{W}_{j,\text{out}}(\boldsymbol{z}_j(t), t)[\boldsymbol{v}_j(t), \boldsymbol{x}_j^{(1)}(t), \boldsymbol{x}_j^{(2)}(t), \cdots, \boldsymbol{x}_j^{(n)}(t)]. \quad (5.28)$$

We adopt a temporal difference RL approach for training the output matrix $W_{j,\text{out}}$ of the deep ESN architecture. In particular, we employ a linear gradient

descent approach using the reward error signal, given by the following update rule [194]:

$$\boldsymbol{W}_{j,\text{out}}(\boldsymbol{z}_j(t),t+1)=\boldsymbol{W}_{j,\text{out}}(\boldsymbol{z}_j(t),t)+\lambda_j(r_j(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))-y_j(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t)))[\boldsymbol{v}_j(t), \boldsymbol{x}_j^{(1)}(t),\boldsymbol{x}_j^{(2)}(t),\cdots,\boldsymbol{x}_j^{(n)}(t)]^T. \quad (5.29)$$

Here, note that the objective of each UAV is to minimize the value of the error function $e_j(\boldsymbol{v}_j(t)) = |r_j(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) - y_j(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t))|$.

5.4.3 Proposed Deep RL Algorithm

Based on the proposed deep ESN architecture and update rule, we next introduce a multi-agent deep RL framework that the UAVs can use to learn an SPNE in behavioral strategies for the game $\mathcal{G}$. The algorithm is divided into two phases: *training and testing*. In the former, UAVs are trained offline before they become active in the network using the architecture of Subsection 5.4.1. The testing phase corresponds to the actual execution of the algorithm after which the weights of $\boldsymbol{W}_{j,\text{out}}, \forall j \in \mathcal{J}$ have been optimized and is implemented on each UAV for execution during run time.

During the training phase, each UAV aims at optimizing its output weight matrix $\boldsymbol{W}_{j,\text{out}}$ such that the value of the error function $e_j(\boldsymbol{v}_j(t))$ at each stage t is minimized. In particular, the training phase is composed of multiple iterations, each consisting of multiple rounds, i.e., the number of steps required for all UAVs to reach their corresponding destinations d_j. At each round, UAVs face a tradeoff between playing the action associated with the highest expected utility, and trying out all their actions to improve their estimates of the reward function in (5.25). This in fact corresponds to the exploration and exploitation tradeoff, in which UAVs need to strike a balance between exploring their environment and exploiting the knowledge accumulated through such exploration [195]. Therefore, we adopt the ϵ-greedy policy in which UAVs choose the action that yields the maximum utility value with a probability of $1 - \epsilon + \frac{\epsilon}{|\mathcal{Z}_j|}$ while exploring randomly other actions with a probability of $\frac{\epsilon}{|\mathcal{A}_j|}$. The strategy over the action space will be:

$$\pi_{j,z_j}(\boldsymbol{v}_j(t)) = \begin{cases} 1 - \epsilon + \frac{\epsilon}{|\mathcal{Z}_j|}, \ \text{argmax}_{z_j \in \mathcal{Z}_j} y_j\left(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t)\right), \\ \frac{\epsilon}{|\mathcal{Z}_j|}, \ \text{otherwise.} \end{cases} \quad (5.30)$$

Algorithm 3: Training phase of the proposed deep RL algorithm

Initialization:

$\boldsymbol{\pi}_{j,z_j}(\boldsymbol{v}_j(t)) = \frac{1}{|\mathcal{A}_j|} \forall t \in T, z_j \in \mathcal{Z}_j$, $y_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t)) = 0$, $\boldsymbol{W}_{j,\text{in}}^{(n)}$, $\boldsymbol{W}_j^{(n)}$, $\boldsymbol{W}_{j,\text{out}}$.

for The number of training iterations **do**

  while At least one UAV j has not reached its destination d_j, **do**

    for all UAVs j (in a parallel fashion) **do**

      Input: Each UAV j receives an input $\boldsymbol{v}_j(t)$ based on (5.16).

      Step 1: Action selection

      Each UAV j selects a random action $\boldsymbol{z}_j(t)$ with probability ϵ,

      Otherwise, UAV j selects $\boldsymbol{z}_j(t) = \text{argmax}_{z_j \in \mathcal{Z}_j} y_j\left(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t)\right)$.

      Step 2: Location, cell association and transmit power update

      Each UAV j updates its location, cell association and transmission power level based on the selected action $\boldsymbol{z}_j(t)$.

      Step 3: Reward computation

      Each UAV j computes its reward values based on (5.25).

      Step 4: Action broadcast

      Each UAV j broadcasts its selected action $\boldsymbol{z}_j(t)$ to all other UAVs.

      Step 5: Deep ESN update

      - Each UAV j updates the state transition vector $\boldsymbol{x}_j^{(n)}(t)$ for each layer (n) of the deep ESN architecture based on (5.26) and (5.27).

      - Each UAV j computes its output $y_j\left(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t)\right)$ based on (5.28).

      - The weights of the output matrix $\boldsymbol{W}_{j,\text{out}}$ of each UAV j are updated based on the linear gradient descent update rule given in (5.29).

    end for

  end while

end for

Based on the selected action $\boldsymbol{z}_j(t)$, each UAV j updates its location, cell association, and transmission power level and computes its reward function according to (5.25). To determine the next network state, each UAV j broadcasts its selected action to all other UAVs in the network. Then, each UAV j updates its state transition vector $\boldsymbol{x}_j^{(n)}(t)$ for each layer (n) of the deep ESN architecture according to (5.26) and (5.27). The output y_j at stage t is then updated based on (5.28). Finally, the weights of the output matrix $\boldsymbol{W}_{j,\text{out}}$ of each UAV j are updated based on the linear gradient descent update rule given in (5.29). Note that, a UAV stops taking any actions once it has reached its destination. A summary of the training phase is given in Algorithm 3.

Meanwhile, the testing phase corresponds to the actual execution of the algorithm. In this phase, each UAV chooses its action greedily for each state $\boldsymbol{v}_j(t)$, i.e., $\max_{z_j \in \mathcal{Z}_j} y_j(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t))$, and updates its location, cell association, and transmission power level accordingly. Each UAV then broadcasts its selected action and updates its state transition vector $\boldsymbol{x}_j^{(n)}(t)$ for each layer n of the deep ESN architecture based on (5.26) and (5.27). A summary of the testing phase is given in

Algorithm 4: Testing phase of the proposed deep RL algorithm

while At least one UAV j has not reached its destination d_j, **do**
  for all UAVs j (in a parallel fashion) **do**
    Input: Each UAV j receives an input $\boldsymbol{v}_j(t)$ based on (5.16).
    Step 1: Action selection
    Each UAV j selects an action $\boldsymbol{z}_j(t) = \text{argmax}_{z_j \in \mathcal{Z}_j} y_j\left(\boldsymbol{v}_j(t), \boldsymbol{z}_j(t)\right)$.
    Step 2: Location, cell association and transmit power update
    Each UAV j updates its location, cell association and transmission power level based on the selected action $\boldsymbol{z}_j(t)$.
    Step 3: Action broadcast
    Each UAV j broadcasts its selected action $\boldsymbol{z}_j(t)$ to all other UAVs.
    Step 4: State transition vector update
    Each UAV j updates the state transition vector $\boldsymbol{x}_j^{(n)}(t)$ for each layer (n) of the deep ESN architecture based on (5.26) and (5.27).
  end for
end while

Algorithm 4.

It is important to note that analytically guaranteeing the convergence of the proposed deep learning algorithm is challenging as it is highly dependent on the hyperparameters used during the training phase. For instance, using too few neurons in the hidden layers results in underfitting which could make it hard for the neural network to detect the signals in a complicated data set. On the other hand, using too many neurons in the hidden layers can either result in overfitting or an increase in the training time that could prevent the training of the neural network. Overfitting corresponds to the case when the model learns the random fluctuations and noise in the training data set to the extent that it negatively impacts the model's ability to generalize when fed with new data. Therefore, in this work, we limit our analysis of convergence by providing simulation results (see Section 5.5) to show that, under a reasonable choice of the hyperparameters, convergence is observed for our proposed game. In such cases, it is important to study the convergence point and the convergence complexity of our proposed algorithm. Next, we characterize the convergence point of our proposed algorithm.

Proposition 2. If Algorithm 3 converges, then the convergence strategy profile corresponds to a SPNE of game $\mathcal{G}$.

Proof. An SPNE is a strategy profile that induces a NE on every subgame. Therefore, at the equilibrium state of each subgame, there is no incentive for any UAV to deviate after observing any history of joint actions. Moreover, given the fact that an ESN framework exhibits adaptive memory that enables it to store necessary

previous state information, UAVs can essentially retain other players' actions at each stage t and thus take actions accordingly. To show that our proposed scheme guarantees convergence to an SPNE, we use the following lemma from [192].

Lemma 3. *For our proposed game $\mathcal{G}$, the payoff functions in (5.25) are bounded, and the number of players, state space and action space is finite. Therefore, $\mathcal{G}$ is a finite game and hence a SPNE exists. This follows from Selten's theorem which states that every finite extensive form game with perfect recall possesses an SPNE where the players use behavioral strategies.*

Here, it is important to note that for finite dynamic games of perfect information, any backward induction solution is a SPNE [127]. Therefore, given the fact that, for our proposed game $\mathcal{G}$, each UAV aims at maximizing its expected sum of *discounted rewards* at each stage t as given in (5.25), one can guarantee that the convergence strategy profile corresponds to a SPNE of game $\mathcal{G}$. This completes the proof. □

Moreover, it is important to note that the convergence complexity of the proposed deep RL algorithm for reaching a SPNE is $O(J \times A^2)$. Next, we analyze the computational complexity of the proposed deep RL algorithm for practical scenarios in which the number of UAVs is relatively small.

Theorem 3. For practical network scenarios, the computational complexity of the proposed training deep RL algorithm is $O(A^3)$ and reduces to $O(A^2)$ when considering a fixed altitude for the UAVs, where A is the number of discretized unit areas.

Proof. Consider the case in which the UAVs can move with a fixed step size in a 3D space. For such scenarios, the state vector $\boldsymbol{v}'_j(t)$ is defined as:

$$\boldsymbol{v}'_j(t)=\left[\{\delta_{j,l,a}(t),\theta_{j,l,a}(t)\}_{l=1}^{L_j},\theta_{j,d_j,a}(t),\{x_j(t),y_j(t),h_j(t)\}_{j\in\mathcal{J}}\right], \tag{5.31}$$

For each state $\boldsymbol{v}'_j(t)$, the action of UAV j is a function of the location, transmission power level and cell association vector of all other UAVs in the network. Nevertheless, the number of possible locations of other UAVs in the network is much larger than the possible number of transmission power levels and the size of the cell association vector of those UAVs. Therefore, by the law of large numbers, one can consider the number of possible locations of other UAVs only when analyzing the convergence complexity of the proposed training algorithm. Moreover,

Table 5.2 System parameters.

Parameters	Values	Parameters	Values
UAV max transmit power ($\overline{P}_j$)	20 dBm	SINR threshold ($\overline{\Gamma}_j$)	-3 dB
UE transmit power ($\widehat{P}_q$)	20 dBm	Learning rate (λ_j)	0.01
Noise power spectral density (N_0)	-174 dBm/Hz	RB bandwidth (B_c)	180 kHz
Total bandwidth (B)	20 MHz	# of interferers (L)	2
Packet arrival rate ($\lambda_{j,s}$)	(0,1)	Packet size (ν)	2000 bits
Carrier frequency ($\hat{f}$)	2 GHz	Discount factor (γ)	0.7
# of hidden layers	2	Step size ($\widetilde{a}_j$)	40 m
Leaky parameter/layer ($\omega_j^{(n)}$)	0.99, 0.99	ϵ	0.3

for practical scenarios, the total number of UAVs in a given area is considered to be relatively small as compared to the number of discretized unit areas i.e., $J \ll A$ 3GPP admission control policy for cellular-connected UAVs [178]). Therefore, by the law of large numbers and given the fact that the UAVs take actions in a parallel fashion, the computational complexity of our proposed algorithm is $O(A^3)$ when the UAVs update their x, y and z coordinates and reduces to $O(A^2)$ when considering fixed altitudes for the UAVs. This completes the proof. □

From Theorem 3, we can conclude that the convergence speed of the proposed training algorithm is significantly reduced when considering a fixed altitude for the UAVs. This in essence is due to the reduction of the state space dimension when updating the x and y coordinates only. It is important to note here that there exists a tradeoff between the computational complexity of the proposed training algorithm and the resulting network performance. In essence, updating the 3D coordinates of the UAVs at each step t allows the UAVs to better explore the space thus providing more opportunities for maximizing their corresponding utility functions. Therefore, from both Theorems 3 and 2, the UAVs can update their x and y coordinates only during the learning phase while operating within the upper and lower altitude bounds derived in Theorem 2.

5.5 Simulation Results and Analysis

For our simulations, we consider an 800 m × 800 m square area divided into 40 m × 40 m grid areas, in which we randomly uniformly deploy 15 BSs. All statistical results are averaged over several independent testing iterations during which the initial locations and destinations of the UAVs and the locations of the BSs and the ground UEs are randomized. The maximum transmit power for each UAV is discretized into 5 equally separated levels. We consider an uncorrelated

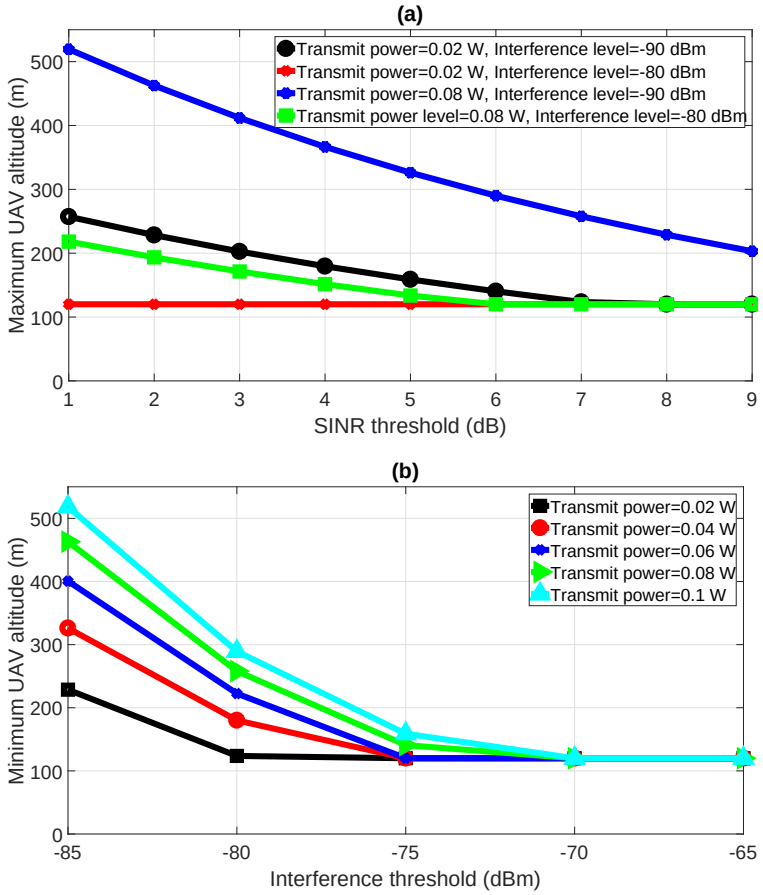

Figure 5.2. The (a) upper bound for the optimal altitude of the UAVs as a function of the SINR threshold value ($\bar{\Gamma}$) and for different transmit power levels and ground network density and (b) lower bound for the optimal altitude of the UAVs as a function of the interference threshold value ($\sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}$) and for different transmit power levels.

Rician fading channel with parameter $\widehat{K} = 1.59$ [196]. The external input of the deep ESN architecture, $\boldsymbol{v}_j(t)$, is a function of the number of UAVs and thus the number of hidden nodes per layer, $N_{j,R}^{(n)}$, varies with the number of UAVs. For instance, $N_{j,R}^{(n)} = 12$ and 6 for $n = 1$ and 2, respectively, for a network size of 1 and 2 UAVs, and 20 and 10 for a network size of 3, 4, and 5 UAVs. Table 5.2 summarizes the main simulation parameters.

Fig. 5.2 (a) shows the upper bound for the optimal altitude of UAV j as a function of the SINR threshold value, $\bar{\Gamma}$, and for different transmit power levels and ground network density, based on Theorem 2. On the other hand, Fig. 5.2 (b) shows the lower bound for the optimal altitude of UAV j as a function of the interference threshold value, $(\sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a})$, and for different transmit power levels, based on Theorem 2. From Fig. 5.2, we can deduce that the optimal altitude range of a given UAV is a function of network design parameters, ground

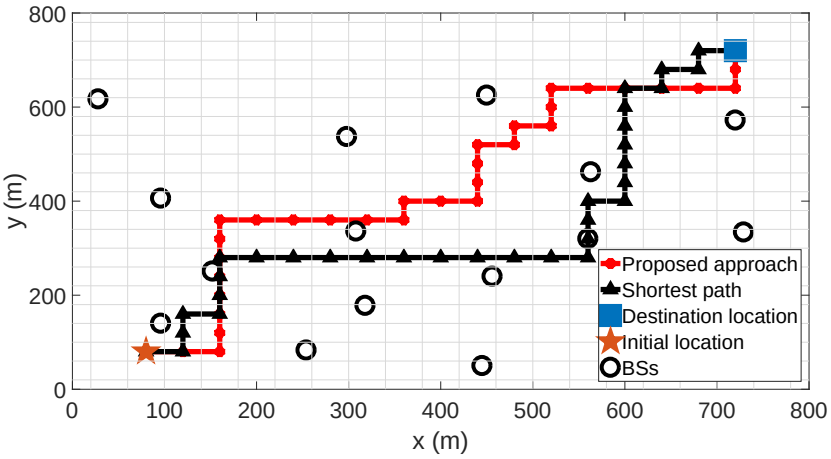

Figure 5.3. Path of a UAV for our approach and shortest path scheme.

Table 5.3 Performance assessment for one UAV.

	# of steps	delay (ms)	average rate per UE (Mbps)
Proposed approach	32	6.5	0.95
Shortest path	32	12.2	0.76

network data requirements, the density of the ground network, and its action $\boldsymbol{v}_j(t)$. For instance, the upper bound on the UAV's optimal altitude decreases as $\bar{\Gamma}$ increases while its lower bound decreases as $\sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}$ increases. Moreover, the maximum altitude of the UAV decreases as the ground network gets denser while its lower bound increases as the ground network data requirements increase. Thus, in such scenarios, a UAV should operate at higher altitudes. A UAV should also operate at higher altitudes when its transmit power level increases due to the increase in the lower and upper bounds of its optimal altitude.

Fig. 5.3 shows a snapshot of the path of a single UAV resulting from our approach and from a shortest path scheme. Unlike our proposed scheme which accounts for other wireless metrics during path planning, the objective of the UAVs in the shortest path scheme is to reach their destinations with the minimum number of steps. Table 5.3 presents the performance results for the paths shown in Fig. 5.3. From Fig. 5.3, we can see that, for our proposed approach, the UAV selects a path away from the densely deployed area while maintaining proximity to its serving BS in a way that would minimize the steps required to reach its destination. This path will minimize the interference level that the UAV causes on the ground UEs and its wireless latency (Table 5.3). From Table 5.3, we can see that our proposed approach achieves 25% increase in the average rate per ground UE and 47% decrease in the wireless latency as compared to the shortest path, while requiring the same number of steps that the UAV needs to reach the

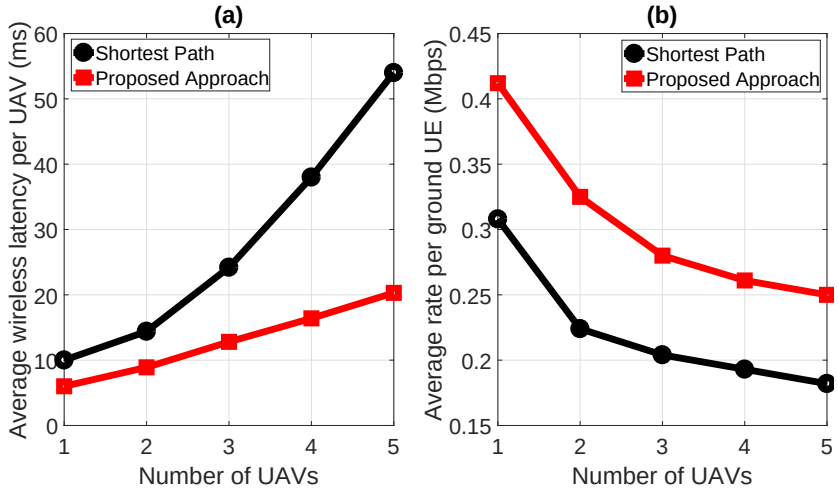

Figure 5.4. Performance assessment of the proposed approach in terms of average (a) wireless latency per UAV and (b) rate per ground UE as compared to the shortest path approach, for different number of UAVs.

Table 5.4 The required number of steps for all UAVs to reach their corresponding destinations based on our proposed approach and that of the shortest path scheme for different number of UAVs.

# of steps	1 UAV	2 UAVs	3 UAVs	4 UAVs	5 UAVs
Proposed approach	4	4	6	7	8
Shortest path	4	4	6	6	7

destination.

Fig. 5.4 compares the average values of the (a) wireless latency per UAV and (b) rate per ground UE resulting from our proposed approach and the baseline shortest path scheme. Moreover, Table 5.4 compares the number of steps required by all UAVs to reach their corresponding destinations for the scenarios presented in Fig. 5.4. From Fig. 5.4 and Table 5.4, we can see that, compared to the shortest path scheme, our approach achieves a lower wireless latency per UAV and a higher rate per ground UE for different numbers of UAVs while requiring a number of steps that is comparable to the baseline. In fact, our scheme provides a better tradeoff between energy efficiency, wireless latency, and ground UE data rate compared to the shortest path scheme. For instance, for 5 UAVs, our scheme achieves a 37% increase in the average achievable rate per ground UE, 62% decrease in the average wireless latency per UAV, and 14% increase in energy efficiency. Indeed, one can adjust the multi-objective weights of our utility function based on several parameters such as the rate requirements of the ground network, the power limitation of the UAVs, and the maximum tolerable wireless

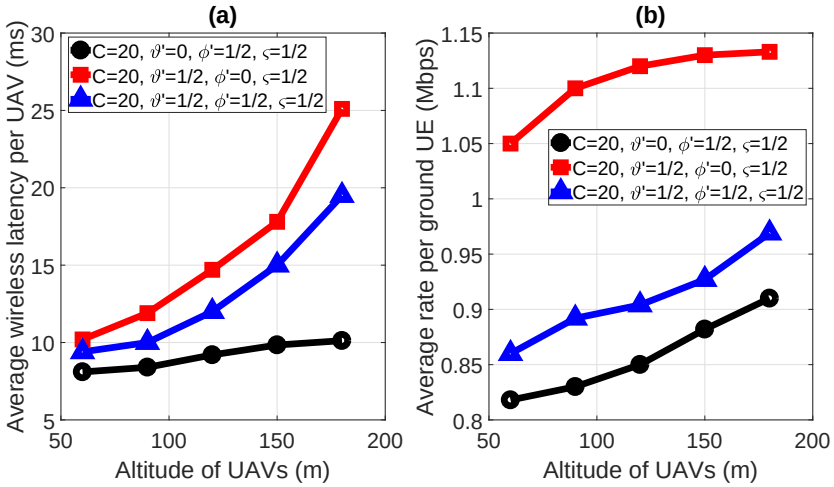

Figure 5.5. Performance assessment of the proposed approach in terms of average (a) wireless latency per UAV and (b) rate per ground UE for different utility functions and for different altitudes of the UAVs.

latency of the UAVs. Moreover, Fig. 5.4 shows that, as the number of UAVs increases, the average delay per UAV increases and the average rate per ground UE decreases, for all schemes. This is due to the increase in the interference level on the ground UEs and other UAVs as a result of the LoS link between the UAVs and the BSs.

Fig. 5.5 studies the effect of the UAVs' altitude on the average values of the (a) wireless latency per UAV and (b) rate per ground UE for different utility functions. From Fig. 5.5, we can see that, as the altitude of the UAVs increases, the average wireless latency per UAV increases for all studied utility functions. This is mainly due to the increase in the distance of the UAVs from their corresponding serving BSs which accentuates the path loss effect. Moreover, higher UAV altitudes result in a higher average data rate per ground UE for all studied utility functions mainly due to the decrease in the interference level that is caused from the UAVs on neighbouring BSs. Here, there exists a tradeoff between minimizing the average wireless delay per UAV and maximizing the average data rate per ground UE. Therefore, alongside the multiobjective weights, the altitude of the UAVs can be varied such that the ground UE rate requirements is met while minimizing the wireless latency for each UAV based on its mission objective.

Fig. 5.6 shows the average transmit power level per UAV along its path as a function of the number of BSs considering two utility functions, one for minimizing the average wireless latency for each UAV and the other for minimizing the interference level on the ground UEs. From Fig. 5.6, we can see that network densification has an impact on the transmission power level of the UAVs. For

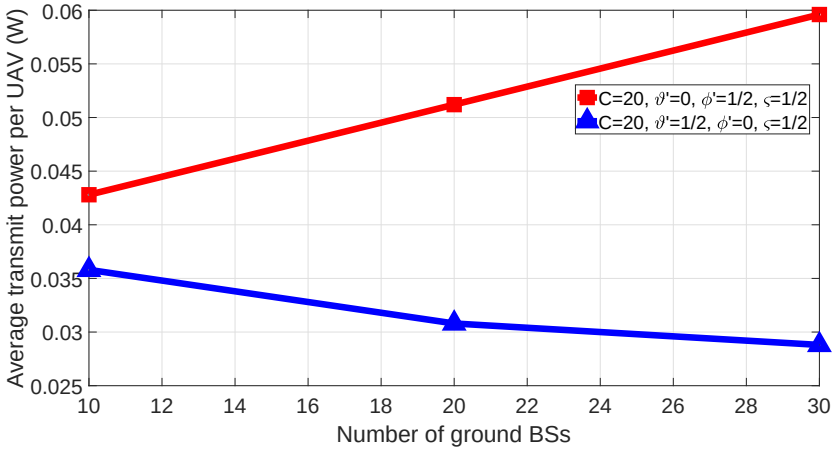

Figure 5.6. Effect of the ground network densification on the average transmit power level of the UAVs along their paths.

instance, when minimizing the wireless latency of each UAV along its path, the average transmit power level per UAV increases from 0.04 W to 0.06 W as the number of ground BSs increases from 10 to 30, respectively. In essence, the increase in the transmit power level is the result of the increase in the interference level from the ground UEs as the ground network becomes denser. As a result, the UAVs will transmit using a larger transmission power level so as to minimize their wireless transmission delay. On the other hand, the average transmit power level per UAV decreases from 0.036 W to 0.029 W in the case of minimizing the interference level caused on neighbouring BSs. This is due to the fact that as the number of BSs increases, the interference level caused by each UAV on the ground network increases thus requiring each UAV to decrease its transmit power level. Note that, when minimizing the wireless latency, the average transmit power per UAV is always larger than the case of minimizing the interference level, irrespective of the number of ground BSs. Therefore, the transmit power level of the UAVs is a function of their mission objective and the number of ground BSs.

Fig. 5.7 presents the (a) wireless latency per UAV and (b) rate per ground UE for different utilities as a function of the number of BSs and for a fixed altitude of 120 m. From this figure, we can see that, as the ground network becomes more dense, the average wireless latency per UAV increases and the average rate per ground UE decreases for all considered cases. For instance, when the objective is to minimize the interference level along with energy efficiency, the average wireless latency per UAV increases from 13 ms to 47 ms and the average rate per ground UE decreases from 0.86 Mbps to 0.48 Mbps as the number of BSs increases from 10 to 30. This is due to the fact that a denser network results in

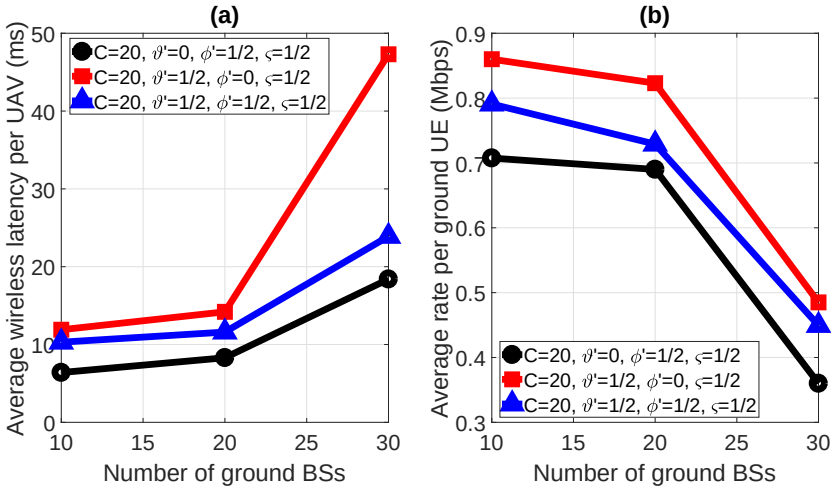

Figure 5.7. Effect of the ground network densification on the average (a) wireless latency per UAV and (b) rate per ground UE for different utility functions and for a fixed altitude of 120m.

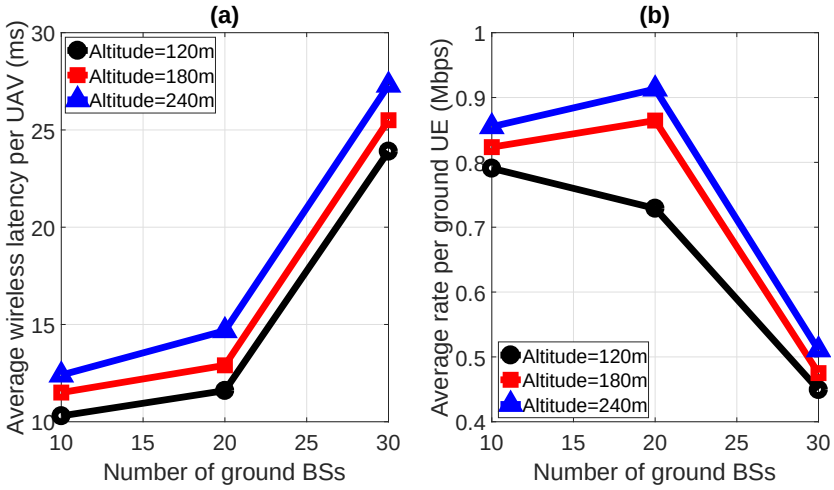

Figure 5.8. Effect of the ground network densification on the average (a) wireless latency per UAV and (b) rate per ground UE for different utility functions and for various altitudes of the UAVs.

higher interference on the UAVs as well as other UEs in the network.

Fig. 5.8 investigates the (a) wireless latency per UAV and (b) rate per ground UE for different values of the UAVs altitude and as a function of the number of BSs. From this figure, we can see that as the UAV altitude increases and/or the ground network becomes denser, the average wireless latency per UAV increases. For instance, the delay increases by 27% as the altitude of the UAVs increases from 120 to 240 m for a network consisting of 20 BSs and increases by 120% as the number of BSs increases from 10 to 30 for a fixed altitude of 180 m. This essentially follows from Theorem 2 and the results in Fig. 5.2 (a) which shows that the maximum altitude of the UAV decreases as the ground network gets denser and thus the UAVs should operate at a lower altitude when the number of

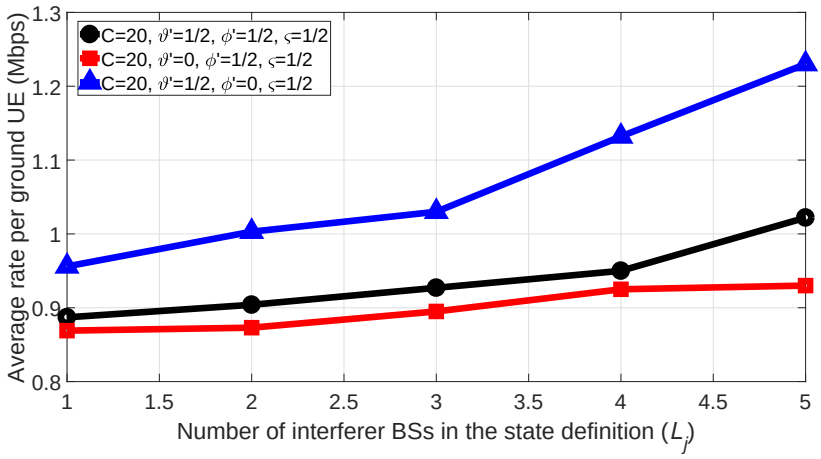

Figure 5.9. The average rate per ground UE as a function of the number of interferer BSs in the state definition (L_j).

BSs increases from 10 to 30. Moreover, the average rate per ground UE decreases as the ground network becomes denser due to the increase in the interference level and increases as the altitude of the UAVs increases. Therefore, the resulting network performance depends highly on both the UAVs altitude and the number of BSs in the network. For instance, in case of a dense ground network, the UAVs need to fly at a lower altitude for applications in which the wireless transmission latency is more critical and at a higher altitude in scenarios in which a minimum achievable data rate for the ground UEs is required.

Fig. 5.9 shows the effect of varying the number of nearest BSs (L_j) in the observed network state of UAV j, $\boldsymbol{v}_j(t)$, on the average data rate per ground UE for different utility functions. From Fig. 5.9, we can see an improvement in the average rate per ground UE as the number of nearest BSs in the state definition increases. For instance, in scenarios in which the UAVs aim at minimizing the interference level they cause on the ground network along their paths, the average rate per ground UE increases by 28% as the number of BSs in the state definition increases from 1 to 5. This gain results from the fact that as L_j increases, the UAVs get a better sense of their surrounding environment and thus can better select their next location such that the interference level they cause on the ground network is minimized. It is important to note here, that as L_j increases, the size of the external input ($\boldsymbol{v}_j$) increases thus requiring a larger number of neurons in each layer. This in turn increases the number of required iterations for convergence. Therefore, a tradeoff exists between improving the performance of the ground UEs and the running complexity of the proposed algorithm.

Fig. 5.10 shows the average of the error function $e_j(\boldsymbol{v}_j(t))$ resulting from the

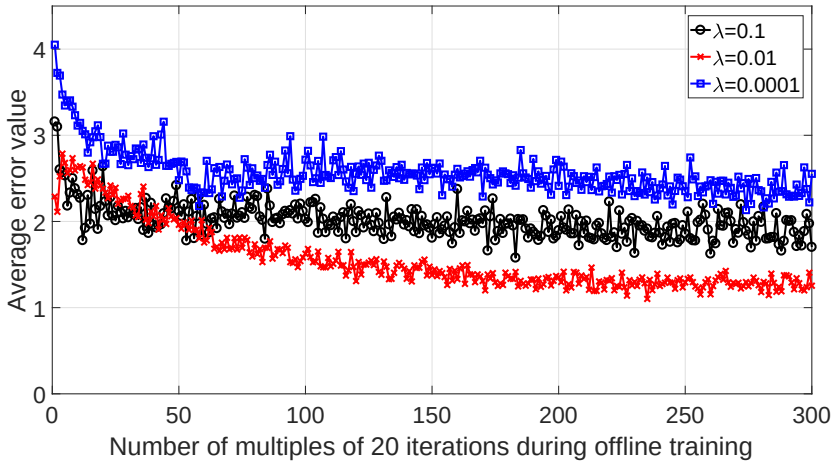

Figure 5.10. Effect of the learning rate on the convergence of offline training.

offline training phase as a function of a multiple of 20 iterations while considering different values for the learning rate, λ. The learning rate determines the step size the algorithm takes to reach the optimal solution and, thus, it impacts the convergence rate of our proposed framework. From Fig. 5.10, we can see that small values of the learning rate, i.e., $\lambda = 0.0001$, result in a slow speed of convergence. On the other hand, for large values of the learning rate, such as $\lambda = 0.1$, the error function decays fast for the first few iterations but then remains constant. Here, $\lambda = 0.1$ does not lead to convergence during the testing phase, but $\lambda = 0.0001$ and $\lambda = 0.01$ result in convergence, though requiring a different number of training iterations. In fact, a large learning rate can cause the algorithm to diverge from the optimal solution. This is because large initial learning rates will decay the loss function faster and thus make the model get stuck at a particular region of the optimization space instead of better exploring it. Clearly, our framework achieves better performance for $\lambda = 0.01$, as compared to smaller and larger values of the learning rate. We also note that the error function does not reach the value of zero during the training phase. This is due to the fact that, for our approach, we adopt the early stopping technique to avoid overfitting which occurs when the training error decreases at the expense of an increase in the value of the test error [3].

5.6 Summary

In this chapter, we have proposed a novel interference-aware path planning scheme that allows cellular-connected UAVs to minimize the interference they cause on a ground network as well as their wireless transmission latency while transmit-

ting online mission-related data. We have formulated the problem as a dynamic noncooperative game in which the UAVs are the players. To solve the game, we have proposed a deep RL algorithm based on ESN cells which is guaranteed to reach an SPNE, if it converges. The proposed algorithm enables each UAV to decide on its next location, transmission power level, and cell association vector in an autonomous manner thus adapting to the changes in the network. Simulation results have shown that the proposed approach achieves better wireless latency per UAV and rate per ground UE while requiring a number of steps that is comparable to the shortest path scheme. The results have also shown that a UAV's altitude plays a vital role in minimizing the interference level on the ground UEs as well as the wireless transmission delay of the UAV. In particular, we have shown that the altitude of the UAV is a function of the ground network density, the UAV's objective and the actions of other UAVs in the network.

Chapter 6

Unmanned Aerial Vehicles for Multi-hop Wireless Backhauling

6.1 Introduction

As we have discussed earlier in Chapter 2, 5G cellular networks will heavily rely on ultra-dense networks in order to cope with the increasing traffic demand. Ultra-dense networks are often HetNets, i.e., multi-layered including legacy high power macro-cells and very dense small cells with lower power. SBSs are multi-RAT capable and represent an essential part of ultra-dense networks, which are considered an imperative 5G solution. Nevertheless, these ultra-dense and heavy traffic cells should be connected to the core network through the backhaul, often with extreme requirements in terms of capacity, latency, availability, energy, and cost efficiency. Therefore, to reap the benefits of SBS deployment, innovative backhaul solutions are needed, as SBSs may be deployed in adverse locations and rural areas in which backhaul access is either non-existent or strictly limited in capacity [38]. As such, in this chapter, we propose a novel backhaul scheme that relies on UAVs as an on-demand flying network. The proposed scheme enables the UAVs to form the necessary multi-hop backhaul network in a decentralized manner thus adapting the backhaul architecture to the dynamics of the network. Therefore, for scenarios in which high-speed ground backhaul links are either unavailable or limited in capacity, the UAVs would serve as a bridge among the SBSs and relay the traffic to a nearby gateway node (with core network access) or as an intermediate relay point between different backhaul transceivers. The work presented in this chapter has appeared in the Proceedings of the IEEE

Table 6.1 Variables and notations.

Notation	Description
J	Number of UAVs
S	Number of SBSs
S_j	Number of SBSs served by UAV j
n	Gateway node
$\xi_{j,s}$	Free-space path loss between UAV j and SBS s
η_{LoS}	Attenuation factor for the LoS links
η_{NLoS}	Attenuation factor for the NLoS links
$P_{j,s}^{\text{LoS}}$	Probability of LoS
$P_{j,s}^{\text{NLoS}}$	Probability of NLoS
$L_{j,s}$	Path loss between UAV j and SBS s
$L_{j,s}$	Path loss between UAV j and UAV i
$\Gamma_{o,d}$	SINR between an origin node o and a destination node d
$\Gamma_{j,i}$	SNR between UAV j and UAV i
$P_{o,d}$	Transmit power of the origin node o to the destination node d
$h_{o,d}$	Channel gain between o and d
$R_j^{DL}(p_j, G)$	Data rate in the DL along path p_j over graph G
$R_j^{UL}(p_j, G)$	Data rate in the UL along path p_j over graph G
$P_j^{DL}(p_j, G)$	Number of relayed packets by UAV j in the DL along path p_j over graph G
$P_j^{UL}(p_j, G)$	Number of relayed packets by UAV j in the UL along path p_j over graph G
$\tau_j^{DL}(p_j, G)$	Average delay in the DL along path p_j over graph G
$\tau_j^{UL}(p_j, G)$	Average delay in the UL along path p_j over graph G
$U_j(p_j, G)$	Utility of UAV j along path p_j over graph G
$d_{j,i}$	Distance between UAV j and UAV i
$d_{j,i}^{\max}$	Maximum distance between UAV j and UAV i such that the SNR threshold $\widehat{\Gamma}$ is guaranteed
$\overrightarrow{\boldsymbol{F}}_{j,i}^{A}$	Attractive virtual force vector from UAV j towards UAV i
$\overrightarrow{\boldsymbol{F}}_{j,i}^{R1}$	Repulsive virtual force vector from UAV j towards UAV i upon deletion of link ji
$\overrightarrow{\boldsymbol{F}}_{j,i}^{R2}$	Repulsive virtual force vector from UAV j towards UAV i for physical collision avoidance
$\overrightarrow{\boldsymbol{F}}_{j}$	Total virtual force exerted from UAV j on UAV i

Global Conference on Communications [46]. Table 6.1 provides a summary for the description of the main notations used in this chapter. Next, we give an overview on the related literature and then we present our problem statement and contribution.

6.1.1 Related Work

Several approaches have been recently proposed for SBS backhauling [38–41]. Such solutions include wired and wireless backhauling to and from core network

aggregators, cooperation through anchor base stations, and multi-hop over short-range links [38, 39]. On the one hand, wired backhaul connection can be costly and does not always exist (e.g., remote/rural areas). On the other hand, wireless backhaul links are capacity limited. Therefore, existing solutions do not account for scenarios in which the high-speed ground backhaul is either congested, unavailable, or limited in capacity. In such scenarios, the backhaul connectivity of SBSs can become a bottleneck thus degrading the performance of the radio access network. Therefore, a novel paradigm shift of backhaul network design for 5G networks and beyond is needed. One promising solution for such scenarios is to deploy UAVs for providing backhaul connectivity to the ground SBSs [181] and [16]. Due to their rapid and flexible deployment capabilities, mobility, ability to fly above obstacles, and relatively low cost, UAVs have received considerable interest for different applications in wireless communications, and in particular, as communication relays [197–199].

In this regard, the authors in [197] consider the formation of a multi-hop relay system based on UAVs in order to extend the communication range of the ground network. In [198] and [199], the authors consider a mobile relay network model in which a UAV serves as a resilient moving relay among the SBSs. In [200], the authors consider the use of UAVs as relays for backhaul connectivity of high altitude balloons in case of temporary failed links. Although the use of UAVs as communication relays has been explored in the literature [197–199], these works are restricted to ad hoc, rather than cellular networks. On the other hand, the authors in [41–43] propose a vertical fronthaul/backhaul framework based on UAVs and free-space optics communication. In [42, 43], the authors propose an optimization formulation for the association problem of the UAVs and the SBSs but ignoring the design of the multi-hop links among the UAVs. Therefore, one challenging area which remains relatively unexplored is the dynamic formation of the aerial graph that connects UAVs to the core network. Indeed, the existing prior art does not provide an efficient scheme, in terms of achievable rate and delay, for the formation of a multi-hop aerial network for SBS backhauling.

6.1.2 Problem Statement and Contribution

The main objective of this chapter is to propose an online framework that allows the UAVs to form a multi-hop aerial network for SBS backhauling. In essence,

providing backhaul connection to SBSs could be challenging in scenarios where the SBSs may be deployed in adverse locations and rural areas in which backhaul access is either non-existent or strictly limited in capacity. Therefore, the design of a UAV-based multi-hop backhaul network is essential for such scenarios. In essence, a multi-hop aerial network enables the extension of the communication range of the UAVs thus connecting more SBSs to the core network. However, a challenging aspect for the UAV-based multi-hop backhaul network is the formation of the A2A links among the UAVs as well as the A2G links between the UAVs and the SBSs and is the main scope of this chapter. Moreover, the UAVs should be capable of adapting their corresponding link formation based on the dynamics of the network such as capacity, data rate requirements, and delay.

The main contribution of this chapter is to introduce a novel backhaul framework that utilizes UAVs as an on-demand flying network linking the SBSs and the core network in scenarios where the ground backhaul is either unavailable or limited in capacity. The design of the aerial backhaul network is formulated as a network formation game in which the players are the UAVs. The objective of the proposed game is to allow the UAVs to autonomously learn which A2A and A2G links to form in order to guarantee the connectivity of the SBSs to the core network. In particular, we consider that the UAVs form a multi-hop aerial network in which each UAV can individually select the path that connects it to the backhaul gateway node through other UAVs while optimizing its own utility. To solve this game, we propose a dynamic network formation algorithm that is guaranteed to reach a pairwise stable network upon convergence. Moreover, to ensure an efficient backhauling process between the UAVs, we incorporate the notion of *virtual force fields* [45] into our dynamic algorithm. In essence, virtual forces allow the UAVs to adjust their location dynamically based on the links they want to form. We show that, using the proposed algorithm, the UAVs are able to self-organize into a stable tree structure rooted at the gateway node. Simulation results show that the proposed approach achieves significant rate and delay improvements.

The rest of this chapter is organized as follows. In Section 6.2, we present the system model. Section 6.3 describes the proposed network formation game. The proposed network formation algorithm is given in Section 6.4. In Section 6.5, simulation results are analyzed. Finally, a summary is provided in Section 6.6.

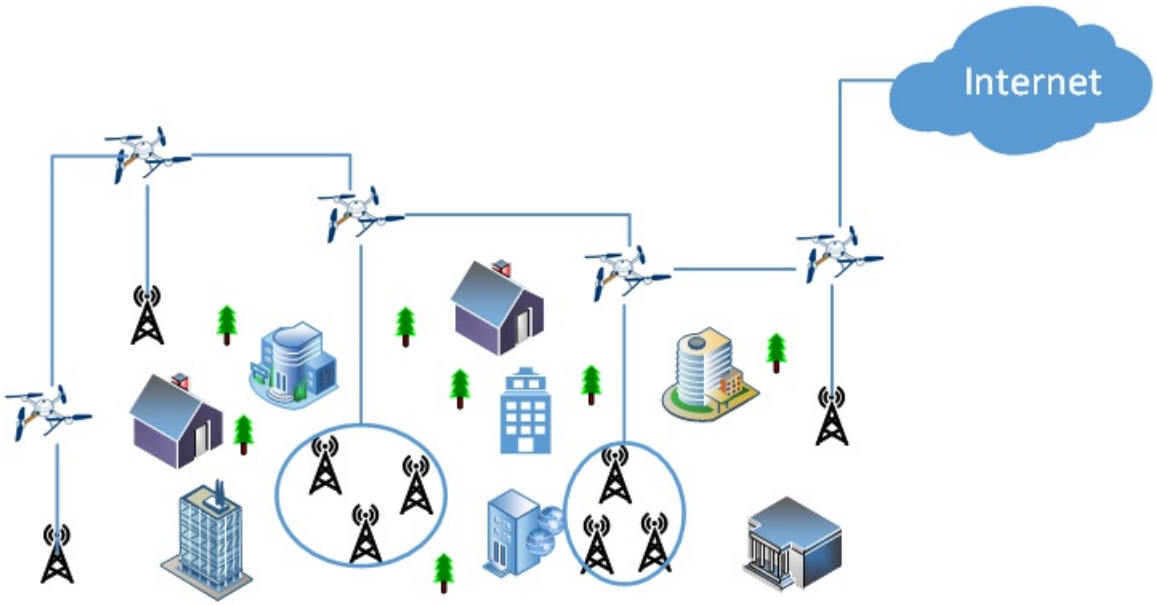

Figure 6.1. Illustration of the system model.

6.2 System Model

Consider a network composed of a set $\mathcal{S}$ of S SBSs and a set $\mathcal{J}$ of J UAVs. We consider networks deployed in rural areas, hotspots, or ultra dense cellular areas in which SBSs are located at adverse locations (e.g., at lamp posts or street levels), and a ground backhaul network that connects the SBSs to the core network is either unavailable or limited in capacity. To overcome such bottleneck, we propose the use of UAVs as a temporary aerial backhaul network for the SBSs. In particular, the UAVs serve as a bridge among the SBSs that relay their traffic to a nearby gateway node (with core network access) or as an intermediate relay between backhaul transceivers. An illustration of the proposed system model is shown in Fig. 6.1.

In our model, the UAVs are initially located based on the deployment approach given in [201] and each UAV j serves a set of $\mathcal{S}_j$ SBSs. Packet forwarding is supported for both UL and DL directions via an FDD model thus allowing the flow of traffic to (from) the SBSs from (to) the core network through a gateway node. We consider the availability of one gateway node n in a given area and assume that at least one UAV has access to this gateway node. Given that the communication range of low-altitude platform UAVs is typically limited to a few hundred meters, after which the signal quality deteriorates [202], the formation of a multi-hop aerial network becomes necessary to extend the communication range of the ground network and provide service to SBSs that are located at distant

or hard to reach areas where infrastructure does not exist. Consequently, a communication link with the infrastructure is formed through either UAV-to-UAV multi-hop links or a UAV-to-infrastructure data link.

6.2.1 A2G and A2A channel models

In our proposed model, UAVs transmit over the licensed cellular band for the A2G and A2A links. We adopt the free-space path loss model, ξ, given by [78]:

$$\xi(\text{dB}) = 20\log_{10}(d_{o,d}) + 20\log_{10}(f_c) - 147.55, \tag{6.1}$$

where f_c is the system center frequency (in Hz) and $d_{o,d} = \frac{\Delta h_{o,d}}{\sin\theta_{o,d}}$ is the Euclidean distance between an origin node o and a destination node d (in m); $\Delta h_{o,d} = z_o - z_d$ is the altitude difference between o and d and $\theta_{o,d}$ is the elevation angle. The use of a free space propagation model is validated by the fact that low-altitude platform UAVs fly at an altitude of $\sim$100m.

We consider a probabilistic LoS and NLoS links for the A2G propagation channel as done in [201]. In such a model, NLoS links experience higher attenuations due to the shadowing and diffraction loss. Therefore, the adopted path loss model between UAV j and SBS s, $L_{j,s}$, is given by:

$$L_{j,s} = \begin{cases} \xi_{j,s} + \eta_{\text{LoS}}, & \text{LoS link}, \\ \xi_{j,s} + \eta_{\text{NLoS}}, & \text{NLoS link}. \end{cases} \tag{6.2}$$

where η_{LoS} and η_{NLoS} correspond to additional attenuation factors added to the free space propoagation model for the LoS and NLoS links, respectively. Here, the probability of a LoS connection depends on the environment, density and height of buildings, the locations of the UAV and the SBS, and the corresponding elevation angle. The LoS probability is given by [78]:

$$P_{j,s}^{\text{LoS}} = \frac{1}{1 + C\exp(-D[\theta_{j,s} - C])}, \tag{6.3}$$

where C and D are constants which depend on the environment (rural, urban, dense urban, or others) and $\theta_{j,s} = \sin^{-1}(\frac{\Delta h_{j,s}}{d_{j,s}})$ is the elevation angle. Clearly, the probability of NLoS is $P_{j,s}^{\text{NLoS}} = 1 - P_{j,s}^{\text{LoS}}$. Therefore, the average path loss between UAV j and SBS s, $\overline{L}_{j,s}$, is given by:

$$\overline{L}_{j,s} = P_{j,s}^{\text{LoS}} \cdot L_{j,s}^{\text{LoS}} + P_{j,s}^{\text{NLoS}} \cdot L_{j,s}^{\text{NLoS}}, \tag{6.4}$$

For the A2A links, we consider LoS links between different UAVs that wish to form a link. Therefore, the path loss between UAV j and UAV i, will be $L_{j,i} = \xi_{j,i} + \eta_{\text{LoS}}$. Based on the given channel model, the average SINR of the A2G link between an origin node o and a destination node d (which can represent the link between UAV j and SBS s or UAV j and the gateway node n) in the DL or UL direction, $\Gamma_{o,d}$, is given by:

$$\Gamma_{o,d} = \frac{P_{o,d} \cdot h_{o,d}}{\sum_{q=1,q\neq o}^{O} I_{q,d} + \sigma^2}, \tag{6.5}$$

where $P_{o,d}$ is the transmit power of the origin node o (which can represent UAV j or SBS s) to the destination node d, $h_{o,d} = 1/10^{\overline{L}_{o,d}/10}$ is the channel gain between o and d, σ^2 is the Gaussian noise and $\sum_{q=1,q\neq o}^{O} I_{q,d}$ is the total interference power at the destination node d from other neighbouring origin nodes q (UAVs in the DL or SBSs in the UL) that are transmitting over the same channel, where $I_{q,d} = P_{q,d}/10^{\overline{L}_{q,d}/10}$. Therefore, the achievable data rate of the A2G link will be given by $R_{o,d} = B_o \log_2(1 + \Gamma_{o,d})$, where B_o is the transmission bandwidth of the origin node o.

For the A2A links, we consider orthogonal channel allocation among all UAVs and, hence, the signal-to-noise ratio (SNR) between UAVs j and i will be $\Gamma_{j,i} = \frac{P_{j,i}}{10^{L_{j,i}/10} \cdot \sigma^2}$. The capacity of the A2A link is $R_{j,i} = B_j \log_2(1 + \Gamma_{j,i})$, where B_j is the transmission bandwidth of UAV j.

Therefore, the achievable end-to-end rate, $R_j^{\text{DL}}(p_j)$, of UAV j along a multi-hop path p_j in the DL direction corresponds to the minimum of the rates achievable over N hops, as given below [203]:

$$R_j^{\text{DL}}(p_j) = \min_{n=1,\cdots,N} R_{j_k,j_{k+1}}^{\text{DL}}, \tag{6.6}$$

where $R_{j_k,j_{k+1}}^{\text{DL}}$ corresponds to the rate over link $j_k j_{k+1}$ in the DL direction. Similarly, for $R_j^{\text{UL}}(p_j)$, the achievable rate in the UL direction along path p_j.

6.2.2 Problem formulation

Given this model, our objective is to form an aerial backhaul network that allows each UAV j to be connected to the gateway node n via at most *one* path, denoted as p_j, whenever this path exists. To realize this, we consider the formation of a *bidirectional tree structure* rooted at the gateway node n. We let $\boldsymbol{\beta}$ be the UAV-UAV association vector with each element $\beta_{j,i} = 1$ if link ji is formed between

UAV j and UAV i, and 0, otherwise and $\boldsymbol{\alpha}$ be the UAV-gateway association vector with each element $\alpha_{j,n} = 1$ if link jn is formed between UAV j and the gateway node n, and 0, otherwise. Therefore, the centralized optimization formulation is given by:

$$\max_{\boldsymbol{\alpha},\boldsymbol{\beta}} \sum_{j=1}^{J} \phi_j(p_j(\alpha_{j,n}, \beta_{j,i})), \tag{6.7}$$

$$\text{s.t.} \quad \sum_{i=1, i\neq j}^{J} \beta_{j,i} + \alpha_{j,n} \geq 1 \ \ \forall j, \qquad \sum_{j=1}^{J} \alpha_{j,n} \geq 1, \tag{6.8}$$

$$\sum_{j=1}^{J} \Big(\alpha_{j,n} + \beta_{j,i} \Big) = J, \tag{6.9}$$

$$\alpha_{j,n} \in \{0,1\}, \ \ \beta_{j,i} \in \{0,1\} \quad \forall j, n. \tag{6.10}$$

where $\phi_j(p_j(\alpha_{j,n}, \beta_{j,i}))$ corresponds to the utility function of UAV j along its path p_j. (6.8) guarantees the formation of at least one path for each UAV j to the gateway node n (via direct or multi-hop). The left-most constraint in (6.8) ensures that UAV j is connected to at least another UAV i in the network or to the gateway node n. The right-most constraint in (6.8) guarantees that at least one UAV j is connected to the gateway node n. (6.9) limits the maximum number of formed edges in the network to J, the number of available UAVs. Thus, (6.8) and (6.9), avoid the formation of cycles in the network and, hence, guarantee the formation of a tree structure rooted at the gateway node. Finally, (6.10) represents the feasibility constraints.

Although a fully centralized approach can be used to form the aerial backhaul network, the need for a distributed solution is desirable for our problem as it has several advantages. For instance, a centralized control system suffers from the single-point failure problem. In contrast, a distributed approach does not rely on a single controller which, if compromised (due to malicious attacks or failures), can disrupt the operation of the entire network. Further, a centralized approach requires the controller to communicate with all UAVs at all times which is infeasible when the UAVs belong to different operators. Moreover, it can yield significant overhead and complexity, namely in networks with a rapidly changing environment due to the mobility of UAVs or incoming traffic load. Given these reasons, a distributed approach for network formation is needed, as proposed next.

6.3 Network Formation Game for UAV Backhauling

Our main objective is to develop a distributed approach that can model the interactions among UAVs that seek to form an aerial multi-hop backhaul network. For this purpose, we adopt the analytical framework of *network formation games* [128, 204] which involves a number of independent decisions makers that interact in order to form a suited network graph that connects them. For our proposed game, the players correspond to the set of UAVs and the action space of each UAV is defined as the set of links which UAV j can delete or form. Therefore, we consider an *undirected* graph $G(\mathcal{V}, \mathcal{E})$ with $\mathcal{V}$ being the set of all vertices (J UAVs and gateway node n) that will be present in the graph and $\mathcal{E}$ the set of all edges (links) that connect different pairs of nodes. Each undirected link $ji \in \mathcal{E}$ between two nodes j and i corresponds to the DL/UL traffic flow between these nodes. Given any network $G(\mathcal{V}, \mathcal{E})$, the path p_j from UAV j to the gateway node n is defined as a sequence of nodes $j_1, \cdots, j_K$ (in $\mathcal{V}$) such that $j_1 = j$, $j_K = n$ and each undirected link $j_k j_{k+1} \in \mathcal{E}$ for each $k \in \{1, \cdots, K-1\}$.

Therefore, each UAV j aims at optimizing its own utility by selecting an appropriate path that connects it to the backhaul gateway node through other UAVs. Subsequently, the UAVs can act as source nodes transmitting the received SBSs/gateway node packets to the gateway node/SBSs through one or more hops in the formed graph. The resulting network graph G is highly dependent on the goals, objectives, and incentives of each UAV. For instance, the number of hops can have an impact on the end-to-end delay, scalability, and throughput and therefore, can affect the performance of the resulting network. Next, we define the proposed utility function for our game.

6.3.1 Utility function

The utility of each UAV j is function of the network topology and the set of links formed among different UAVs. To this end, we propose a utility function that captures key metrics such as rate, delay, and number of relayed packets.

6.3.1.1 Achievable data rate

To maximize the performance of the SBSs, each UAV j aims at maximizing its end-to-end achievable data rate along its path p_j in the DL and UL directions, denoted as $R_j^{DL}(p_j, G)$ and $R_j^{UL}(p_j, G)$ respectively.

6.3.1.2 Number of relayed packets

To provide incentives for UAVs to route each others' packets, each UAV j is given a positive utility equivalent to the number of packets it transmits successfully to/from the gateway node via DL and UL, $P_j^{DL}(p_j, G)$ and $P_j^{UL}(p_j, G)$, respectively. These packets correspond to packets originating from the set $\mathcal{S}_j$ of SBSs connected to UAV j and from all other UAVs connected to UAV j.

6.3.1.3 Delay cost

We assume an M/D/1 queueing system and we define $\tau_j(p_j, G)$ as the average delay over path $p_j = \{j_1, \ldots, j_K\}$ from SBS s connected via UAV j to the core network (or vice versa). Therefore, $\tau_j(p_j, G)$ can be given by [204]:

$$\tau_j(p_j,G)= \sum_{(j_k,j_{k+1})\in q_j} \left(\frac{\Psi_{j_k,j_{k+1}}}{2\mu_{j_k,j_{k+1}}(\mu_{j_k,j_{k+1}}\text{-}\Psi_{j_k,j_{k+1}})} + \frac{1}{\mu_{j_k,j_{k+1}}} \right), \tag{6.11}$$

where $\Psi_{j_k,j_{k+1}} = \Lambda_{j_k} + \Delta_{j_k}$ is the total packet arrival rate (packets/sec) traversing link $(j_k, j_{k+1}) \in p_j$ between UAV j_k and UAV j_{k+1} and originating from the set $\mathcal{S}_j$ of SBSs connected to UAV j_k and from all other UAVs that are connected to UAV j_k (considering the Kleinrock approximation [205]). $\Lambda_j = \sum_{s\in\mathcal{S}_j} \lambda_s$ where λ_s corresponds to the average arrival rate of the traffic of SBS s and $\Delta_{j_k} = \sum_{i\in A_{j_k}} \Lambda_i$ where A_{j_k} is the set of UAVs that have a link formed with UAV j. $\mu_{j_k,j_{k+1}} = R_{j_k,j_{k+1}}/\upsilon$ is the service rate over link (j_k, j_{k+1}) where $R_{j_k,j_{k+1}}$ is the rate of the direct transmission between UAV j and UAV $j+1$ and υ is the packet length. According to (6.11), the delay will be infinite when $\mu_{j_k,j_{k+1}} < \Psi_{j_k,j_{k+1}}$.

6.3.1.4 Total utility

Hence, the utility function $U_j(p_j, G)$ of UAV j along path p_j for $\mathcal{S}_j \neq \emptyset$, is defined as:

$$U_j(p_j, G) = \left(R_j^{DL}(p_j, G) + R_j^{UL}(p_j, G) \right) + \delta_j \left(P_j^{DL}(p_j, G) + P_j^{UL}(p_j, G) \right) - \gamma_j \left(\tau_j^{DL}(p_j, G) + \tau_j^{UL}(p_j, G) \right), \tag{6.12}$$

where δ_j and γ_j are multi-objective weights.

Here, we note that there is no incentive for any UAV j to be disconnected from the gateway node, otherwise, its delay cost and, subsequently, its utility

function would be infinite. Therefore, for any network formation algorithm, the resulting tree graph of our proposed game is always connected.

6.3.2 Pairwise stability

Given the fact that, in network formation games, the consent of two nodes is required to form a single link, the stability of the outcome can be accurately characterized by considering bilateral deviations. To satisfy this requirement, we consider the notion of *pairwise stability* that was introduced in [128].

Definition 4. A network G is *pairwise stable* with respect to the proposed utility function $U_j(p_j, G)$ if:

1. for all $ji \in \mathcal{E}$, $U_j(p_j, G) \geq U_j(p_j - ji, G - ji)$ and $U_i(p_i, G) \geq U_i(p_i - ji, G - ji)$, and
2. for all $ji \notin \mathcal{E}$, if $U_j(p_j + ji, G + ji) > U_j(G)$ then $U_i(p_i + ji, G + ji) < U_i(G)$,

where $G - ji$ refers to deleting link ji from G and $G + ji$ refers to adding link ji to G.

Definition 5. When a network G is not pairwise stable, it is said to be *defeated* by G' if either $G' = G + ji$ and 2) is violated for ji, or if $G' = G - ji$ and 1) is violated for ji.

Therefore, a given backhaul graph is pairwise stable if there is no incentive for any UAV j to break a link that is formed with another UAV i (unilateral deviation) and no pair of UAVs j and i have an incentive to establish a new link (bilateral deviation). Under pairwise stability, one can ensure that each UAV j will not change its link formation strategy and therefore guarantee the promised performance for other UAVs in the network, and more specifically, to those connected to it or belong to its path p_j. Moreover, given that the graph resulting from our proposed network formation game is always a tree structure, G is pairwise stable if and only if no pair of UAVs can profitably deviate by *simultaneously* breaking one link and forming another. In other words, given UAVs j and i and any link $jk \in \mathcal{E}$, let $G' = G - jk + ji$, $p'_j = p_j - jk + ji$ and $p'_i = p_i + ji$ then:

$$U_j(p_j, G) < U_j(p'_j, G') \Rightarrow U_i(p_i, G) > U_i(p'_i, G'). \tag{6.13}$$

Note, however, that pairwise stable networks may not always exist. In particular, a pairwise stable network does not exist in case each network is defeated by some adjacent network, and these *improving paths* form *cycles*, as defined next.

Definition 6. An *improving path* is a sequence of networks $\{G_1, G_2, \cdots, G_k\}$ where each network G_k is defeated by the subsequent network G_{k+1}.

Definition 7. A *cycle* is an improving path $\{G_1, G_2, \cdots, G_k\}$ such that $G_1 = G_k$.

Consequently, a network is pairwise stable if and only if it has no improving paths emanating from it. In fact, for any network graph G, there exists either a pairwise stable network (or more) or a cycle of networks [129]. For network formation games, given that the strategy space is typically discrete, it is customary to characterize pairwise stable networks using an algorithmic approach, as the derivation of closed-form equilibrium policies is often infeasible [128]. As such, next, we propose a dynamic network formation algorithm that is guaranteed to reach a pairwise stable network upon convergence.

6.4 Distributed Dynamic Network Formation

Our proposed network formation algorithm allows the UAVs to adapt their location based on the resulting formed graph thus ensuring an efficient backhauling process. To realize this, we adopt the notion of virtual (artificial) force field which is introduced as follows.

6.4.1 Virtual force field

Given the initial locations of the UAVs, the formation of an aerial backhaul network might not be feasible in case UAVs are located outside each others' communication range. Therefore, to adjust the locations of the UAVs based on the links they want to form, a dynamic and self-organizing approach that allows adaptation to the dynamics of the network, is necessary. In this regard, we adopt the notion of virtual forces for UAVs [45]. A virtual force field allows a UAV to adjust its location by exerting forces of attraction and repulsion towards other UAVs. For our model, we consider the SNR as a metric for updating the value of the virtual force vector. In particular, to guarantee an efficient backhauling process, a minimum threshold value of SNR, denoted as $\widehat{\Gamma}$, should be achieved over

each of the formed links. This in turn allows the determination of the maximum distance between UAVs j and i, $d_{j,i}^{\max}$.

Remark 2. To guarantee a minimum threshold value of SNR between UAV j and UAV i, the distance between the two UAVs should not exceed $d_{j,i}^{\max}$, which is defined as:

$$d_{j,i}^{\max} = \sqrt{\frac{P_{j,i}}{\widehat{\Gamma} \cdot \sigma^2 \cdot 10^{\eta_{\text{LoS}}/10} \cdot (\frac{4\pi f_c}{c})^2}}, \tag{6.14}$$

where $P_{j,i}$ is the transmit power from UAV j to UAV i and c is the speed of light. The derivation of the expression of $d_{j,i}^{\max}$ follows from the SNR between UAV j and UAV i, $\Gamma_{j,i}$.

In fact, a virtual force can be expressed by a polar coordinate notation (r, θ) where r is its magnitude and θ its orientation angle. It can act as an attractive or a repulsive force, adapting to the actions of each UAV. For our proposed model, we consider an attractive virtual force from UAV j towards UAV i, when both UAVs agree on the formation of link ji but are out of each other's communication range. Therefore, the attractive force vector from UAV j towards UAV i, $\overrightarrow{\boldsymbol{F}}_{j,i}^{A}$, is given by:

$$\overrightarrow{\boldsymbol{F}}_{j,i}^{A} = \Big(u_A \cdot (d_{j,i} - d_{j,i}^{\max}), \theta_{j,i} \Big), \tag{6.15}$$

where u_A corresponds to the virtual force attractive coefficient and $d_{j,i}$ is the Euclidean distance between UAV j and UAV i. On the other hand, a repulsive force is exerted from UAV j towards UAV i, in case link ji is deleted and is given by:

$$\overrightarrow{\boldsymbol{F}}_{j,i}^{R1} = \Big(u_{R1} \cdot (d_{j,i} - d_{j,i}^{\max}), \theta_{j,i} + \pi \Big), \tag{6.16}$$

where u_{R1} is the virtual force repulsive coefficient and $d_{j,i}$ is the Euclidean distance based on initial locations. Moreover, for physical collision avoidance between different UAVs, we define the following repulsive force from UAV j towards UAV i:

$$\overrightarrow{\boldsymbol{F}}_{j,i}^{R2} = \Big(u_{R2} \cdot \frac{1}{d_{j,i}}, \theta_{j,i} + \pi \Big), \tag{6.17}$$

where u_{R2} corresponds to the virtual force repulsive coefficient for collision avoidance. Therefore, the total virtual force exerted from UAV j on UAV i can be

written as:

$$\overrightarrow{\boldsymbol{F}}_j = \sum_{i=1,i\neq j}^{J} \overrightarrow{\boldsymbol{F}}^{A}_{j,i} + \sum_{i=1,i\neq j}^{J} \overrightarrow{\boldsymbol{F}}^{R1}_{j,i} + \sum_{i=1,i\neq j}^{J} \overrightarrow{\boldsymbol{F}}^{R2}_{j,i}, \tag{6.18}$$

In our model, we consider that UAVs broadcast their initial locations at $t = 0$ and, hence, they can compute the corresponding virtual force vector even if they are not within each other's communication range. Therefore, given the strategies of each UAV j, its corresponding location is updated as follows:

$$x'_j = x_j + \overrightarrow{\boldsymbol{F}}^{x}_j, \;\; y'_j = y_j + \overrightarrow{\boldsymbol{F}}^{y}_j, \text{and} \;\; z'_j = z_j + \overrightarrow{\boldsymbol{F}}^{z}_j, \tag{6.19}$$

where x_j and x'_j are the initial and updated x-coordinate of UAV j, and $\overrightarrow{\boldsymbol{F}}^{x}_j$ is the x-component of $\overrightarrow{\boldsymbol{F}}_j$. Consequently, this location update procedure improves the achievable data rate for each UAV along its path and thus ensures an efficient backhauling process.

6.4.2 Dynamic network formation algorithm

Taking into account the location update of UAVs based on the defined virtual forces, we propose a myopic dynamic network formation algorithm. In particular, myopic players update their strategic decisions considering only the current state of the network without taking into account the future evolution of the network. To ensure the formation of a tree network architecture, link addition can be seen as link replacement and thus the strategy space of UAV j can be regarded as either a delete operation or a *replace* operation using which UAV j replaces its previously connected link with its parent node (if it exists) with a new link. Let $\mathcal{W}$ denote the set of possible nodes with which UAV j can possibly form or delete a link. We refer to $w \in \mathcal{W}$ as the *activated* node which corresponds to any of the other $(J-1)$ UAVs or the gateway node n. The adopted rules for the formation of the undirected network graph are:

1. UAV j can add a link with node w if both nodes j and w agree to add this link i.e., link addition is bilateral. Link jw is formed via a link replacement strategy if $U_j(p_j - jl + jw, G - jl + jw) > U_j(p_j, G)$ and $U_w(p_w + jw, G - jl + jw) > U_w(p_w, G)$ where node l corresponds to the parent node of UAV j (if it exists).

2. UAV j can delete link jw if $U_j(p_j - jw, G - jw) > U_j(p_j, G)$ i.e., link deletion can be unilateral.

3. Link replacement or deletion do not occur simultaneously.

Note that the gateway node is considered to be a passive player in our game.

For our network formation dynamics, we consider initially a star topology for G_0. Each iteration of our proposed algorithm consists of J rounds during which the UAVs engage in the network formation game in an arbitrary but sequential order. At a given round, UAV j chooses randomly (following a uniform distribution) another node w and takes an action with respect to w. Following the network formation rules, if link jw exists between the two nodes, then node j can delete this link if it is beneficial for it. If link jw is deleted, a repulsive force $\overrightarrow{\boldsymbol{F}}^R_{w,j}$ is exerted from UAV w towards UAV j thus returning UAV w to its initial location, in case of location update during previous iterations. On the other hand, if link jw does not exist, then UAV j can split from its parent node l and add link jw, if such a change is beneficial for both UAV j and the activated node w. Here, both nodes j and w can communicate with each other via a direct temporarily communication link that is established in order to decide whether link jw should be formed. Note that an attractive force $\overrightarrow{\boldsymbol{F}}^A_{j,w}$ is exerted from UAV j towards node w in case the corresponding two nodes are not within each other's communication range. If, at the end of the round, both nodes agree on the formation of link jw, UAV j updates its location to the current position. Otherwise, a repulsive force $\overrightarrow{\boldsymbol{F}}^R_{j,w}$ is exerted from UAV j towards node w, thus returning UAV j to its initial position at the beginning of this round. Note that $\overrightarrow{\boldsymbol{F}}^A_{j,w}$ and $\overrightarrow{\boldsymbol{F}}^R_{j,w}$ are exerted only when node w is not the gateway node. At the end of each round, UAV j and the activated node w update their corresponding location and path and broadcast such information to all other UAVs. After the convergence of the network formation algorithm, the UAVs are connected through a tree structure rooted at the gateway node. Consequently, data packets from (to) the SBSs to (from) the core network can now be transmitted using the resulting formed network tree structure G_{final}. The convergence complexity of our proposed myopic network formation algorithm is $O(J^2)$. A summary and a flow chart illustration of the proposed algorithm are given, respectively, in Algorithm 5 and Fig. 6.2.

Given the definition of pairwise stability and the proposed network formation rules, it can be clearly seen that, if the network formation process converges to a final network G, then G must be pairwise stable. However, proving the convergence of the network formation rules is challenging. In fact, if a pairwise stable network does not exist, then the proposed algorithm would involve cycles of

Algorithm 5: Proposed network formation algorithm.

Initialization:

Consider initially a star network G_0 where each UAV J is connected to the gateway node via a direct link.

Myopic network formation:

while G has not yet converged to a stable network, **do**

In a random but sequential order, the UAVs engage in the network formation game.

Step 1. UAV j activates another node w, in a random fashion but following a uniform distribution.

if $jw \in \mathcal{E}$ **then**

Step 2. UAV j deletes link jw if $U_j(p_j - jw, G - jw) > U_j(p_j, G)$.

if link jw is deleted **then**

Step 3. A repulsive virtual force $\overrightarrow{\boldsymbol{F}}^R_{w,j}$ is exerted from UAV w towards UAV j thus returning UAV w to its initial location.

end if

end if

if $jw \notin \mathcal{E}$ **then**

if UAV j and node w are not within each other's communication range **then**

Step 4. An attractive virtual force $\overrightarrow{\boldsymbol{F}}^A_{j,w}$ is exerted from UAV j towards node w thus updating the location of UAV j.

end if

Step 5. UAV j establishes a temporarily communication link with node w.

if $U_j(p_j - jl + jw, G - jl + jw) > U_j(p_j, G)$ and $U_w(p_w + jw, G - jl + jw) > U_w(p_w, G)$ where node l corresponds to the parent node of UAV j (if it exists) **then**

Step 6. Link jw is formed via a link replacement strategy.

else

Step 7. A repulsive virtual force $\overrightarrow{\boldsymbol{F}}^R_{j,w}$ is exerted from UAV j towards node w thus returning UAV j to its initial location, in case of location update at Step 6.

end if

end if

Step 8. UAV j and node w broadcast their updated locations and paths to all other nodes in the network.

end while

networks which are randomly visited over time [129]. Therefore, using simulation, we show in the following section that our proposed algorithm will converge.

6.5 Simulation Results and Analysis

For our simulations, we consider a 5 km $\times$ 5 km square area in which we randomly deploy a number of SBSs and UAVs. Table 6.2 summarizes the main simulation parameters. Note that the bandwidth per UAV is defined as the ratio of the total channel bandwidth B to the number of UAVs. All statistical results are averaged over 1000 independent runs.

Fig. 6.3 shows a snapshot of the tree graph resulting from the proposed algorithm for a network with $J = 10$ randomly deployed UAVs. From Fig. 6.3, we can see that most of the UAVs that are located far from the gateway engage in a

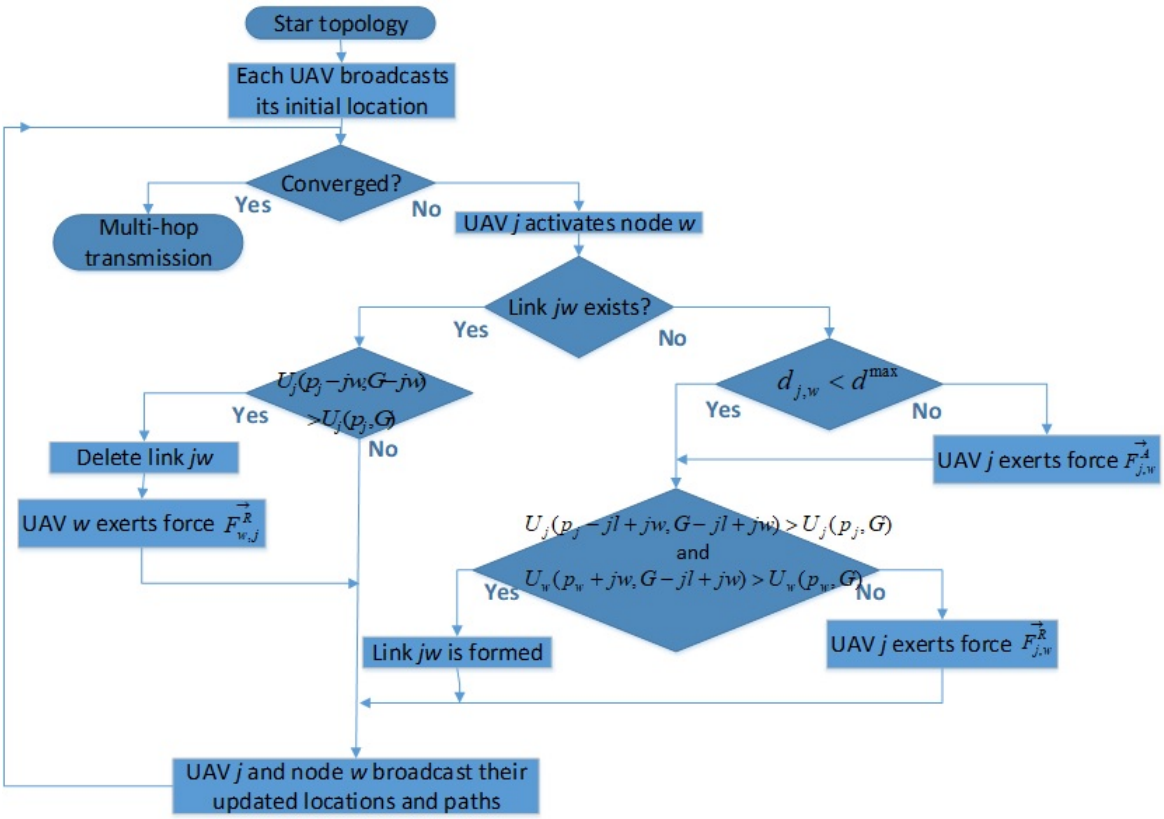

Figure 6.2. The flow chart of the proposed network formation algorithm.

Table 6.2 System parameters.

Parameters	Values	Parameters	Values
Max transmit power (P_o)	20 dBm	η_{LoS}	5 dB
SNR threshold ($\widehat{\Gamma}$)	-4 dB	η_{NLoS}	20 dB
Speed of light (c)	3×10^8 m/s	C	11.9
Channel bandwidth (B)	500 MHz	D	0.13
Noise variance (σ^2)	-90 dBm	u_A	1
Carrier frequency (f_c)	2 GHz	u_{R1}, u_{R2}	10
Packet arrival rate (λ_s)	(0, 1)	Packet size (υ)	2000 bits

multi-hop transmission with other UAVs that are located closer to the gateway thus extending the communication range of the network. Moreover, from this snapshot, we can see that the UAVs select their paths not only based on distance but also on the number of hops and traffic over a given path. For instance, UAV 9 connects to UAV 6, although UAV 7 is closer. This is due to the fact that the path for UAV 9 along UAV 7 involves 5 hops and is more congested as compared to 3 hops and less traffic when connected to UAV 6. This in turn decreases the latency along its path and thus improves its utility. From Fig. 6.3, we can also see the effect of the virtual force vector on the location of the UAVs. For instance,

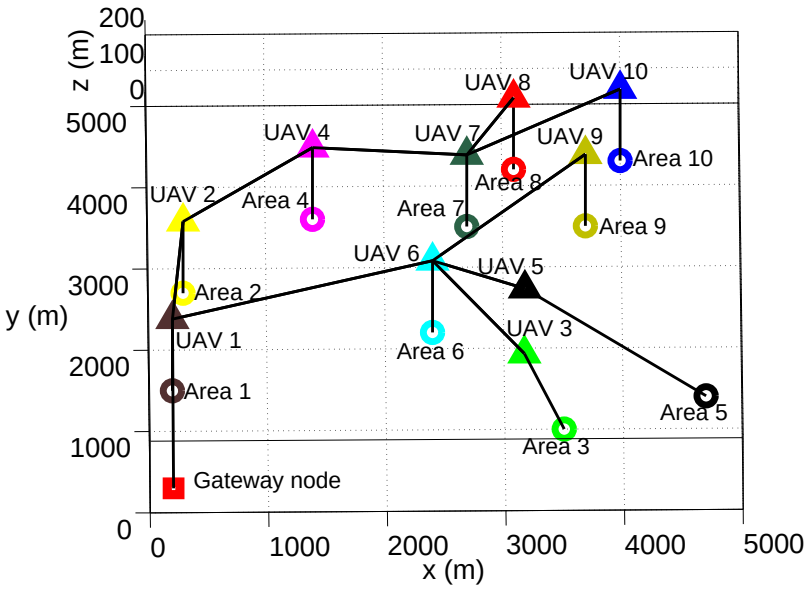

Figure 6.3. Snapshot of a tree graph formed using the proposed algorithm for a network with $J = 10$ randomly deployed UAVs. Circles represent target areas having one or multiple SBSs.

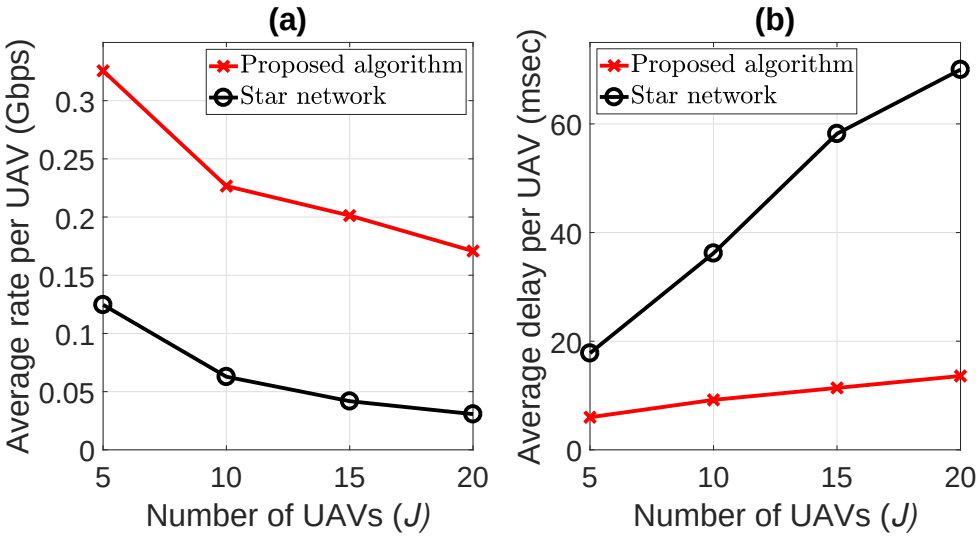

Figure 6.4. Performance assessment of the proposed network formation algorithm in terms of average (a) rate and (b) delay per UAV as compared to the star network, for different number of UAVs.

UAVs 3 and 5 adjust their initial location in order to guarantee an efficient communication link with UAV 6. Here, note that one could deploy more UAVs in case the location update of a particular UAV causes severe degradation in the A2G link connecting it to its serving SBSs.

Fig. 6.4 shows the average achievable rate and delay per UAV of the resulting network for our proposed scheme and the direct transmission approach consid-

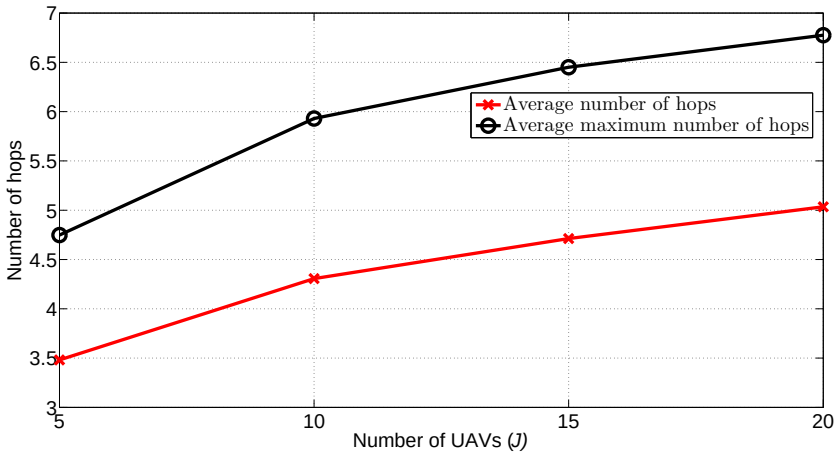

Figure 6.5. Average and average maximum number of hops in the final tree structure as a function of the number of UAVs J in the network.

ering a star topology. From Fig. 6.4, we can see that, at all network sizes, the proposed network formation algorithm yields significant performance gains in terms of both rate and delay reaching, respectively, up to 3.8 and 4-fold increase relative to the star network (for a network with 15 UAVs). The reason for this gain stems from the fact that multi-hop transmission allows UAVs having bad channel conditions with the gateway node to form links with other UAVs having better channel conditions. Here, note that the rate of the A2A links is higher than that of the A2G links due to the availability of a LoS communication links between different UAVs as well as the orthogonal channel allocation. Therefore, although more hops are formed, the average achieved rate over the multi-hop path is improved as compared to a direct link having weaker channel conditions. This in turn results in a higher service rate and thus a lower delay over the formed path. Here, note that the transmission bandwidth of each UAV is a function of the number of UAVs in the network. This in turn justifies the decrease in the average rate per UAV for both schemes as the number of UAVs increases. Fig. 6.4 (b) demonstrates that, although the delay for both schemes increases as the number of UAVs in the network increases from 5 to 20, the speed at which the delay increases for our proposed scheme (12.6%) is much smaller compared to that of the star network (29.3%). This is due to the fact that, for a given UAV j, the number of possible paths to the gateway node increases as the number of UAVs increases.

In Fig. 6.5, we show the average and the average maximum number of hops in the resulting network structure as the number of UAVs increases. Note that

the number of hops shown in this figure includes both the A2G links and the A2A links along a given path that connects the SBSs of each target area to the gateway node. From Fig. 6.5, we can conclude that as the number of UAVs increases, both the average and the average maximum number of hops in the tree structure increase. Specifically, the average and the average maximum number of hops vary, respectively, from 3.5 and 4.7 for $J = 5$, up to around 5 and 6.7 for $J = 20$. Note that both the average and average maximum number of hops increase slowly with the increase in the number of UAVs due to the delay cost induced for multi-hop transmission. For instance, one can notice that the average number of hops increases by around 1.5 hops and the average maximum number of hops increases by around 2 hops when the number of UAVs increases from 5 to 20.

Fig. 6.6 shows the minimum, average, and maximum number of iterations needed till convergence of our proposed network formation algorithm as the number of UAVs increases. From Fig. 6.6, we can see that our proposed network formation algorithm converges after a number of iterations and therefore a stable graph is reached. Moreover, we can note that as the number of UAVs increases, the total number of iterations required for the convergence of the algorithm increases. This result is due to the fact that, as J increases, the number of possible activated nodes w for a particular UAV j increases, and, thus, more actions (i.e., iterations) are required prior to convergence. For instance, the minimum, average and maximum number of iterations vary, respectively, from 4, 7.2, and 23 at $J = 5$ UAVs up to 18, 81, and 170 at $J = 20$ UAVs. Here, it is worth noting that practical UAV-based backhaul solutions will typically use only a relatively small number of UAVs, thus the convergence time resulting from our approach is practically reasonable.

6.6 Summary

In this chapter, we have proposed a novel UAV-based backhaul network design for wireless networks. We have formulated the problem as a network formation game in which the UAVs seek to form a multi-hop aerial network that connects SBSs to the core network. In particular, each UAV can take an individual decision to optimize its utility by exploiting the possible paths that connects it to the gateway node. To solve the game, we have proposed a distributed myopic algorithm which

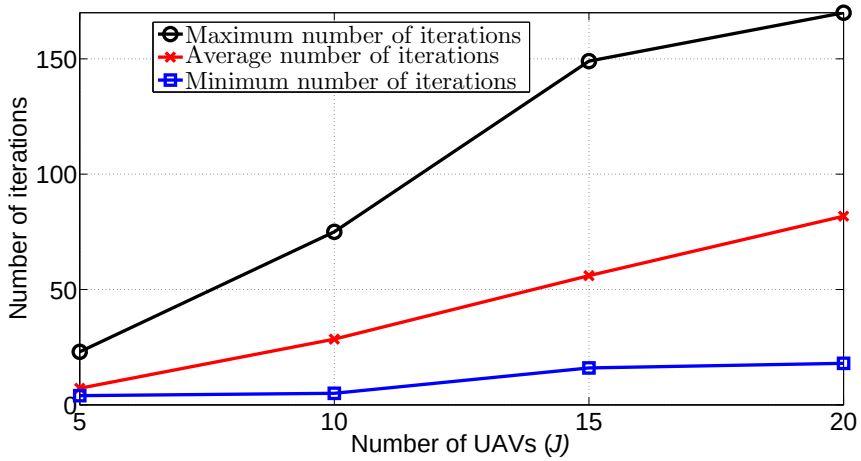

Figure 6.6. Minimum, average, and maximum number of iterations till convergence as a function of the number of UAVs J in the network.

is guaranteed to reach a pairwise stable network upon convergence. Simulation results have shown that the proposed approach yields significant performance gains in terms of delay and rate.

Chapter 7

Conclusion and Future Work

7.1 Summary and Conclusions

Next-generation wireless cellular networks must support extremely high data rates and radically new applications. To realize the 5G vision and allow the integration of a heterogeneous mix of wireless-enabled devices, these devices must be capable of adapting to the dynamics of the environment in an online and self-organized manner. Current centralized communication models and corresponding technologies may not be able to provide such massive connectivity. In essence, current models are intractable and would yield significant overhead for a network consisting of massive IoT devices. Consequently, there is a need for new communication models for making the 5G vision come into reality.

In this regard, the main scope of this thesis was to propose novel algorithms and methodologies that would incorporate self-organizing and intelligent decision making techniques across various components of future wireless cellular networks. Such edge intelligence is a key enabler of self-organizing solutions for optimizing the management of the available radio resources, user association, and data offloading. In this thesis, we have mainly focused on spectrum management for 5G networks, and in particular, the coexistence of LTE and WiFi in the unlicensed spectrum. We have also addressed some of the challenges that arise in the context of UAV-enabled wireless networks. Specifically, we have studied the problem of interference management for cellular-connected UAVs and the use of UAVs for providing backhaul connectivity to SBSs. The main contributions of this thesis can be summarized as follows:

- First, we have proposed a proactive resource allocation framework for the

coexistence of LTE-LAA and WiFi in the unlicensed band. We have formulated a game model where each SBS seeks to maximize its rate over a given time horizon while achieving long-term equal weighted fairness with WLAN and other LTE-LAA operators transmitting on the same channel. To solve this problem, we have developed a novel deep learning algorithm based on LSTM cells. The proposed algorithm enables each SBS to autonomously update their channel selection and channel access probabilities based on the traffic of WLAN on each of the unlicensed channels over a given time window. We have shown that the proposed deep RL algorithm reaches a mixed-strategy NE upon convergence. Simulation results have shown that the proposed scheme yields significant rate improvements compared to conventional reactive solutions such as instantaneous equal weighted fairness, proportional fairness and total network throughput maximization. The results have also shown that the proposed scheme prevents disruption to WLAN operation in the case large number of LTE operators selfishly deploy LTE in the unlicensed spectrum.

- Second, we have developed a holistic approach for LTE-LAA small cell traffic balancing by jointly optimizing the use of the licensed and unlicensed bands. The proposed scheme aims at achieving a proportional fair coexistence of WiFi, small cell and macro cell users by adapting the transmission probability of the LTE-LAA SBSs in the licensed and unlicensed bands. We have derived a closed form solution for the aforementioned optimization problem and proposed a transmission mechanism for the operation of the LTE-LAA SBS on both bands. Results have shown that our proposed scheme provides a better performance for WLAN when coexisting with LTE and an efficient utilization of the radio resources compared to alternative approaches from the literature as it allows a better tradeoff between maximizing the total network throughput and achieving fairness among all network flows.

- Third, we have presented an interference-aware path planning scheme for a network of cellular-connected UAVs. In particular, the proposed scheme enables each UAV at achieving a tradeoff between maximizing energy efficiency and minimizing both wireless latency and the interference level caused on the ground network along its path. We have formulated the prob-

lem as a dynamic noncooperative game in which the UAVs are the players. To solve the game, we have proposed a deep RL algorithm based on ESN cells in order to allow each UAV to decide on its next location, transmission power level, and cell association vector in an autonomous manner thus adapting to the dynamics of the network. We have shown that the proposed algorithm reaches an SPNE upon convergence. We have also derived upper and lower bounds for the altitude of the UAVs thus reducing the computational complexity of the proposed algorithm. Simulation results have shown that the proposed approach achieves better wireless latency per UAV and rate per ground UE while requiring a number of steps that is comparable to the shortest path scheme.

- Fourth, we have proposed a novel UAV-based backhaul scheme for linking the SBSs and the core network in scenarios where the ground backhaul is either unavailable or limited in capacity. The design of the aerial backhaul network is formulated as a network formation game in which the players are the UAVs. To solve the game, we have proposed a dynamic network formation algorithm that is guaranteed to reach a pairwise stable network upon convergence. The objective of the proposed network formation algorithm is to allow the UAVs to autonomously learn which A2A and A2G links to form in order to guarantee the connectivity of the SBSs to the core network. We have shown that the proposed approach achieves performance gains in terms of rate and delay reaching, respectively, up to 3.8 and 4-fold increase compared to the formation of direct communication links with the gateway node.

Next, we highlight on various challenges that warrant detailed exploration in future work.

7.2 Future Work

The proposed schemes in this thesis addressed some of the major challenges that arise in the context of spectrum management and the use of UAVs for wireless communications in next-generation cellular networks. In this section, we elaborate more on the possibility of extending the work presented in this thesis. Moreover, we overview some of other interesting challenges in the context of fu-

ture mobile networks that are worth investigating as part of our future work. In what follows, we summarize these problems:

- **Traffic balancing for multi-mode SBSs:** Future SBSs would essentially operate over multiple bands simultaneously such as the sub-6 GHz licensed band, sub-6 GHz unlicensed band, and the high-frequency mmWave band (licensed or unlicensed). In this context, it would be essential to design a traffic balancing scheme for steering the traffic of a multi-mode BS among the mmWave, the microwave, and the 5 GHz unlicensed band based on the availability of a LoS link, congestion on the microwave band and the availability of the 5 GHz unlicensed band. Moreover, it would be interesting for this framework to incorporate a context-aware scheduling algorithm for allocating the associated UEs of a multi-mode BS to the different types of available bands. This can build on some early work in this literature, such as [52].

- **Antenna tilting for efficient communication over the mmWave band:** Transmission over the mmWave band requires a LoS link due to the sensitivity of the mmWave spectrum to blockages. In this regard, it is essential to investigate machine learning techniques for adjusting the antenna tilt angle of SBSs operating over the mmWave spectrum. For instance, DNNs are capable of learning several features of the network environment and thus predicting the users mobility pattern. This in turn can allow the SBSs to determine the optimal tilt angle based on the availability of a LoS link and the UE data rate requirements.

- **Coexistence of WiGig and 5G radios:** As we mentioned earlier in Chapter 2, the mmWave spectrum is envisaged to be an important part of 5G multi-RAT ecosystem. In this regard, it is essential to propose novel and efficient interference mitigation and coexistence mechanisms for spectrum sharing with other technologies such as WiGig (802.11 ad). Therefore, one future direction in this context is to investigate a self-organizing beam scheduling scheme for the coexistence of WiGig and 5G radios in the 60 GHz band.

- **Classification of ground and aerial UEs:** Field measurements have shown that the radio propagation environment experienced by UAV-UEs

differs from that experienced by ground UEs [26]. In that case, network operators might want to allocate their radio resources differently between airborne UEs and ground UEs in order to maximize the total network performance. This in turn requires network operators to be able to differentiate an airborne UE from a ground one. To realize this, it is worthwhile investigating machine learning techniques, such as classification algorithms, for distinguishing airborne UEs from ground UEs, based on their wireless signal.

- **Interference mitigation for cellular-connected UAVs:** The ability of UAV-UEs to establish LoS connectivity with multiple cellular BSs can lead to substantial inter-cell mutual interference among the UAV-UEs and to the ground users, as discussed in Chapter 5. To address this challenge, one would need to investigate new improvements in the design of future cellular networks such as advanced receiver, cell coordination, 3D frequency reuse, and 3D beamforming. For example, UAVs can be equipped with a directional antenna whose beamwidth can be adjusted. Here, it is important to note that our proposed scheme in Chapter 5 can be readily used to accommodate to any of the aforementioned changes in the design of the network. Therefore, it would be interesting to combine the work in Chapter 5 with these solution methods in order to provide more efficient communication links for both, ground and aerial, users.

- **Admission control for cellular-connected UAVs:** Alongside the above mentioned solutions for interference management for UAV-UEs, a network operator may need to limit the admission of aerial vehicles in the network so that the perceived throughput performance of the conventional UEs is not deteriorated. In this regard, it is interesting to investigate admission control schemes for cellular-connected UAVs based on the data rate requirements of the cellular ground network.

- **Handover mechanism for cellular-connected UAVs:** Cellular-connected UAVs might face frequent handover and handover to distance cells thus resulting in ping-pong effect and an inefficient handover signalling. This is the consequence of the fact that UAVs can exhibit LoS links with multiple neighbouring BSs simultaneously which, along with dynamic channel variations, can result in a fluctuation in the quality of signal with these

neighbouring BSs. In this context, it is worth investigating ML techniques, and in particular RNNs, for addressing the issue of handover for cellular-connected UAVs. In fact, given their capability of dealing with time series data, RNNs can enhance mobility and handover in highly mobile wireless environments by learning the optimal mobility patterns of the UAVs thus decreasing the ping-pong effect among different ground BSs. For instance, a predictive mobility management framework can address critical handover issues, including frequent handover, handover failure, and excessive energy consumption for seamless handover in future wireless cellular networks.

Appendix A

Appendix

A.1 Possible Combinations for Solution Candidates of the Optimization Problem in Chapter 4

To derive a closed form solution for the proposed PF coexistence optimization problem in Chapter 4, we initially compute the solution candidates for a pair of optimal muting patterns on the licensed and unlicensed bands, (α,β), based on the Stationarity and the Complementary Slackness conditions of the KKT conditions. The proposed optimization problem involves inequality constraints and thus inequality Lagrangian multipliers λ_is. These multipliers can have a value of NZ or zero, depending on whether they bind at optimality or not, respectively. Thus, there exists 64 possible combinations for λ_is and the solution candidates satisfying the Complementary Slackness equations belong to one of these 64 possible combinations. However, out of the 64 possible combinations of λ_is, 8 of them only satisfying constraints (4.11)-(4.13) simultaneously. In what follows, we give a discussion about the reduction of the possible combinations of λ_is from 64 to 8.

If a constraint has finite values of both lower and upper bounds, one would need to consider the possible combinations when at most one Lagrange multiplier for that constraint is NZ. This is due to the fact that, at optimality, one or the other, or both, of the multipliers will always be equal to zero since only one of the bounds can be active at a time. For a given combination of λ_is, only one of the multipliers of either the lower or the upper bounds of the two variables α and β

could be NZ and thus the total number of possible combinations of λ_is is reduced from 64 to 36. An alternative way would be to combine the Lagrange multiplier of the lower bound with that of the upper bound into the same variable and use the sign of the variable to show which of the bounds is active. In other words, the value of the multiplier would be negative when the lower bound of a constraint is active, and positive when the upper bound of that same constraint is active or vice versa.

Consider the possible combinations when one of the multipliers of either the lower or the upper bounds of α and β are NZ. In that case, the values of α and β are deduced directly from the upper and lower bounds constraints as given below:

- The four different combinations, (NZ,NZ,NZ,0,NZ,0), (NZ,0,NZ,0,NZ,0), (0,NZ,NZ,0,NZ,0) and (0,0,NZ,0,NZ,0) result in (α,β) = (1,1).
- The four different combinations, (NZ,NZ,NZ,0,0,NZ), (NZ,0,NZ,0,0,NZ), (0,NZ,NZ,0,0,NZ) and (0,0,NZ,0,0,NZ) yield (α,β) = (0,1).
- The four different combinations, (NZ,NZ,0,NZ,NZ,0), (NZ,0,0,NZ,NZ,0), (0,NZ,0,NZ,NZ,0) and (0,0,0,NZ,NZ,0) result in (α,β) = (1,0).
- The four different combinations, (NZ,NZ,0,NZ,0,NZ), (NZ,0,0,NZ,0,NZ), (0,NZ,0,NZ,0,NZ) and (0,0,0,NZ,0,NZ) have the value of (0,0) for (α,β).

According to constraint (4.12), α should be less than or equal to β and hence the four different combinations resulting in (α,β)=(1,0) violate constraint (4.12) and thus cannot be a possible candidate solution to the optimization problem defined by (4.10)-(4.13).

For the optimization problem defined by (4.10)-(4.13), α is zero when either $\overline{R}_w$ or N_w are zero. However, in our system model, we consider only active users and assume that there is always at least one active user. Therefore, at optimality, α cannot be equal to zero and thus the possible combinations for λ_is resulting in (α,β) = (0,0) or (α,β) = (0,1) cannot be possible candidate solutions to the optimization problem defined by (4.10)-(4.13)

Moreover, the different possible combinations resulting in (α,β) = (1,1), i.e., (NZ,0,0,NZ,0,NZ) and (0,0,0,NZ,0,NZ), can be ignored. For these 2 possible combinations, λ_4 and λ_6 are NZ and thus the optimal solution lies at the intersection of the two lines α=1 and β=1. However, λ_2 is zero which states that the optimal

value does not lie on the straight line $\alpha = \beta$. This is a contradiction since the line $\alpha = \beta$ intersects with the point (1,1) and thus two combinations result in an infeasible solution. Therefore, the total number of possible combinations of λ_is is reduced from 36 to 22.

Following the previous discussion that α cannot be zero, the following combinations of λ_is can also be ignored since λ_3 is NZ for these cases: (NZ,NZ,NZ,0,0,0), (NZ,0,NZ,0,0,0), (0,NZ,NZ,0,0,0), (0,0,NZ,0,0,0). Moreover, β is equal to zero and $\alpha = \beta$ for the combinations (NZ,NZ,0,0,NZ,0) and (0,NZ,0,0,NZ,0). It implies that α should be zero; however, λ_3 is zero. Therefore, this is a contradiction and thus these two combinations result in an infeasible solution. Therefore, the total number of possible combinations of λ_is is reduced from 22 to 16.

Meanwhile, consider the possible combinations (NZ,NZ,0,NZ,0,0) and (0,NZ,0,NZ,0,0) where λ_2 and λ_4 are NZ and thus α=1 and α=β. This implies that $\beta = 1$, however, $\lambda_6 = 0$ and thus there is a contradiction and these 2 possible combinations can be neglected. Similarly for the combinations (NZ,NZ,0,0,0,NZ) and (0,NZ,0,0,0,NZ) where $\lambda_4 = 0$ which also results in a contradiction. Therefore, the total number of possible combinations of λ_is is reduced from 16 to 12.

Now, for the possible combinations (NZ,0,0,NZ,0,0) and (0,0,0,NZ,0,0), $\alpha = 1$. According to constraint (4.12), $\alpha \leq \beta$ and thus β should be 1 so that these combinations could be considered as possible candidates for the optimal solution of α and β. However, $\lambda_6 = 0$ i.e., this constraint does not bind at optimality and hence β cannot be equal to 1. Therefore, these 2 cases lead to an infeasible solution. Moreover, for the possible combinations (NZ,0,0,0,NZ,0) and (0,0,0,0,NZ,0), β is zero. Following constraint (4.12), $\alpha \leq \beta$ and thus α should be 0 for these cases to be feasible. However, λ_3 is 0 and thus the solution of α does not lie on the straight line α=0 and hence α cannot be zero. Therefore, these 2 combinations lead to infeasible solutions. Consequently, the total number of possible combinations of λ_is is reduced from 12 to 8, 2 of which have the same solution (α^*,β^*)=(1,1).

Consequently, the candidates for a pair of optimal value for (α, β) satisfying the stationarity and complementary slackness conditions are: (NZ,0,0,0,0,NZ), (0,0,0,0,0,NZ), (NZ,NZ,0,0,0,0), (NZ,0,0,0,0,0), (0,NZ,0,0,0,0), (0,0,0,0,0,0),

(NZ,NZ,0,NZ,0,NZ) and (0,NZ,0,NZ,0,NZ).

A.2 Proof of Theorem 2

For a given network state $\boldsymbol{v}_j(t)$ and a particular action $\boldsymbol{z}_j(t)$, the upper bound for the altitude of UAV j can be derived when UAV j aims at minimizing its delay function only, i.e., $\vartheta' = 0$. For such scenarios, UAV j should guarantee an upper limit, $\overline{\Gamma}_j$, for the SINR value $\Gamma_{j,s,c,a}$ of the transmission link from UAV j to BS s on RB c at location a as given in constraint (5.13). Therefore, $\hat{h}_j^{\max}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ corresponds to the altitude at which UAV j achieves $\overline{\Gamma}_j$ and beyond which (5.13) is violated. The derivation of the expression of $\hat{h}_j^{\max}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ is:

$$\sum_{c=1}^{C_{j,s}(t)} \Gamma_{j,s,c,a} = \overline{\Gamma}_j, \tag{A.1}$$

$$\sum_{c=1}^{C_{j,s}(t)} \frac{\frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t))}{C_{j,s}(t)} \cdot g_{j,s,c,a}(t)}{\left(\frac{4\pi \hat{f} d_{j,s,a}^{\max}}{\hat{c}}\right)^2 \cdot (I_{j,s,c}(t) + B_c N_0)} = \overline{\Gamma}_j, \tag{A.2}$$

$$\frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t))}{C_{j,s}(t)} \cdot \frac{1}{\left(\frac{4\pi \hat{f} d_{j,s,a}^{\max}}{\hat{c}}\right)^2} \cdot \sum_{c=1}^{C_{j,s}(t)} \frac{g_{j,s,c,a}(t)}{I_{j,s,c}(t) + B_c N_0} = \overline{\Gamma}_j, \tag{A.3}$$

$$(d_{j,s,a}^{\max})^2 = \frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t))}{C_{j,s}(t)} \cdot \frac{1}{\overline{\Gamma}_j \cdot \left(\frac{4\pi \hat{f}}{\hat{c}}\right)^2} \cdot \sum_{c=1}^{C_{j,s}(t)} \frac{g_{j,s,c,a}(t)}{I_{j,s,c}(t) + B_c N_0}, \tag{A.4}$$

where $d_{j,s,a}$ is the Euclidean distance between UAV j and its serving BS s at location a. Assume that the altitude of BS s is negligible, i.e., $z_s = 0$, $\hat{h}_j^{\max}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ can be expressed as:

$$\hat{h}_j^{\max}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) = \sqrt{\frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t))}{C_{j,s}(t) \cdot \overline{\Gamma}_j \cdot \left(\frac{4\pi \hat{f}}{\hat{c}}\right)^2} \cdot \sum_{c=1}^{C_{j,s}(t)} \frac{g_{j,s,c,a}(t)}{I_{j,s,c}(t) + B_c N_0} - (x_j - x_s)^2 - (y_j - y_s)^2}, \tag{A.5}$$

where x_s and y_s correspond to the x and y coordinates of the serving BS s and $\hat{c}$ is the speed of light.

Meanwhile, for a given network state $\boldsymbol{v}_j(t)$ and a particular action $\boldsymbol{z}_j(t)$, the lower bound for the altitude of UAV j can be derived when the objective function of UAV j is to minimize the interference level it causes on the ground network only, i.e., $\phi' = 0$ and $\varsigma = 0$. For such scenarios, the interference level that UAV j causes on neighbouring BS r at location a should not exceed a predefined value given by $\sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}$[1]. Therefore, $\hat{h}_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ corresponds to the altitude at which UAV j achieves $\sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}$ and below which the level of interference it causes on BS r exceeds the value of $\sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}$. The derivation of the expression of $\hat{h}_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ is given by:

$$\sum_{c=1}^{C_{j,s}(t)} \sum_{r=1,r\neq s}^{S} \frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t)) h_{j,r,c,a}(t)}{C_{j,s}(t)} = \sum_{c=1}^{C_{j,s}(t)} \sum_{r=1,r\neq s}^{S} \bar{I}_{j,r,c,a}, \tag{A.6}$$

$$\sum_{c=1}^{C_{j,s}(t)} \sum_{r=1,r\neq s}^{S} \frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t)) \cdot g_{j,r,c,a}(t)}{C_{j,s}(t) \cdot \left(\frac{4\pi \hat{f} d_{j,r,a}^{\min}}{\hat{c}}\right)^2} = \sum_{c=1}^{C_{j,s}(t)} \sum_{r=1,r\neq s}^{S} \bar{I}_{j,r,c,a}, \tag{A.7}$$

To find $\hat{h}_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$, we need to solve (A.7) for each neighbouring BS r separately. Therefore, for a particular neighbouring BS r, (A.7) can be written as:

$$\sum_{c=1}^{C_{j,s}(t)} \frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t)) \cdot g_{j,r,c,a}(t)}{C_{j,s}(t) \cdot \left(\frac{4\pi \hat{f} d_{j,r,a}^{\min}}{\hat{c}}\right)^2} = \sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}, \tag{A.8}$$

$$\frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t)) \cdot \sum_{c=1}^{C_{j,s}(t)} g_{j,r,c,a}(t)}{C_{j,s}(t) \cdot \left(\frac{4\pi \hat{f} d_{j,r,a}^{\min}}{\hat{c}}\right)^2} = \sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}, \tag{A.9}$$

$$(d_{j,r,a}^{\min})^2 = \frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t)) \cdot \sum_{c=1}^{C_{j,s}(t)} g_{j,r,c,a}(t)}{C_{j,s}(t) \cdot \left(\frac{4\pi \hat{f}}{\hat{c}}\right)^2 \cdot \sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}}, \tag{A.10}$$

[1] $\sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}$ is a network design parameter that is a function of the ground network density, number of UAVs in the network and the data rate requirements of the ground UEs. The value of $\bar{I}_{j,r,c,a}$ is in fact part of the admission control policy which limits the number of UAVs in the network and their corresponding interference level on the ground network [178].

where $d_{j,r,a}$ is the Euclidean distance between UAV j and its neighbouring BS r at location a. Assume that the altitude of BS r is negligible, i.e., $z_r = 0$, we have:

$$\hat{h}_{j,r}^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) = \sqrt{\frac{\widehat{P}_{j,s,a}(\boldsymbol{v}_j(t)) \cdot \sum_{c=1}^{C_{j,s}(t)} g_{j,r,c,a}(t)}{C_{j,s}(t) \cdot \left(\frac{4\pi \hat{f}}{\hat{c}}\right)^2 \cdot \sum_{c=1}^{C_{j,s}(t)} \bar{I}_{j,r,c,a}} - (x_j - x_r)^2 - (y_j - y_r)^2}, \tag{A.11}$$

Therefore, $\hat{h}_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ corresponds to the maximum value of $\hat{h}_{j,r}^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t))$ among all neighbouring BSs r and is expressed as:

$$\hat{h}_j^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)) = \max_r \hat{h}_{j,r}^{\min}(\boldsymbol{v}_j(t),\boldsymbol{z}_j(t),\boldsymbol{z}_{-j}(t)), \tag{A.12}$$

where x_r and y_r correspond to the x and y coordinates of other neighbouring BSs r. This completes the proof.

Bibliography

[1] F. Liu, E. Bala, E. Erkip, M. Beluri, and R. Yang, "Small cell traffic balancing over licensed and unlicensed bands," *IEEE Transactions on Vehicular Technology (TVT)*, Dec. 2014.

[2] ETSI, "ENI-experiential networked intelligence."

[3] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Machine learning for wireless networks with artificial intelligence: A tutorial on neural networks," *arXiv:1710.02913*, Oct. 2017.

[4] Ericsson, "Spectrum sharing," [White Paper], Oct. 2013.

[5] B. W. Forum, "Spectrum for small cells and authorised shared access," Oct. 2012.

[6] T. Rappaport, J. Murdock, and F.Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1390–1436, Aug. 2011.

[7] Rohde and Schwarz, "LTE basics webinar - part 03," [Online] https://ytd2525.wordpress.com/2013/01/23/lte-video-presentations/, May 2009.

[8] T. Nihtila, V. Tykhomyrov, O. Alanen, M. Uusitalo, A. Sorri, M. Moisio, S. Iraji, R. Ratasuk, and N. Mangalvedhe, "System performance of LTE and IEEE 802.11 coexisting on a shared frequency band," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*. Shanghai, China, Apr. 2013.

[9] "A beginner guide to recurrent networks and LSTMs," https://deeplearning4j.org/lstm.html.

[10] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[11] P. Popovski, "Ultra-reliable communication in 5G wireless systems," in *Proc. of International Conference on 5G for Ubiquitous Connectivity (5GU)*, Akaslompolo, Finland, Feb. 2014.

[12] N. A. Johansson, Y. P. E. Wang, E. Eriksson, and M. Hessler, "Radio access for ultra-reliable and low-latency 5G communications," in *Proc. of IEEE International Conference on Communication Workshop (ICCW)*, London, UK, Sept. 2015.

[13] O. Yilmaz, Y. Wang, N. Johansson, N. Brahmi, S. Ashraf, and J. Sachs, "Analysis of ultra-reliable and low-latency 5G communication for a factory automation use case," in *Proc. of IEEE International Conference on Communication Workshop (ICCW)*, London, UK, Sept. 2015, pp. 1190–1195.

[14] J. Ferber, *Multi-agent systems: An introduction to distributed artificial intelligence*. Addison-Wesley Reading, 1999, vol. 1.

[15] Z. Guan and T. Melodia, "CU-LTE: Spectrally-efficient and fair coexistencebetween LTE and Wi-Fi in unlicensed bands," in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*. San Francisco, CA, USA, Apr. 2016.

[16] M. Chen, W. Saad, and C. Yin, "Echo state networks for self-organizing resource allocation in LTE-U with uplink-downlink decoupling," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 3–16, Jan. 2017.

[17] S. Sagari, S. Baysting, D. Saha, I. Seskar, W. Trappe, and D. Raychaudhuri, "Coordinated dynamic spectrum management of LTE-U and Wi-Fi networks," in *Proc. of IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. Stockholm, Sweden, Sept. 2015.

[18] J. Perez-Romero, O. Sallent, R. Ferrus, and R. Agusti, "A robustness analysis of learning-based coexistence mechanisms for LTE-U operation in non-stationary conditions," in *Proc. of the 82nd IEEE Vehicular Technology Conference (VTC Fall)*. Glasgow, Scotland, Sept. 2015.

[19] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[20] U. Challita, L. Dong, and W. Saad, "Deep learning for proactive resource allocation in LTE-U networks," in *Proc. of the 23rd European Wireless Conference*. Dresden, Germany, May 2017.

[21] U. Challita, and L. Dong, and W. Saad, "Proactive resource management in LTE-U systems: A deep learning perspective," *arXiv:1702.07031*, Feb. 2017.

[22] E. Almeida, A. Cavalcante, R. Paiva, F. Chaves, F. Abinader, R. Vieira, S. Choudhury, E. Tuomaala, and K. Doppler, "Enabling LTE/WiFi coexistence by LTE blank subframe allocation," *IEEE International Conference on Communications (ICC)*, June 2013.

[23] M. Xing, Y. Peng, T. Xia, H. Long, and K. Zheng, "Adaptive spectrum sharing of LTE co-existing with WLAN in unlicensed frequency bands," *IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015.

[24] C. Cano and D. Leith, "Coexistence of WiFi and LTE in unlicensed bands: a proportional fair allocation scheme," *IEEE International Conference on Communication Workshop (ICCW)*, June 2015.

[25] U. Challita and M. K. Marina, "Holistic small cell traffic balancing across licensed and unlicensed bands," in *Proc. of ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, Malta, Nov. 2016.

[26] B. V. der Bergh, A. Chiumento, and S. Pollin, "LTE in the sky: Trading off propagation benefits with interference costs for aerial nodes," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 44–50, May 2016.

[27] X. Lin, V. Yajnanarayana, S. Muruganathan, S. Gao, and H. Asplund, "The sky is not the limit: LTE for unmanned aerial vehicles," *arXiv:1707.07534*, July 2017.

[28] M. Azari, F. Rosas, A. Chiumento, and S. Pollin, "Coexistence of terrestrial and aerial users in cellular networks," in *Proc. of IEEE Global Communications Conference (GLOBECOM) workshops*. Singapore, Dec. 2017.

[29] M. Azari, F. Rosas, and S. Pollin, "Reshaping cellular networks for the sky: The major factors and feasibility," *arXiv:1710.11404*, Oct. 2017.

[30] J. Yoon, Y. Jin, N. Batsoyol, and H. Lee, "Adaptive path planning of UAVs for delivering delay-sensitive information to ad-hoc nodes," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*. San Francisco, CA, USA, Mar. 2017.

[31] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3747–3760, June 2017.

[32] M. Messous, S. Senouci, and H. Sedjelmaci, "Network connectivity and area coverage for UAV fleet mobility model with energy constraint," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*. Doha, Qatar, Apr. 2016.

[33] X. Wang, A. Chowdhery, and M. Chiang, "Networked drone cameras for sports streaming," in *Proc. of International Conference on Distributed Computing Systems (ICDCS)*. Atlanta, Georgia, USA, June 2017.

[34] S. Zhang, Y. Zeng, and R. Zhang, "Cellular-enabled UAV communication: Trajectory optimization under connectivity constraint," *arXiv:1710.11619*, Oct. 2017.

[35] M. Lukoševicius, *A Practical Guide to Applying Echo State Networks*. Springer Berlin Heidelberg, 2012.

[36] U. Challita, W. Saad, and C. Bettstetter, "Deep reinforcement learning for interference-aware path planning of cellular-connected UAVs," in *Proc. of IEEE International Conference on Communicatios (ICC)*, Kansas City, MO, USA, May 2018.

[37] U. Challita, and W. Saad, and C. Bettstetter, "Cellular-connected UAVs over 5G: Deep reinforcement learning for interference management," *arXiv:1801.05500*, Jan. 2018.

[38] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "5G backhaul challenges and emerging research directions: A survey," *IEEE Access*, vol. 4, pp. 1743–1766, Apr. 2016.

[39] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, "Wireless backhauling of 5G small cells: Challenges and solution approaches," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 22–31, Oct. 2015.

[40] O. Semiari, and W. Saad, and M. Bennis, and Z. Dawy, "Inter-operator resource management for millimeter wave, multi-hop backhaul networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5258–5272, Aug. 2017.

[41] M. Alzenad, M. Shakir, H. Yanikomeroglu, and M. Alouini, "FSO-based vertical backhaul/fronthaul framework for 5G+ wireless networks," in *arXiv:1607.01472*, 2016.

[42] S. Shah, T. Khattab, M. Shakir, and M. Hasna, "A distributed approach for networked flying platform association with small cells in 5G+ networks," *arXiv:1705.03304*, Apr. 2017.

[43] S. Shah, and T. Khattab, and M. Shakir, and M. Hasna, "Association of networked flying platforms with small cells for network centric 5G+ C-RAN," *arXiv:1707.03510*, July 2017.

[44] M. Jackson, *Group Formation in Economics: Networks, Clubs, and Coalitions*, G. Demange and M. Wooders, Eds. Chapter 1: A Survey of Models of Network Formation: Stability and Efficiency: Cambridge University Press, 2005.

[45] A. Howard, M. J. Mataric, and G. S. Sukhatme, "Mobile sensor network deployment using potential field: a distributed scalable solution to the area coverage problem," in *Proc. of the International Conference on Distributed Autonomous Robotic Systems*. Fukuoka, Japan, June 2002.

[46] U. Challita and W. Saad, "Network formation in the sky: Unmanned aerial vehicles for multi-hop wireless backhauling," in *Proc. of the IEEE Global Communications Conference (GLOBECOM)*, Singapore, Dec. 2017.

[47] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[48] 5GPPP, "The 5G infrastructure public private partnership: the next generation of communication networks and services." Feb. 2015.

[49] J. Andrews, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Communications Magazine*, vol. 51, no. 3, pp. 136–144, Mar. 2013.

[50] Q. Li, H. Niu, A. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 71–78, Mar. 2014.

[51] O. Semiari, W. Saad, M. Bennis, and M. Debbah, "Performance analysis of integrated sub-6 GHz-millimeter wave wireless local area networks," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Washington DC, USA, Dec. 2017.

[52] O. Semiari, and W. Saad, and M. Bennis, and M. Debbah, "Joint millimeter wave and microwave resources allocation in cellular networks with dual-mode base stations," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4802–4816, July 2017.

[53] I. C. Network, "Authorised shared access (ASA). an evolutionary spectrum authorisation scheme for sustainable economic growth and consumer benefit," Paper presented at the 72nd Meeting of the WG FM, Miesbach, Germany, May 2011.

[54] 3GPP, "TR 36.889 - feasibility study on licensed-assisted access to unlicensed spectrum," Tech. Rep., July 2015.

[55] Nokia, "Nokia LTE for unlicensed spectrum," 2014, [White Paper].

[56] S. Rangan, T. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. of IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[57] T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[58] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[59] 3GPP, "3GPP RAN Workshop on LAA," Aug. 2015.

[60] Qualcomm, "Extending LTE advanced to unlicensed spectrum," White Paper, Dec. 2013.

[61] A. Cavalcante, E. Almeida, R. Vieira, S. Choudhury, E. Tuomaala, and K. Doppler, "Performance evaluation of LTE and Wi-Fi coexistence in unlicensed bands," *IEEE 77th Vehicular Technology Conference (VTC Spring)*, June 2013.

[62] "IEEE 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," *IEEE Std 802.11-2012*, pp. 1–2793, Mar. 2012.

[63] U. Challita, L. Al-Kanj, and Z. Dawy, "On LTE cellular network planning under demand uncertainty," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*. Instanbul, Turkey, Apr. 2014.

[64] U. Challita, Z. Dawy, G. Turkiyyah, and J. Naoum-Sawaya, "A chance constrained approach for LTE cellular network planning under uncertainty," *Computer Communications*, vol. 73, pp. 34–45, Jan. 2016.

[65] H. Cui, V. Leung, S. Li, and X. Wang, "LTE in the unlicensed band: Overview, challenges, and opportunities," *IEEE Wireless Communications*, Aug. 2017.

[66] Qualcomm Technologies, "Qualcomm research LTE in unlicensed spectrum: harmonious coexistence with Wi-Fi," Qualcomm, Tech. Rep., June 2014.

[67] Intel, "Alternative LTE solutions in unlicensed spectrum: overview of LWA, LTE-LAA and beyond," White Paper, 2016.

[68] Qualcomm, "More spectrum - licensed, unlicensed and ASA," [Online]: https://www.qualcomm.com/invention/technologies/1000x/spectrum, Nov. 2013.

[69] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, Feb. 2016.

[70] A. Ferdowsi and W. Saad, "Deep learning-based dynamic watermarking for secure signal authentication in the internet of things," in *Proc. of IEEE International Conference on Communications*. Kansas City, MO, USA, May 2018.

[71] I. Lee and K. Lee, "The internet of things (IoT): Applications, investments, and challenges for enterprises," *Science Direct*, vol. 54, no. 4, pp. 431–440, Aug. 2015.

[72] G. Wallace, "Amazon says drone deliveries are the future," [Online] http://money.cnn.com/2013/12/01/technology/amazon-drone-delivery/index.html, Dec. 2013.

[73] Google, "Project loon," [Online] https://x.company/loon/.

[74] 3D Robotics, "IRIS," [Online] https://3dr.com/support/iris/.

[75] Y. Zeng, R. Zhang, and T. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, May 2016.

[76] 3GPP, "Enhanced LTE support for aerial vehicles," Mar. 2017. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3231

[77] Qualcomm, "Paving the path to 5G: Optimizing commercial LTE networks for drone communication," [Online] https://www.qualcomm.com/news/onq/2016/09/06/paving-path-5g-optimizing-commercial-lte-networks-drone-communication, Sept. 2016.

[78] A. Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*. Austin, TX, USA, Dec. 2014.

[79] A. Sanjab, W. Saad, and T. Başar, "Prospect theory for enhanced cyber-physical security of drone delivery systems: A network interdiction game," in *Proc. of IEEE International Conference on Communications (ICC)*, Paris, France, May 2017.

[80] Google, "Waymo," [Online] https://waymo.com/.

[81] A. Ferdowsi, U. Challita, and W. Saad, "Deep learning for reliable mobile edge analytics in intelligent transportation systems," *arXiv:1712.04135*, Dec. 2017.

[82] J. Allen and F. Chevalier, "Report for vodafone-mobile backhaul market: Phase 2 report," Analysys Mason, Cambridge, U.K., Tech. Rep., 2014.

[83] U. Siddique, H. Tabassum, E. Hossain, and D. Kim, "Wireless backhauling of 5G small cells: Challenges and solution approaches," *IEEE Wireless Communications*, Oct. 2015.

[84] H. Shi, V. Prasad, E. Onur, and I. Niemegeers, "Fairness in wireless networks: Issues, measures and challenges," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 5–24, First Quarter 2014.

[85] C. Shannon and W. Weaver, *The mathematical theory of communication*. Urbana: University of Illinois Press, 1949.

[86] A. Sediq, R. Gohary, and H. Yanikomeroglu, "Optimal tradeoff between efficiency and jain's fairness index in resource allocation," in *Proc. of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. Sydney, Australia, Sept. 2012.

[87] "Machine learning: What it is and why it matters," https://www.sas.com/en_us/insights/analytics/machine-learning.html.

[88] E. Alpaydin, *Introduction to machine learning*. MIT press, 2014.

[89] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of Artificial Intelligence Research*, vol. 4, pp. 237–285, 1996.

[90] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural network design*. Martin Hagan, 2014.

[91] J. W. Kalat, *Biological psychology*. Nelson Education, 2015.

[92] J. Wang, "Analysis and design of a k-winners-take-all model with a single state variable and the Heaviside step activation function," *IEEE Transactions on Neural Networks*, vol. 21, no. 9, pp. 1496–1506, Sept. 2010.

[93] B. Müller, J. Reinhardt, and M. T. Strickland, *Neural networks: An introduction*. Springer Science & Business Media, 2012.

[94] Y. Chauvin and D. E. Rumelhart, *Backpropagation: Theory, architectures, and applications*. Psychology Press, 1995.

[95] P. Baldi, "Gradient descent learning algorithm overview: A general dynamical systems perspective," *IEEE Transactions on Neural Networks*, vol. 6, no. 1, pp. 182–195, Jan. 1995.

[96] M. Mazur, "A step by step backpropagation example," https://mattmazur.com/2015/03/17/a-step-by-step-backpropagation-example/.

[97] S. Klein, J. P. W. Pluim, M. Staring, and M. A. Viergever, "Adaptive stochastic gradient descent optimisation for image registration," *International Journal of Computer Vision*, vol. 81, no. 3, pp. 227–239, Aug. 2009.

[98] S. Ruder, "An overview of gradient descent optimization algorithms," *available online: arXiv:1609.04747*, Sept. 2016.

[99] K. Greff, R. Srivastava, J. Koutnik, B. Steunebrink, and J. Schmidhuber, "LSTM: a search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, Oct. 2016.

[100] Y. Nesterov, "A method of solving a convex programming problem with convergence rate O (1/k2)," in *Proc. of Soviet Mathematics Doklady*, vol. 27, no. 2, 1983.

[101] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of Machine Learning Research*, vol. 12, pp. 2121–2159, July 2011.

[102] M. D. Zeiler, "Adadelta: an adaptive learning rate method," *available online: arXiv:1212.5701*, Dec. 2012.

[103] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding," *available online arXiv:1510.00149*, Feb. 2016.

[104] D. P. Mandic and J. A. Chambers, *Recurrent neural networks for prediction: learning algorithms, architectures and stability*. Wiley Online Library, 2001.

[105] P. J. Werbos, "Backpropagation through time: What it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, Oct. 1990.

[106] R. J. Williams and D. Zipser, "Experimental analysis of the real-time recurrent learning algorithm," *Connection Science*, vol. 1, no. 1, pp. 87–111, Jan. 1989.

[107] H. Jaeger, "Short term memory in echo state networks," in *GMD Report*, 2001.

[108] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review*, vol. 3, no. 3, pp. 127–149, Aug. 2009.

[109] B. Farhang-Boroujeny, *Adaptive filters: theory and applications*. John Wiley & Sons, 2013.

[110] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, Apr. 2004.

[111] H. Jaeger, "Adaptive nonlinear system identification with echo state networks," in *Proc. of Advances in Neural Information Processing Systems*, Cambridge, MA, USA, 2002.

[112] J. J. Steil, "Backpropagation-decorrelation: online recurrent learning with O(n) complexity," in *Proc. of IEEE International Joint Conference on Neural Networks*, Budapest, Hungary, July 2004.

[113] D. Sussillo and L. F. Abbott, "Generating coherent patterns of activity from chaotic neural networks," *Neuron*, vol. 63, no. 4, pp. 544–557, Aug. 2009.

[114] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT press, 2016.

[115] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier networks," in *Proc. of Artificial Intelligence and Statistics (AISTATS)*, Fort Lauderdale, FL, USA, June 2011.

[116] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, ch. 7: Regularization for Deep Learning, http://www.deeplearningbook.org.

[117] T. Tieleman and G. Hinton, "Lecture 6.5—RmsProp: Divide the gradient by a running average of its recent magnitude," Technical report, 2012.

[118] J. Schmidhuber, "Learning complex, extended sequences using the principle of history compression," *Neural Computation*, vol. 4, pp. 234–242, Mar. 1992.

[119] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, June 2016.

[120] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. of International Conference on Learning Representations (ICLR)*. Banff, Canada, Apr. 2014.

[121] I. Sutskever, O. Vinyals, and Q. Le, "Sequence to sequence learning with neural networks," in *Proc. of International Conference on Neural Information Processing Systems (NIPS)*. Montreal, Canada, Dec. 2014.

[122] F. Gers, N. Schraudolph, and J. Schmidhuber, "Learning precise timing with LSTM recurrent networks," *Journal of Machine Learning Research*, vol. 3, pp. 115–143, Aug. 2002.

[123] K. Cho, B. Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. of Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, Oct. 2014.

[124] K. Greff, R. Srivastava, J. Koutnik, B. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.

[125] C. Gallicchio and A. Micheli, "Echo state property of deep reservoir computing networks," *Cognitive Computation*, vol. 9, pp. 337–350, May 2017.

[126] C. Gallicchio and A. Micheli, "Deep reservoir computing: A critical analysis," in *Proc. of European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, Bruges, Belgium, Apr. 2016.

[127] Z. Han, D. Niyato, W. Saad, T. Başar, and A. Hjorungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge University Press, 2011.

[128] M. O. Jackson and A. Wolinsky, "A strategic model of social and economic networks," *Journal of Economic Theory*, vol. 71, no. 1, pp. 44–74, Oct. 1996.

[129] M. Jackson and A. Watts, "The existence of pairwise stable networks," *Seoul Journal of Economics*, vol. 14, no. 3, pp. 299–321, 2001.

[130] A. Calvô-Armengo and R. Ílkiliç, "Pairwise-stability and nash equilibria in network formation," *International journal on Game theory*, vol. 38, pp. 51–79, 2009.

[131] R. Zhang, M. Wang, L. Cai, Z. Zheng, X. Shen, and L. Xie, "LTE-unlicensed: the future of spectrum aggregation for cellular networks," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 150–159, June 2015.

[132] B. Jia and M. Tao, "A channel sensing based design for LTE in unlicensed bands," in *Proc. of IEEE International Conference on Communication Workshop (ICCW)*. London, UK, June 2015.

[133] A. Mukherjee, J.-F. Cheng, S. Falahati, L. Falconetti, A. Furuskr, B. Godana, D. H. Kang, H. Koorapaty, D. Larsson, and Y. Yang, "System architecture and coexistence evaluation of licensed-assisted access LTE with IEEE 802.11," in *Proc. of IEEE International Conference on Communications (ICC)*. London, UK, June 2015.

[134] C. Hasan, M. K. Marina, and U. Challita, "On LTE-WiFi coexistence and inter-operator spectrum sharing in unlicensed bands: Altruism, cooperation and fairness," in *Proc. of ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*. Paderborn, Germany, July 2016.

[135] Y. Gu, Y. Zhang, L. X. Cai, M. Pan, L. Song, and Z. Han, "Exploiting student-project allocation matching for spectrum sharing in LTE-Unlicensed," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*. San Diego, CA, USA, Dec. 2015.

[136] S. Ha, S. Sen, C. Joe-Wong, Y. Im, and M. Chiang, "TUBE: Time dependent pricing for mobile data," in *Proceedings of Special Interest Group on Data Communication (ACM SIGCOMM)*. Helsinki, Finland, Aug. 2012.

[137] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Q. S. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Communication Magazine*, vol. 18, no. 3, pp. 22–30, June 2011.

[138] 3GPPP TS 36.300 v10.2.0, "E-UTRA and E-UTRAN; Overall description; Stage 2 (Release 10)," 2011.

[139] 3GPP, "Tr 36.889 - feasibility study on licensed-assisted access to unlicensed spectrum," July 2015.

[140] ETSI EN 301 893 V2.1.1, "5 GHz RLAN; harmonised standard covering the essential requirements of article 3.2 of directive 2014/53/eu," May 2017.

[141] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle sense: an optimal access method for high throughput and fairness in rate diverse wireless LANs," in *Proceedings of the annual conference of the Special Interest Group on Data Communication (SIGCOMM)*. Philadelphia, PA, Oct. 2005.

[142] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[143] J. Yi, W. Sun, S. Park, and S. Choi, "Performance analysis of LTE-LAA network," *IEEE Communications Letters*, no. 99, Dec. 2017.

[144] H. Gintis, *Game theory evolving: A problem-centered introduction to modeling strategic behavior*. Princeton University Press, 2000.

[145] E. Fehr and K. Schmidt, "A theory of fairness, competition and cooperation," *Quarterly Journal of Economics (QJE)*, vol. 114, no. 3, pp. 817–868, 1999.

[146] Y. Xing, R. Chandramouli, S. Mangold, and S. Shankar, "Dynamic spectrum access in open spectrum wireless networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 24, no. 3, pp. 626–637, Mar. 2006.

[147] M. Fukushima, "Restricted generalized Nash equilibria and controlled penalty algorithm," *Computational Management Science (CMS)*, vol. 8, no. 3, pp. 201–218, Aug. 2011.

[148] R. Sutton, D. McAllester, S. Singh, and Y. Mansour, "Policy gradient methods for reinforcement learning with function approximation," *Advances in Neural Information Processing Systems*, vol. 12, pp. 1057–1063, 2000.

[149] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, pp. 286–295, 1951.

[150] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*. The MIT Press, 1998.

[151] W. Zaremba, I. Sutskever, and O. Vinyals, "Recurrent neural network regularization," in *Proceedings of International Conference on Learning Representations (ICLR)*. San Diego, CA, May 2015.

[152] H. Sak, A. Senior, and F. Beaufay, "Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition," in *arXiv preprint arXiv:1402.1128*, 2014.

[153] R. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Machine Learning*, vol. 8, no. 3, pp. 229–256, May 1992.

[154] D. Rumelhart, G. Hinton, and R. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.

[155] E. Rozner, Y. Mehta, A. Akella, and L. Qiu, "Traffic-aware channel assignment in enterprise wireless LANs," in *Proc. of IEEE International Conference on Network Protocols (ICNP)*. Beijing, China, Oct. 2007.

[156] N. Burlutskiy, M. Petridis, A. Fish, A. Chernov, and N. Ali, *An Investigation on Online Versus Batch Learning in Predicting User Behaviour*, ser. Research and Development in Intelligent Systems XXXIII, M. Bramer and M. Petridis, Eds. Springer, 2016.

[157] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, June 2014.

[158] L. Peshkin, K. Kim, N. Meuleau, and L. P. Kaelbling, "Learning to cooperate via policy search," in *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI-2000)*. San Francisco, CA, USA, June 2000.

[159] K. J. Arrow and L. Hurwicz, "Stability of the gradient process in n-person games." *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 2, pp. 280–295, June 1960.

[160] H. Robbins and S. Munro, "A stochastic approximation method," *Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400–407, Sept. 1951.

[161] M. Balazinska and P. Castro, "IBM Watson Research Center," CRAWCAD, Feb. 2003.

[162] H. Shi, R. V. Prasad, E. Onur, and I. Niemegeers, "Fairness in wireless networks: issues, measures and challenges," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 5–24, First Quarter 2014.

[163] R. Jain, D.-M. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," *Digital Equipment Corporation, DEC-TR-301*, vol. 38, 1984, tech. rep.

[164] COINOR, "Basic open-source nonlinear mixed integer programming," [Online] http://www.coin-or.org/Bonmin/.

[165] C. Daniel, J. Taylor, and S. Nowozin, "Learning step size controllers for robust neural network training," in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*. Phoenix, Arizona USA, Feb. 2016.

[166] R. Zhang, M. Wang, L. Cai, Z. Zheng, X. Shen, and L. Xie, "LTE-Unlicensed: The future of spectrum aggregation for cellular networks," vol. 22, no. 3, pp. 150–159, June 2015.

[167] Y. Jian, C. Shih, B. Krishnaswamy, and R. Sivakumar, "Coexistence of Wi-Fi and LAA-LTE: Experimental evaluation, analysis and insights," *IEEE International Conference on Communication Workshop (ICCW)*, June 2015.

[168] M. Bennis, M. Simsek, A. Czylwik, W. Saad, S. Valentin, and M. Debbah, "When cellular meets WiFi in wireless small cell networks," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 44–50, June 2013.

[169] A. Elsherif, W. Chen, A. Ito, and Z. Ding, "Resource allocation and inter-cell interference management for dual-access small cells," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1082–1096, June 2015.

[170] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Q. S. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Communication Magazine*, vol. 18, no. 3, pp. 22–30, June 2011.

[171] Y. Wang and K. Pedersen, "Time and power domain interference management for LTE networks with macro-cells and HeNBs," in *Vehicular Technology Conference (VTC Fall)*, Sept. 2011.

[172] C. Cano and D. Leith, "Unlicensed LTE/WiFi coexistence: Is LBT inherently fairer than CSAT?" *IEEE International Conference on Communications (ICC)*, May 2016.

[173] IEEE 802.11 WG, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE Std 802.11, 2012.

[174] K. Duffy, "Mean field Markov models of wireless local area networks," *Markov Processes and Related Fields*, vol. 16, no. 2, Apr. 2010.

[175] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[176] 3GPP R4-092042, "Simulation assumptions and parameters for FDD HeNB RF requirement," 3GPP, Tech. Rep., 2009.

[177] A. Bhorkar, C. Ibars, and P. Zong, "Performance analysis of LTE and WiFi in unlicensed band using stochastic geometry," *IEEE 25th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2014.

[178] 3GPP, "Study on licensed-assisted access to unlicensed spectrum (release 13)," *TR 36.889 V0.1.1*, Nov. 2014.

[179] Qualcomm, "Paving the path to 5G: Optimizing commercial LTE networks for drone communication," [Online], Sept. 2016. [Online]. Available: https://www.qualcomm.com/news/onq/2016/09/06/paving-path-5g-optimizing-commercial-lte-networks-drone-communication

[180] T. Andre, K. Hummel, A. Schoellig, E. Yanmaz, M. Asadpour, C. Bettstetter, P. Grippa, H. Hellwagner, S. Sand, and S. Zhang, "Application-driven design of aerial communication networks," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 129–137, May 2014.

[181] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3949–3963, June 2016.

[182] Q. Wu, J. Xu, and R. Zhang, "Capacity characterization of UAV-enabled two-user broadcast channel," *arXiv:1801.00443*, Jan. 2018.

[183] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE Journal on Selected Areas on Communications (JSAC), Special Issue on Human-In-The-Loop Mobile Networks*, vol. 35, no. 5, pp. 1046–1061, May 2017.

[184] Y. Zeng, R. Zhang, and T. Lim, "Throughput maximization for UAV-enabled mobile relaying systems," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 4983–4996, Dec. 2016.

[185] M. Bekhti, M. Abdennebi, N. Achir, and K. Boussetta, "Path planning of unmanned aerial vehicles with terrestrial wireless network tracking," in *Proc. of Wireless Days*. Toulouse, France, May 2016.

[186] U. Mengali and A. D'Andrea, *Synchronization Techniques for Digital Receivers*, Plenum Press, Ed., New York, 1997.

[187] 3GPP TR 25.942 v2.1.3, "3rd generation partnership project; technical specification group (TSG) RAN WG4; RF system scenarios," Tech. Rep., 2000.

[188] D. Bertsekas and R. Gallager, *Data Networks*. Prentice Hall, Mar. 1992.

[189] A. Kumar, A. Abdelhadi, and C. Clancy, "A delay optimal MAC and packet scheduler for heterogeneous M2M uplink," *arXiv:1606.06692*, June 2016.

[190] S. Jahromizadeh and V. Rakocevic, "Joint rate control and scheduling for providing bounded delay with high efficiency in multihop wireless networks," *IEEE/ACM Transactions on Networking*, vol. 22, no. 5, pp. 1686–1698, Oct. 2013.

[191] W. Kwon, I. Suh, S. Lee, and Y. Cho, "Fast reinforcement learning using stochastic shortest paths for a mobile robot," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*. San Diego, CA, USA, Nov. 2007.

[192] M. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2004.

[193] H. Jaeger, M. Lukosevicius, D. Popovici, and U. Siewert, "Optimization and applications of echo state networks with leaky-integrator neurons," *Neural Networks*, vol. 20, no. 3, pp. 335–352, 2007.

[194] I. Szita and A. L. V. Gyenes, *Reinforcement Learning with Echo State Networks*. Springer, Berlin, Heidelberg, 2006, vol. 4131.

[195] R. Sutton and A. Barto, *Introduction to Reinforcement Learning*, 1998.

[196] A. Ghaffarkhah and Y. Mostofi, "Path planning for networked robotic surveillance," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3560–3575, July 2012.

[197] A. Kopeikin, S. Ponda, and G. Inalhan, *Control of Communication Networks for Teams of UAVs*. Dordrecht: Springer Netherlands, Aug. 2015, pp. 1619–1654.

[198] F. Ono, H. Ochiai, and R. Miura, "A wireless relay network based on unmanned aircraft system with rate optimization," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7699–7708, Nov. 2016.

[199] Y. Zeng, R. Zhang, and T. J. Lim, "Throughput maximization for UAV enabled mobile relaying systems," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 4983–4996, Dec. 2016.

[200] F. Ahdi and S. Subramanian, "Using unmanned aerial vehicles as relays in wireless balloon networks," in *Proc. of IEEE International Conference on Communications (ICC)*. London, UK, Sept. 2015.

[201] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile internet of things: Can UAVs provide an energy-efficient mobile architecture?" in *Proc. of IEEE Global Communications Conference (Globecom)*. Washington, DC, USA, Dec. 2016.

[202] Y. Zeng, R. Zhang, and T. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, May 2016.

[203] M. Sikora, J. Laneman, M. Haenggi, D. Costello, and T. Fuja, "Bandwidth- and power-efficient routing in linear wireless networks," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2624–2633, June 2006.

[204] W. Saad, Z. Han, T. Başar, M. Debbah, and A. Hjørungnes, "Network formation games among relay stations in next generation wireless networks," *IEEE Transactions on Communications*, vol. 59, no. 9, pp. 2528–2542, Sept. 2011.

[205] D. Bertsekas and R. Gallager, *Data networks*. Prentice Hall, Mar. 1992.