FACULTY OF ENGINEERING AND ARCHITECTURE

Efficient Spectrum Utilization and Coexistence of Heterogeneous Wireless Technologies Using Dynamic Spectrum Sharing

Irfan Jabandžić

Doctoral dissertation submitted to obtain the academic degree of
Doctor of Electrical Engineering

Supervisors
Prof. Ingrid Moerman, PhD - Spilios Giannoulis, PhD

Department of Information Technology
Faculty of Engineering and Architecture, Ghent University

September 2021

GHENT UNIVERSITY

FACULTY OF ENGINEERING
AND ARCHITECTURE

Efficient Spectrum Utilization and Coexistence of Heterogeneous Wireless Technologies Using Dynamic Spectrum Sharing

Irfan Jabandžić

Doctoral dissertation submitted to obtain the academic degree of
Doctor of Electrical Engineering

Supervisors
Prof. Ingrid Moerman, PhD - Spilios Giannoulis, PhD

Department of Information Technology
Faculty of Engineering and Architecture, Ghent University

September 2021

GHENT
UNIVERSITY

ISBN 978-94-6355-523-4
NUR 959, 958
Wettelijk depot: D/2021/10.500/71

Members of the Examination Board

Chair

Prof. Em. Daniël De Zutter, PhD, Ghent University

Other members entitled to vote

Prof. Jeroen Hoebeke, PhD, Ghent University
Vasilis Maglogiannis, PhD, Ghent University
Prof. Miklos Maroti, PhD, University of Szeged, Hungary
Michael Mehari, PhD, Ghent University
Emiliano Re, PhD, European Space Agency, the Netherlands

Supervisors

Prof. Ingrid Moerman, PhD, Ghent University
Spilios Giannoulis, PhD, Ghent University

Acknowledgements

Writing this part of the dissertation was the hardest for me. On the one hand, there are not enough words to express my gratitude to many great people for their contributions towards achieving my goals. On the other hand, I am not good at expressing my feelings. Therefore, I will keep this section short and to the point, but every word I mean from the bottom of my heart.

First, I would like to express my sincere gratitude to my supervisors Spilios Giannoulis and Ingrid Moerman who have guided me during the previous 5 years and were there for me not just as mentors, but as friends as well.

I would like to thank all my friends and colleagues who made my daily life enjoyable, offered their support in hard times, shared laughs and jokes in good times, and motivated me to improve both professionally and personally. When I was moving to Ghent for my PhD research, I was worried about the unknown and thorny road that awaits me. However, during my stay in Ghent, I met a lot of people who, little by little, became an important part of my life and helped me in my endeavors with their love and care, which I deeply appreciate.

Finally, special thanks belong to my parents Hamdija and Sabiha, and my precious sister Irfana, for their unconditional love, encouragement and support, this journey would not be possible without them. Naposljetku, posebne zahvale pripadaju mojim roditeljima Hamdiji i Sabihi, te mojoj dragoj sestri Irfani, za njihovu bezuslovnu ljubav, bodrenje i podršku, ovaj put ne bi bio moguć bez njih.

Ghent, September 2021
Irfan Jabandžić

Table of Contents

List of Figures

List of Tables

List of Acronyms

0-9

3G	3^{rd} Generation
4G	4^{th} Generation
5G	5^{th} Generation
6P	6top Protocol
6TiSCH	IPv6 over the TiSCH mode of IEEE 802.15.4e
6top	6TiSCH operation sublayer

A

ACI	Adjacent Channel Interference
ACK	acknowledgment
AI	Artificial Intelligence
API	Application Programming Interface
APSK	Amplitude and Phase Shift Keying
AP	access point

B

BS	base station
BT	Bonus Threshold

C

CAB	Coordinated Access Band

CAGR Compound Annual Growth Rate
CBRS Citizen Broadband Radio Service
CBR Constant Bit Rate
CB Control-Broadcast
CCA Clear Channel Assessment
CCC common control channel
CCH Color Constraint Heuristic
CC control channel
CDMA Code Division Multiple Access
CEPT European Conference of Postal and Telecommunications Administrations
CE Central entity
CIL CIRN Interaction Language
CIRN Collaborative Intelligent Radio Network
CL Control Layer
CNN Convolutional Neural Network
CODYSUN COllaborative and DYnamic approaches to increasing Spectrum UtilizatioN
CPU Central Processing Unit
CP collaboration protocol
CQI Channel Quality Indicator
CR Cognitive Radio
CSA-CCH Centralized Slot Assignment CCH
CSMA/CA Carrier Sense Multiple Access with Collision Avoidance
CSMA Carrier Sense Multiple Access
CTS Clear to Send
Ctrl Control

D

DARPA Defense Advanced Research Projects Agency
DBA Dynamic Bandwidth Allocation
DDMC-TDMA Dynamic Distributed Multi-Channel TDMA
DIFS DCF Interframe Space

DIMSUMnet Dynamic intelligent management of spectrum for ubiquitous mobile-access network

DRiVE Dynamic Radio for IP Services in Vehicular Environments

DSA-CCH Distributed Slot Assignment CCH

DSAP Dynamic Spectrum Access Protocol

DSA Dynamic Spectrum Access

DSS dynamic spectrum sharing

DTSR Dynamic TDMA Slot Reservation

DVB-RCS2 2^{nd} Generation Interactive Digital Video Broadcasting Satellite System

E

E-T-DRAND Energy-Topology DRAND

E-T Energy-Topology

EB-ET-DRAND Exponential Backoff and Energy-Topology DRAND

ECEF Earth Centered Earth Fixed

EIRP Equivalent Isotropic Radiation Power

EPFD Equivalent Power Flux Density

EP entry point

ES-CC Earth Station control channel

ESA European Space Agency

ESIM Earth Station in Motion

ES Earth Station

F

FCC Federal Communications Commission

FDD Frequency Division Duplexing

FDMA Frequency Division Multiple Access

FFT Fast Fourier Transform

FIR Finite Impulse Response

FPGA Field Programmable Gate Array

FPRP Five-Phase Reservation Protocol

FSM	Finite State Machine
FSS	Fixed Satellite Services
FS	Fixed Service
FWA	Fixed Wireless Access

G

GPS	Global Positioning System
GPU	Graphics Processing Unit
GSO-ES	geostationary Earth Stations
GSO-SAT	geostationary satellites
GSO	geostationary
GS	Ground Station
GW	gateway

I

ID	Identification
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
INCP	Incumbent Protection
IPv6	Internet Protocol version 6
IP	Internet Protocol
ISM	Industrial, Scientific and Medical
ISO	International Standards Organization
ITU	International Telecommunication Union
IoT	Internet of Things

K

KPI	Key Performance Indicator

L

L-DRAND	Localized-DRAND
LBT	Listen Before Talk
LEO	Low-Earth Orbit
LSA	Licensed Shared Access
LTE-LAA	LTE Licensed Assisted Access
LTE-U	LTE-Unlicensed
LTE	Long Term Evolution

M

M2M	Machine to Machine
MAC	Medium Access Control
MCS	Modulation and Coding Scheme
MF-TDMA	Multi-Frequency Time Division Multiple Access
MFS	Minimum Frequency Separation
MIMO	Multiple-Input Multiple-Output
ML	Machine Learning
MSF	Minimal Scheduling Function
MW	Microwave
McF-TDMA	Multi-Concurrent-Frequency Time Division Multiple Access

N

NGSO-ES	non-geostationary Earth Stations
NGSO-SAT	non-geostationary satellites
NGSO	non-geostationary
NI	National Instruments
NR-U	New Radio Unlicensed
NR	New Radio
NTP	Network Time Protocol

O

OFDM Orthogonal Frequency Division Multiplexing
OSI Open System Interconnection
OverDRiVE Spectrum Efficient Uni- and Multicast Over Dynamic Radio Networks in Vehicular Environments

P

PDU Protocol Data Unit
PER Packet Error Rate
PHY Physical layer
PPDU PLCP Protocol Data Unit
PSR Packet Success Rate
PTP Precision Time Protocol
PU Primary User

Q

QoS Quality of Service

R

RFMON RF Monitor
RF Radio Frequency
RSSI Received Signal Strength Indicator
RTS Request to Send
Rx Reception

S

SAT-CC satellite control channel

SATCOM	satellite communications
SC2	Spectrum Collaboration Challenge
SDR	Software Defined Radio
SF0	Scheduling Function Zero
SFX	Experimental Scheduling Function
SF	Scheduling Function
SIFS	Short Interframe Space
SINR	Signal-to-Noise Ratio
SMACS	Self-Organizing Medium Access Control for Sensor Networks
SML	Spectrum Management Layer
SNS3	Satellite Network Simulator 3
STDMA	Spatial TDMA
SU	Secondary User
sub-ms	submillisecond

T

TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TDM	Time Division Multiplexing
TR	Technology Recognition
Tx	Transmission

U

UDM	User Data Management
UDP	User Datagram Protocol
UHD	USRP Hardware Driver
USRP	Universal Software Radio Peripheral

V

V2X	Vehicle to Everything

W

WPAN Wireless Personal Area Network

WSN Wireless Sensor Network

Samenvatting
– Summary in Dutch –

Sinds de introductie in 1969 voor militaire doeleinden is het Internet snel uitgegroeid tot een onvervangbaar onderdeel van ons dagelijks leven. Tegenwoordig verbindt het bijna alles tot een groot wereldwijd communicatienetwerk, waardoor onze dagelijkse taken gemakkelijker te beheren zijn. Draadloze communicatie is een belangrijk onderdeel van het Internet dat nog continue evolueert. Steeds meer nieuwe draadloze technologieën komen op de markt, terwijl bestaande technologieën extra toestellen uitrollen om te blijven voldoen aan de toenemende eisen van hun gebruikers en om een bevredigende Quality of Service (QoS) te bieden.

Alle bestaande en opkomende draadloze technologieën maken gebruik van het elektromagnetische spectrum, dat per definitie vast is en niet kan opschalen om het hoofd te bieden aan de toenemende vraag naar draadloze bandbreedte. Het elektromagnetische spectrum is dus een schaarse bron die zo efficiënt mogelijk moet worden benut. Het gebruik van het radiospectrum wordt tegenwoordig nog altijd op dezelfde manier gereguleerd als decennia geleden, toen de vraag naar draadloos verkeer onvergelijkbaar lager was dan nu. Het radiospectrum is toegewezen volgens een statisch frequentieplan, waarbij het spectrum is opgedeeld in een groot aantal vaste frequentiebanden. Voor de meeste van deze banden zijn licenties toegekend aan welbepaalde draadloze operatoren of technologieën voor exclusief gebruik. Operatoren met een exclusieve licentie zijn bovendien niet verplicht om het spectrum optimaal te gebruiken, waardoor breedte van de frequentieband wordt bepaald door verkeerspieken, wanneer de vraag naar data het hoogst is. Het gebruik van deze exclusieve banden wordt daarom gekenmerkt door een ernstige over-dimensionering, waarbij het spectrum uiteindelijk zwaar onderbenut wordt, zowel in de tijd als in de ruimte. Andere frequentiebanden zijn licentievrij en draadloze technologieën kunnen deze banden kosteloos gebruiken in zoverre ze de vooropgestelde regelgeving respecteren. Licentievrije banden zijn omwille van deze gunstige eigenschappen zeer sterk gegeerd, wat leidt tot overbezetting van deze banden door meerdere draadloze technologieën. Zelfs al bestaan er regels om de onderlinge interferentie tussen verschillende technologieën in licentievrije banden te beperken, toch ondervinden draadloze netwerken in deze banden vaak ernstige hinder met verminderde prestaties tot gevolg.

Het huidige model om spectrum te wijzen is niet houdbaar voor de toekomst van draadloze communicatie. Het verder volgen van dit model zal de introductie van toekomstige draadloze diensten en technologieën sterk belemmeren, waardoor het onmogelijk zal worden om bestaande diensten en technologieën nog verder uit te breiden. Om dit te voorkomen, wordt algemeen erkend dat er nieuwe mechanismen moeten worden ontwikkeld voor een meer dynamische toewijzing van het spectrum, zowel in gelicentieerde als licentievrije spectrumbanden. De behoefte aan dynamisch delen van spectrum heeft de interesse gewekt in de onderzoeksgemeenschap rond draadloze netwerken, wat ertoe heeft geleid dat een aantal Dynamic Spectrum Access (DSA)-technieken reeds hun intrede hebben gedaan. In vergelijking met toewijzing van vaste spectrumbanden, zijn er merkbare verbeteringen vastgesteld qua spectrumefficiëntie bij het toepassen van deze technieken. Toch is het spectrumgebruik nog verre van optimaal. Daarom is het centrale thema van deze doctoraatsstudie het ontwerp en de implementatie van nieuwe modellen en mechanismen voor het dynamisch delen van spectrum.

Een ander belangrijk onderzoeksthema, waar technieken om het spectrum dynamisch te delen ook veel baat bij kunnen hebben, is het ontwerp van een betrouwbare Medium Access Control (MAC)-laag. De belangrijkste taak van de MAC-laag is om te bepalen wanneer en hoe meerdere draadloze toestellen toegang kunnen krijgen tot het gedeelde draadloze medium. Een degelijk ontworpen MAC-laag zal bijdragen tot een efficiënter spectrumgebruik, minder interferentie, en een harmonieuzere samenleving tussen verschillende netwerken, etc. Eén van de meest voorkomende MAC-protocollen in draadloze netwerken is Time Division Multiple Access (TDMA). In een TDMA netwerk krijgt elk toestel een welbepaald deel van het spectrum (een tijdslot) toegewezen voor verzenden en ontvangen van data, waarbij het vermijden van botsingen een primaire taak is om de netwerkprestaties te maximaliseren. Om het draadloze medium zo efficiënt mogelijk te delen en dus een hoge spectrumefficiëntie te garanderen, is een protocol nodig om een degelijk TDMA-schema op te stellen. Er bestaan vandaag vele TDMA-planningsprotocollen, elk ontwikkeld voor een specifiek toepassingsdomein. Er bestaat echter nog geen oplossing die tegemoet kan komen aan meerdere draadloze scenario's en toepassingsdomeinen.

Enkele van de belangrijkste uitdagingen die tijdens dit doctoraatsonderzoek aan bod komen, zijn de volgende:

- Dynamisch toewijzen van radiospectrum;
- Coëxistentie van heterogene technologieën in gedeelde frequentiebanden;
- Gedistribueerde en dynamische planning voor TDMA MAC-protocollen;
- Toepassen van dynamische spectrumdeling in satellietcommunicatie.

De volgende paragrafen zullen meer inzicht geven in de eerdergenoemde uitdagingen en lijsten de bijdragen in dit proefschrift op voor het oplossen van deze uitdagingen.

Dynamische allocatie van radiospectrum

Vanwege de schaarste aan radiospectrum en het toenemende spectrumgebruik is vaste toewijzing van radiospectrum geen haalbaar model voor de toekomst van draadloze communicatie. Bestaande DSA-modellen zijn ook geen ideale oplossingen; er zijn nog verschillende beperkingen die moeten worden aangepakt. Zo hebben bestaande DSA-modellen de neiging om prioriteit te geven aan primaire gebruikers ('Primary Users' of PU's), terwijl secundaire gebruikers ('Secondary Users' of SU's) zich aan strikte regels moeten houden om PU's te beschermen. Als gevolg hiervan kunnen SU's wellicht niet de vereiste QoS-garanties leveren. Deze doctoraatsstudie moedigt een verschuiving aan van licenties voor vaste exclusieve spectrumbanden naar dynamische allocatie van spectrum. In plaats van spectrum op te delen in vele vaste frequentiebanden, wordt aanbevolen om het spectrum te fragmenteren in zowel frequentie- als tijddomein, waardoor er meer flexibiliteit wordt geboden voor toekennen van licenties. Centrale entiteiten (CE's) van regelgevende autoriteiten kunnen het gebruik van spectrumfragmenten dynamisch beheren op basis van de eisen van draadloze operatoren. Het spectrum wordt beheerd via een gecentraliseerd onderhandelingsprotocol, dat bestaat uit allocatie, vrijmaking en verplaatsing van radiospectrumfragmenten.

Coëxistentie van heterogene technologieën in gedeelde frequentiebanden

Wanneer meerdere draadloze technologieën actief zijn in dezelfde frequentiebanden, is de grootste uitdaging ervoor te zorgen dat onderlinge interferentie wordt geminimaliseerd en een aanvaardbare coëxistentie wordt bereikt. In licentievrije banden worden bepaalde regels opgelegd aan draadloze technologieën, bijvoorbeeld limieten op het zendvermogen of verplicht gebruik van de Listen Before Talk (LBT)-schema's. Er bestaan verschillende coëxistentieschema's met als voornaamste doel het verbeteren van de paarsgewijze coëxistentie tussen populaire draadloze technologieën die gebruik maken van de licentievrije banden. Het is niet eenvoudig en hoogstwaarschijnlijk zelfs niet mogelijk om dezelfde benadering te gebruiken om algemene coëxistentie tussen alle reeds bestaande en toekomstige nieuwe draadloze technologieën te bereiken. Daarom stellen we in dit proefschrift gedistribueerde modellen voor om de harmonieuze coëxistentie van meerdere heterogene technologieën in dezelfde frequentiebanden te beheren. Ten eerste stellen we een raamwerk voor dat gebaseerd is op onderhandelingen tussen draadloze netwerken met overlappend zendbereik over het dynamisch gebruik van het licentievrij spectrum (via toegangspunten ('Entry Points' of EP's)). In de tweede benadering wordt coëxistentie bereikt door onderhandelingen tussen de knooppunten in een netwerk. Elk knooppunt kan dynamisch spectrum toekennen aan een TDMA slot en kan ook toegekende TDMA-sloten verwijderen. Op die manier kan elk knooppunt in een netwerk reageren op veranderingen in de vraag naar datatrafiek en bij detectie van externe interferentie.

Gedistribueerde en dynamische planning voor TDMA MAC-protocollen

Zoals eerder vermeld, bieden de meeste bestaande TDMA-protocollen enkel een oplossing om interne botsingen te vermijden in zeer specifieke draadloze scenario's. Om optimale netwerkprestaties te bereiken, verwachten we van een TDMA-protocol dat het rekening houdt met externe interferentie. Daarnaast moet het protocol zich ook bewust zijn van blootgestelde knooppunten ('exposed nodes'), die de capaciteit van draadloze netwerken en hergebruik van radiospectrum sterk kunnen belemmeren. De meeste bestaande TDMA-protocollen negeren echter de aanwezigheid van externe interferentie en 'exposed nodes'. Binnen dit proefschrift worden deze problemen aangepakt door het ontwerp van het Dynamic Distributed Multi-Channel TDMA (DDMC-TDMA) planningsprotocol. DDMC-TDMA richt zich niet op één specifiek draadloos scenario, maar biedt botsingsvrije schema's voor verschillende soorten draadloze netwerken in verschillende toepassingsdomeinen. Het nieuwe planningsprotocol werd in eerste instantie geïmplementeerd en geëvalueerd tijdens de Defense Advanced Research Projects Agency (DARPA) Spectrum Collaboration Challenge (SC2) voor kleinschalige draadloze netwerken. De schaalbaarheid wordt bewezen door een model te implementeren en simulaties uit te voeren in de ns-3 simulatoromgeving.

Dynamisch delen van spectrum in satellietcommunicatie

Net zoals bij aardse draadloze communicatiesystemen, wordt satellietcommunicatie gekenmerkt door een inefficiënte vaste toewijzing van het radiospectrum. Nieuwe satellietconstellaties worden massaal uitgerold in de ruimte om tegemoet te komen aan de toenemende vraag naar draadloos verkeer en dit resulteert in een explosie van toestellen die gebruik maken van satellietfrequentiebanden. Helaas worden veel satellietbanden ook gebruikt door aardse systemen en hierdoor is het probleem van co-existentie van meerdere heterogene systemen die in dezelfde banden nog prangender. Dynamisch delen van spectrum is dan ook een logische stap om tegemoet te komen aan de onderbenutting van spectrum en coëxistentie van heterogene netwerken in satellietbanden. De meeste bestaande oplossingen zijn beperkt tot coëxistentie tussen twee systemen die gebruik maken van dezelfde satellietbanden. In dit proefschrift worden twee gedistribueerde technieken voorgesteld die toelaten het spectrum dynamische te delen tussen drie verschillende systemen: twee satellietsystemen (geostationair (GSO) en niet-geostationair (NGSO)) en een systeem op aarde. De voorgestelde technieken worden geëvalueerd in de ns-3 simulatoromgeving voor downlink- en uplinkcommunicatie in Ka-band.

Summary

Since its introduction in 1969 for military purposes, Internet has quickly expanded to be an irreplaceable part of our daily life. Nowadays it connects almost everything into a big global communication network, making our daily tasks easier to manage. As an important part of the Internet network, wireless communications are constantly developing. New wireless technologies are emerging and entering the market and existing technologies need to deploy additional devices to comply with increasing user demands and offer satisfactory Quality of Service (QoS) to their users.

All the existing and emerging wireless technologies have to fit into the electromagnetic spectrum, which is fixed and cannot scale to follow the increase in requested wireless bandwidth. These factors lead to the electromagnetic spectrum being a scarce resource, which should be utilized as efficiently as possible. However, the utilization of radio spectrum is managed today in the same way as it was decades ago when wireless traffic demands were incomparably lower than today. The radio spectrum has been allocated according to a static frequency plan, dividing the spectrum into many fixed frequency bands. The majority of these bands are licensed for exclusive use by specific wireless operators or technologies. Operators assigned with these bands are not required to use the spectrum optimally. They usually lease chunks of the spectrum that satisfy their demands in a worst-case scenario, when traffic demands are the highest. Thus, usage of these exclusive bands is characterized by severe overprovisioning and underutilization both in time and geographically. The remainder of frequency bands are license-free and wireless technologies can freely operate in these bands as long as they adhere to a predefined set of standards. Due to their favorable characteristics, license-free bands are becoming the habitat of many wireless technologies, which results in overcrowding of these bands. Even with a set of rules that tend to limit mutual interference between different technologies operating in license-free bands, wireless networks in these bands may experience serious performance degradation.

Due to its limitations, the current model of spectrum allocation is not sustainable for the future of wireless communications. If preserved, this model may hinder the introduction of future wireless services and technologies, as well as a natural expansion of existing ones. In order to avoid this, it has been recognized that new mechanisms need to be explored for more dynamic spectrum allocation, both in licensed and license-free spectrum bands. The need for dynamic spectrum sharing has sparked research interest in the wireless networking community, resulting in several Dynamic Spectrum Access (DSA) techniques already being applied

in the real world. If compared to fixed spectrum allocation, there are noticeable improvements when applying these techniques. Even so, achieved spectrum utilization is still far from optimal. For these reasons, the design and implementation of new models and frameworks for dynamic spectrum sharing is the main focus of this PhD research.

Another important research topic, from which dynamic spectrum sharing techniques can greatly benefit, is the design of a reliable Medium Access Control (MAC) layer. The main task of the MAC layer is to determine when and how the wireless devices should access the shared wireless medium. Therefore, a well-designed MAC layer may contribute towards increased spectrum utilization, interference mitigation, coexistence between different networks, etc. One of the most common MAC protocols in wireless networks is Time Division Multiple Access (TDMA). In networks based on TDMA, each device is assigned a dedicated portion of the spectrum for transmission and reception, with the creation of a collision-free schedule being a primary task for achieving seamless network performance. A properly designed scheduling protocol for the TDMA scheme may assist in dynamic spectrum sharing and help towards the final goal of optimization of spectrum usage. To this date, a large number of TDMA scheduling protocols have been proposed, each addressing certain applications domain. However, there is currently no solution that addresses all the different wireless scenarios and application domains.

Some of the main challenges addressed throughout this PhD research are the following:

- Dynamic allocation of radio spectrum;
- Coexistence of heterogeneous technologies operating in the same frequency bands;
- Distributed and dynamic scheduling for TDMA-based MAC protocols;
- Applying dynamic spectrum sharing in satellite communications.

The following sections will give more insights into the previously mentioned challenges and present contributions within this thesis towards solving these challenges.

Dynamic allocation of radio spectrum

Due to radio spectrum scarcity and increasing spectrum usage, fixed allocation of radio spectrum is not a feasible model for the future of wireless communications. Likewise, existing DSA models are not ideal solutions; they possess a few shortcomings that need to be addressed. As one example, they tend to prioritize Primary Users (PUs), while Secondary Users (SUs) need to adhere to strict rules to protect PUs. As a result, SUs may not be able to deliver the required level of QoS. This PhD research recommends a shift from fixed spectrum licensing to dynamic spectrum assignment. Instead of spectrum division into many fixed frequency bands, the spectrum is fragmented both in frequency and time domain, thus offering more flexibility for licensing scheme. Central entities (CEs) of Regulatory Committees can dynamically manage the usage of spectrum fragments based on the demands of wireless operators. Spectrum management is achieved under the proposed centralized radio spectrum negotiation protocol, which consists of allocation, deallocation, and move operations of radio spectrum fragments.

Coexistence of heterogeneous technologies operating in the same frequency bands

When multiple wireless technologies operate in the same frequency bands, the main challenge is to ensure that mutual interference is minimized and proper coexistence achieved. In license-free bands, certain rules are imposed on wireless technologies, for example, limits on transmission power or required usage of the Listen Before Talk (LBT) schemes. Several coexistence schemes exist, but the main goal of these schemes is to improve coexistence between single pairs of major wireless technologies operating in unlicensed bands. Using this approach to achieve full coexistence between all existing and emerging wireless technologies is an impractical and probably impossible task. Thereby, in this dissertation, we propose distributed models for managing the coexistence of heterogeneous technologies that are sharing the same frequency bands. First, we propose a framework based on dynamic spectrum negotiation between wireless networks (their entry points (EPs)) operating in the same collision domain in unlicensed bands. In the second approach, coexistence is achieved on a node level. Each node executes dynamic spectrum allocation and removal operations of TDMA slots, reacting accordingly to changes in traffic demands and detected external interference.

Distributed and dynamic scheduling for TDMA-based MAC protocols

As already mentioned, many existing TDMA scheduling protocols provide internal collision-free schedules focusing on single wireless scenarios. To achieve seamless network performance, TDMA scheduling protocols should also take into account external interference. Moreover, they should address the exposed node problem, which may hinder the capability of wireless networks for spatial reuse of radio spectrum. However, existing protocols are often neglecting the possibility of external interference and exposed node problem. Within this dissertation, these problems are tackled by the design of the Dynamic Distributed Multi-Channel TDMA (DDMC-TDMA) scheduling protocol. DDMC-TDMA does not focus on a single wireless scenario but provides collision-free schedules for various types of wireless networks in different application domains. It is initially implemented and evaluated during Defense Advanced Research Projects Agency (DARPA) Spectrum Collaboration Challenge (SC2) for small-scale wireless networks. Later, its scalability is proven by implementing a model and executing simulations in the ns-3 simulator environment.

Applying dynamic spectrum sharing in satellite communications

Similar to terrestrial wireless communications, satellite communications are characterized by inefficient fixed allocation of the radio spectrum. The surge in

wireless traffic demands is also reflected in plans for the deployment of new satellite constellations, leading to a significant increase in the number of devices operating in satellite frequency bands. In addition, many satellite bands are already used by terrestrial systems, further emphasizing the problem of the coexistence of multiple heterogeneous systems operating in the same bands. Dynamic spectrum sharing represents a natural step towards solving problems of underutilization and coexistence of heterogeneous networks in satellite bands. Several solutions exist, each investigating the coexistence between two types of systems operating in satellite bands. We propose two distributed dynamic spectrum sharing techniques that offer coexistence between all three systems (geostationary (GSO), non-geostationary (NGSO), terrestrial) sharing the same satellite bands. The proposed techniques are evaluated in the ns-3 simulator environment for downlink and uplink communication in Ka-band.

1
Introduction

"I will either find a way, or make one."

–Hannibal (247 BC - 182 BC)

This chapter situates the conducted research work, summarizes the main contributions and outlines the structure of this dissertation. It also provides an overview of the publications that were authored during this research period.

1.1 Background

Since its introduction, the growth and importance of the Internet have overcome all expectations. The Internet has progressed from a specialized military network to a necessity, with modern life unimaginable without it. From interconnecting stationary infrastructures like cities, houses, schools, etc., enabling easier task management in the health, education, industrial, and entertainment sector, to providing connections to mobile users like humans, devices, and even animals, the Internet is ever-present. Even the slightest degradation in the performance of the Internet may lead to various problems in any of the abovementioned application domains. With constant increases in user demands, in the number of available wireless technologies, and the number of Internet-based devices, technological advancements need to guarantee seamless user experience or currently relatively reliable tools might become a thing of the past.

Not only that more and more people have access to the Internet and use it in their daily activities, but in recent decades there was an introduction of the Internet of Things (IoT), Machine to Machine (M2M) communication, Vehicle to Everything (V2X) communication, deployment of many new satellite constellations, etc., all leading to growth in Internet demands. Based on the predicted Compound Annual Growth Rate (CAGR), the Cisco forecast (Fig. 1.1) shows that growth in user demands is mainly in the area of wireless communications, with wireless (Wi-Fi and mobile) devices accounting for 79% of Internet traffic in 2022, which is an increase from 65% in 2017. Globally, it is predicted that the average number of devices and connections per capita will grow from 2.4 in 2017 to 3.6 by 2022 [1].

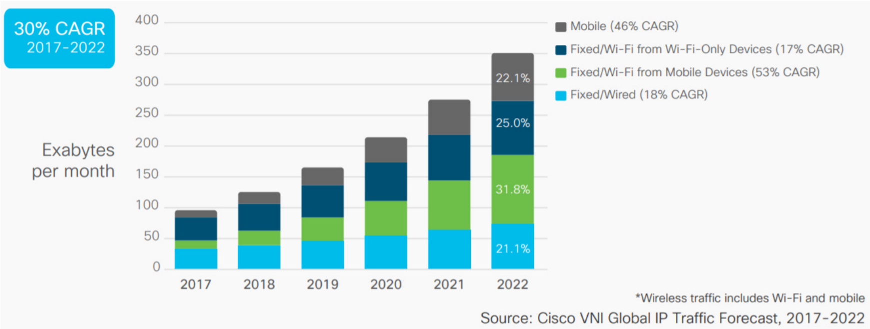


Figure 1.1: Global internet traffic, wired and wireless [1].*

New wireless devices and technologies need to be incorporated into the existing available spectrum, which is fixed and cannot follow the growth in wireless demands. As a consequence, this will soon lead to a spectrum shortage. Internet reliability mainly depends on the reliability of its wireless communications segment and the ability of various wireless technologies to coexist and not interfere with each other in the frequency bands they share. The importance of the radio spectrum was recognized early and its usage started to get monitored and controlled from the 1930s, with the introduction of administrative licensing. Radio spectrum got divided into many licensed and unlicensed frequency bands, with such division being maintained up to today.

Traditionally, licensed frequency bands are assigned to the highest bidders, enabling them to use assigned chunks of the spectrum for a predefined period of time (typically years) and for stated purposes using stated emission parameters [2]. Licensees are given associated licenses for a use of assigned frequency bands with imposed requirements not to induce interference to adjacent frequency bands. Fixed spectrum assignment is an effective way to avoid harmful inter-network interference. Therefore such an approach is effective as long as the total bandwidth demands can fit into the available radio spectrum. However, it is recognized that fixed spectrum allocation is not feasible for the future, as once all the licensed bands are assigned, there would be no free spectrum for new operators. This could further hinder future innovations in the field of wireless communications.

Licensing scheme that is employed today often leads to wireless services/operators applying for radio spectrum that would satisfy their demands in densely populated areas during rush-hours, leading to vast underutilization of licensed spectrum, both in time and geographically. So the actual lack of available spectrum is not a result of spectrum scarcity but rather a consequence of static and inefficient assignment policy. This fact is confirmed by several studies [3–6] where spectrum utilization is measured over time in various geographical areas. The Federal Communications Commission (FCC) report from 2002 [7] showed that most of the allocated spectrum was vastly underutilized, with temporal and geographical fluctuations in the real usage of the assigned spectrum ranging from 15% to 85% in the frequency bands below 3 GHz. In the frequency bands above 3 GHz, underutilization is an even more highlighted issue.

As it is recognized that the traditional model of spectrum allocation is not feasible to support wireless demands in the future, there are already numerous research efforts and standards that are suggested as a potential replacement or supplement to the existing licensing model. Among these proposals, several have proven to be effective solutions for reducing the spectrum underutilization, such as Cognitive Radio (CR), TV white spaces, Citizen Broadband Radio Service (CBRS), and Licensed Shared Access (LSA). These proposals are mainly based on the concept of Dynamic Spectrum Access (DSA), where the spectrum is dynamically shared between the wireless networks. Networks dynamically adapt their transmission characteristics to opportunistically utilize the available spectrum, taking the advantage of abovementioned temporal and spatial underutilization of the spectrum. In DSA, wireless devices belonging to unlicensed networks, so-called Secondary Users (SUs), are allowed to utilize the spectrum that is temporarily unused by wireless devices of licensed networks, so-called Primary Users (PUs). To do so, SUs are required to adhere to agreements and constraints imposed by PUs, ensuring that SUs are not imposing significant interference to PUs. For example, SUs may be required to use Listen Before Talk (LBT), in order to monitor the activity of PUs and back off from using the spectrum if PUs transmissions are detected.

The problem of radio spectrum underutilization was recognized by American Defense Advanced Research Projects Agency (DARPA) that launched Spectrum Collaboration Challenge (SC2) [8] in 2016 as an initiative for creating a future spectrum sharing paradigm. The goal of the SC2 competition was to ensure that the exponentially growing number of wireless devices would have full access to the increasingly crowded electromagnetic spectrum. Competitors reimagined spectrum access strategies and developed a new wireless paradigm in which radio networks autonomously collaborated and reasoned about how to share the radio spectrum, avoid interference and jointly exploit opportunities to achieve the most efficient use of the available spectrum. The teams in the competition applied CR, Artificial Intelligence (AI), and Machine Learning (ML) techniques for efficiently sharing spectrum and meeting the Quality of Service (QoS) requirements in a distributed way. Further, novel ideas in the areas of Medium Access Control (MAC) and Physical layer (PHY) were introduced. The basic concept of spectrum sharing in DARPA SC2 is illustrated in Fig. 1.2. By removing exclusively assigned spectrum and offering a wider shared spectrum band to multiple independent and collaborative networks, the current problem of underutilization in exclusive spectrum bands can be mitigated. As can be observed in Fig. 1.2, if collaboration and coexistence are achieved, there are much fewer unused spectral resources in shared spectrum bands. In DARPA SC2 competition, increased spectrum efficiency has been proven in a shared band compared to a fixed allocation of multiple exclusive bands to individual networks.

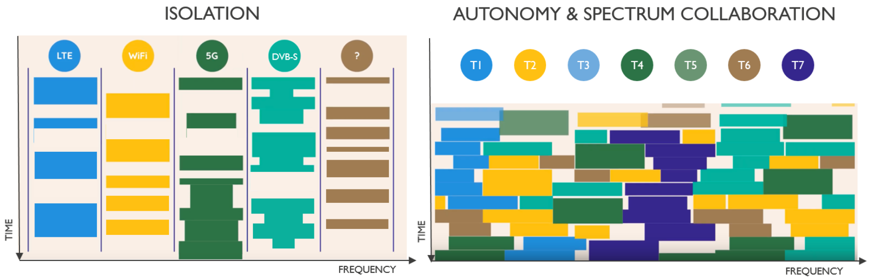


Figure 1.2: Concept of dynamic spectrum sharing adopted in the DARPA Spectrum Collaboration Challenge: from isolation (left) to autonomous collaboration (right) of radio networks for sharing spectrum.

In comparison to licensed frequency bands, unlicensed frequency bands are free for use by an unlimited number of users. Technologies operating in these bands only need to adhere to certain standards and rules, which aim is to eliminate potential interference and provide equal conditions to all present technologies. Two examples of these rules are tight limits on transmission power and requirements on advanced coexistence capabilities. Industrial, Scientific and Medical (ISM) bands at 900 MHz and 2.4 GHz are the oldest unlicensed bands, which are later expanded with unlicensed bands at different frequencies, for example, at 5 GHz and 5.8 GHz. Due to their favorable physical characteristics and freedom of use, ISM bands are very popular for various wireless technologies, like Wi-Fi, Bluetooth, ZigBee, etc., while also LTE and 5th Generation (5G) New Radio Unlicensed (NR-U) aspire to enter these bands. With an increase in the number of wireless technologies, demand for unlicensed bands increases as well, leading to overcrowding and congestion of these bands. Thereby, unlicensed band deployments need to operate in the presence of high interference levels [9]. Overcrowding and congestion affect the quality of wireless communication in unlicensed bands and require huge efforts to enable the coexistence of various heterogeneous technologies residing within these bands.

With the increasing demands in wireless communications, other areas of wireless technologies and standards need further improvement. One of those areas is the MAC layer. For achieving the coexistence of various heterogeneous technologies operating in the unlicensed bands, as well as supporting dynamic spectrum sharing in licensed bands, the MAC layer plays an important role. The MAC layer is a part of the Open System Interconnection (OSI) model, which is introduced in 1978 by the International Standards Organization (ISO) [10] with a purpose to establish a set of standards that communication hardware and software must abide by in order to connect with other devices. The main task of the MAC layer in wireless communications is to regulate the access of devices to a shared medium and minimize collisions. This means that MAC should determine when and how wireless devices access the medium and send their data. Further, a well-designed MAC protocol should target optimal spectrum utilization, maximize throughput, as well as minimize the delay. With an increase in traffic demands and the number of wireless devices, MAC should be able to achieve all these tasks while being able to adapt to any changes in the wireless networks, such as a change in network size or density, mobility of the devices, dynamic traffic patterns, etc.

While there are various MAC protocols employed in communication systems nowadays, contention-based and scheduling/reservation-based MAC protocols are the most prominent. Contention-based MAC protocols typically use the Carrier Sense Multiple Access (CSMA) technique where wireless devices are required to listen to the shared medium before transmitting a packet. These protocols are attractive because they are easy to implement with high flexibility and efficiency for low traffic load conditions and small network sizes [11]. However, they suffer from collisions when the traffic load and network density increases. On the other hand, in scheduling/reservation-based MAC protocols, a transmission schedule is generated with devices assigned dedicated chunks of spectrum for transmission and reception, thus avoiding any contention for accessing the shared medium [12]. As real-time traffic demands continue to grow, these protocols are becoming more important in recent years, especially protocols based on Time Division Multiple Access (TDMA). The main challenge for scheduling/reservation-based MAC protocols is generating a schedule that not only avoids internal collisions but can dynamically adapt to the conditions of the shared medium and mitigate potential sources of external interference. To avoid external interference, scheduling protocols should offer dynamic reallocation of slots to non-interfered sections of the available radio spectrum. Properly implemented dynamic scheduling protocol can be an important tool in enabling dynamic spectrum sharing and guaranteeing coexistence between various heterogeneous technologies operating in the same frequency bands, be it licensed or unlicensed spectrum bands.

1.2 Challenges

This dissertation tackles several challenges regarding the problems of underutilization and coexistence in the radio spectrum. These challenges are the consequence of inefficient radio spectrum usage and a constantly growing number of wireless technologies and devices. Research challenges addressed through this dissertation are as follows:

1. Dynamic allocation of radio spectrum.

2. Coexistence of heterogeneous technologies operating in the same frequency bands.

3. Distributed and dynamic scheduling for TDMA-based MAC protocols.

4. Applying dynamic spectrum sharing in satellite communications.

1.2.1 Dynamic allocation of radio spectrum

As mentioned in Section 1.1, fixed radio spectrum allocation leads to huge underutilization of electromagnetic spectrum and is not considered as a feasible model for the future of wireless communications. New models of spectrum assignment are needed, either by upgrading the existing licensing scheme or switching to more dynamic spectrum schemes. The actual implementation of DSA approaches is imposing several challenges. For example, the implementation of any dynamic spectrum sharing approach requires the complementary development of technologies and new regulatory standards. At the same time, prioritization of PUs and restrictions imposed on SUs may significantly degrade the performance and user experience of SU networks. There are many proposals for new spectrum sharing schemes, and a couple of them are already adopted and gaining more attention (see Section 1.1). If compared to fixed spectrum allocation, there are noticeable improvements. However, those solutions are still quite restrictive, and they possess a few shortcomings that need to be addressed (no guarantees that SUs will be able to deliver the required QoS is presenting the main downside). Achieved spectrum utilization with proposed schemes is still far from optimal and it is unlikely that they represent a long-term solution for the problem of radio spectrum underutilization.

1.2.2 Coexistence of heterogeneous technologies operating in the same frequency bands

While the coexistence of different technologies in licensed bands is achieved by division of the spectrum into frequency bands and exclusive band usage by different technologies, in unlicensed bands technologies are required to adhere to predefined standards and rules. However, even in this case, two or more technologies operating in close proximity typically lead to performance degradation. For example, when networks based on Wi-Fi and IoT technology operate in the same collision domain, mutual interference can lead to a throughput loss of 30% for Wi-Fi devices and 60% for IoT devices [13]. To further improve the performance of competing heterogeneous technologies in unlicensed bands, several coexistence schemes are implemented. These coexistence schemes are often technology-specific, only focusing on coexistence between pairs of mainstream wireless technologies. With an increasing number of technologies residing within unlicensed bands, it would require significant research and implementation efforts to adapt existing coexistence schemes or develop new ones. Congestion in unlicensed bands and the risk of performance degradation might lead to the reluctance of new technologies entering unlicensed bands.

1.2.3 Distributed and dynamic scheduling for TDMA-based MAC protocols

For TDMA-based MAC protocols, the main objective is to create internally collision-free schedules. There are different approaches to achieve this objective, with TDMA scheduling protocols mainly classified as centralized or distributed protocols. Distributed protocols are more scalable and adaptive solutions, thereby, they are the preferred protocols for networks with dynamically changing topology and traffic demands. Many distributed scheduling protocols, which generate internal collision-free schedules, are proposed, yet they often target single application domains. Besides, they do not take into account the potential occurrence of external interference, assuming that a TDMA-based network is operating in exclusive frequency bands. As such, these solutions are not suitable to support dynamic spectrum sharing. Other important issues that scheduling protocols should tackle are hidden and exposed node problems [14]. The hidden node problem occurs when two nodes, not in the same collision domain, transmit simultaneously to the same receiver, thus leading to collisions at the receiver node. On the other hand, exposed node problem occurs when two nodes are within the same collision domain and, for this reason, one or both nodes withdraw from the simultaneous transmission, even if their transmissions would not interfere at respective receivers. Whereas most of the existing scheduling protocols solve hidden node problem, exposed node problem is usually neglected. Hence, it is important to have a generic scheduling protocol that can work in various scenarios, solve the abovementioned issues, and be applicable as part of dynamic spectrum sharing frameworks.

1.2.4 Applying dynamic spectrum sharing in satellite communications

Similar to Section 1.2.1, the problem of spectrum underutilization due to fixed spectrum allocation exists in satellite communications. Due to its rising importance nowadays and different characteristics if compared to terrestrial wireless systems, the problem of spectrum sharing in satellite communications is investigated separately. It is expected that in order to accommodate new satellite constellations, dynamic spectrum sharing will replace exclusive spectrum assignment. Dynamic spectrum sharing schemes in satellite communications need to provide coexistence between geostationary (GSO) and non-geostationary (NGSO) satellite systems, as well as terrestrial networks, all operating in the same frequency bands. Moreover, these schemes need to be highly scalable and adaptable to network changes, taking into account the mobility of satellites and Ground Stations (GSs). There are only a few research proposals aiming to mitigate emerging problems when moving from static to dynamic spectrum assignment in satellite communications. However, none of these proposals investigated the most complex scenarios where GSO, NGSO, and terrestrial systems all operate and need to coexist in the same frequency bands.

1.3 Outline

This dissertation is composed of a number of publications that were realized within the scope of this PhD. The selected publications provide an integral and consistent overview of the work performed. The different research contributions are detailed in Section 1.4 and the complete list of publications that resulted from this work is presented in Section 1.5. Within this section we give an overview of the remainder of this dissertation and explain how the different chapters are linked together.

First, Chapter 2 gives a proposal for migrating from fixed spectrum licensing to dynamic spectrum management, which is accomplished following the centralized radio spectrum negotiation protocol. Instead of static frequency division, the radio spectrum is divided both in frequency and time domain, creating spectrum slices. Central entities (CEs) of Regulatory Committees dynamically assign the spectrum slices based on the demands of operators. Operators are charged per amount of assigned spectrum slices, driving them to minimize their spectrum usage. The majority of operators have the same priority, while only a few critical services may be assigned a higher priority. Each operator can request for allocation, deallocation, or move of assigned radio spectrum slices. Upon successful execution of negotiation protocol and (re)allocation or removal of the assigned radio spectrum, CEs update associated databases/repositories with relevant information on the spectrum usage. Further, in the case of local area networks that operate in unlicensed bands, a decentralized spectrum sharing model is proposed, which works in absence of CEs. Entry points (EPs) (access points (APs), base stations (BSs), or gateways (GWs)) of different wireless networks operating in the same geographical area negotiate for available radio spectrum, with each EP executing a decentralized spectrum negotiation protocol. EPs store information on the spectrum usage in their vicinity within a local database. In both centralized and decentralized spectrum management models, devices of each network are notified about acquired spectrum via supplementary MAC Spectrum Management Layer (SML), which is an extension to the standardized OSI model. Afterward, devices of the network may execute any MAC scheme within the acquired radio spectrum.

Chapter 3 describes the architecture of the end-to-end wireless communication system of the SCATTER team that participated in the DARPA SC2 competition. The main idea behind the design of the whole system was collaborative spectrum usage and coexistence with other unknown wireless systems that share the same spectrum. Participating wireless systems had the same priority to access the spectrum, while in few scenarios higher-priority incumbent networks were present. In Chapter 3, all the layers of the system are briefly introduced and their main functionalities described. Moreover, Chapter 3 describes the interactions between different layers of the system, allowing the whole system to work towards a mutual goal of optimal spectrum usage.

Detailed design of distributed scheduling functionality of MAC layer, employed by the SCATTER team, is presented in Chapter 4. The proposed Dynamic Distributed Multi-Channel TDMA (DDMC-TDMA) scheduling protocol is based on slot allocation and removal operations where nodes exchange corresponding control messages. The resulting Multi-Frequency Time Division Multiple Access (MF-TDMA) schedules are collision-free and dynamically adaptive as traffic demands, network topology, and/or external interference change over time. Every node maintains a scheduling table with information on spectrum usage in its vicinity. In Chapter 4, it is demonstrated how storing different slot types within scheduling tables helps to avoid hidden and exposed node problems. Further, it is shown that DDMC-TDMA offers a solution for problems of underutilization and coexistence between different networks, both in licensed and unlicensed bands.

Chapter 5 presents results from applying our previous work to the field of satellite communications. After investigating the benefits and disadvantages of distributed and centralized models for sharing the spectrum in satellite communications, we designed and implemented two distributed spectrum sharing techniques. With an increasing number of deployed satellites, the proposed techniques offer an efficient solution for solving the spectrum scarcity problem in the area of satellite communications. We investigated the coexistence of dual satellite systems (GSO and NGSO satellite systems) with terrestrial networks, all operating in the same satellite bands. Satellite systems are considered SUs, while terrestrial networks are PUs and need to be protected. Spectrum granularity is achieved by applying MF-TDMA both for downlink and uplink communication. The performance of the two dynamic spectrum sharing techniques is evaluated in uplink and downlink use cases in Ka-band (17.7–19.7 GHz for downlink and 27.5–29.5 GHz for uplink) and compared to fixed spectrum assignment model.

Finally, this dissertation is concluded in Chapter 6 together with an overview of future work.

For easier navigation through this dissertation, Table 1.1 shows the challenges that were highlighted in Section 1.2 and indicates which were targeted per chapter.

Table 1.1: An overview of the contributions per chapter in this dissertation.

	Ch.2	Ch.3	Ch.4	Ch.5
Dynamic allocation of radio spectrum	•	•	•	•
Coexistence of heterogeneous technologies operating in the same frequency bands	•	•	•	•
Distributed and dynamic scheduling for TDMA-based MAC protocols		•	•	
Applying dynamic spectrum sharing in satellite communications				•

1.4 Research contributions

In Section 1.2, the problems and challenges for optimal utilization of radio spectrum and dynamic spectrum sharing are formulated. They are tackled in the remainder of this PhD dissertation for which the outline is given in Section 1.3. To conclude, we present an elaborated list of the research contributions within this dissertation:

- Proposal of licensing model for dynamic spectrum management in licensed bands and dynamic spectrum sharing model for unlicensed bands (Ch. 2).
  - Extension of the conventional OSI model with the design of MAC SML.
  - The centralized licensing model for licensed bands based on negotiation between networks' EPs and Regulatory Committees.
  - Distributed spectrum sharing model for unlicensed bands based on information exchange between EPs of locally deployed networks.

  - Analysis of costs and benefits of proposed spectrum sharing models.

- Design and implementation of a wireless communication system that supports dynamic collaborative spectrum allocation and usage, offering coexistence with other wireless networks (Ch. 3)[1].
  - Design and implementation of full-blown MAC layer that supports mandatory MAC features, while providing dynamic spectrum sharing capabilities.
  - Testing the performance of the system in presence of other unknown wireless networks in an emulated wireless environment.

- Design and implementation of the DDMC-TDMA slot management protocol (Ch. 4).
  - Distributed slot allocation and removal functionalities.
  - Internal collision-free and traffic-adaptive scheduling.
  - Avoidance of both hidden and exposed node problems.
  - A robust solution against interference from external unknown networks.
  - Increased spectrum utilization by taking advantage of spatial spectrum reuse.
  - Implementation and evaluation of a proposed protocol for various wireless network topologies and scenarios in the ns-3 simulator.

- Development of dynamic spectrum sharing models for satellite communications (Ch. 5).
  - Proposal of two distributed and dynamic spectrum sharing techniques for increasing spectrum utilization and supporting coexistence in satellite-terrestrial systems:
    * The first demonstration of distributed dynamic spectrum sharing techniques for efficient spectrum sharing across multiple satellite and terrestrial systems (GSO, NGSO, and terrestrial networks);
    * Dynamic spectrum sharing technique based on collaboration protocol (CP);
    * Dynamic spectrum sharing technique based on decentralized spectrum sensing.
  - Comparison of centralized and distributed approaches for dynamic spectrum sharing in satellite communications.

[1]The author of this thesis was responsible for the implementation and testing of the MAC layer and dynamic spectrum sharing functionalities of the system

– Implementation and evaluation of proposed spectrum sharing techniques in the ns-3 simulator.
– Performance comparison of proposed spectrum sharing techniques against today's fixed spectrum assignment.

1.5 Publications

The research results obtained during this PhD research have been published in scientific journals and presented at a series of international conferences. The following list provides an overview of the publications during my PhD research.

1.5.1 Publications in international journals (listed in the Science Citation Index [2])

1. Felipe Augusto Pereira de Figueiredo, Xianjun Jiao, Wei Liu, Ruben Mennes, **Irfan Jabandžić**, and Ingrid Moerman. *A Spectrum Sharing Framework for Intelligent Next a Generation Wireless Networks.* Published in IEEE Access, Volume 6, p. 60704–60735, 2018.

2. Felipe Augusto Pereira de Figueiredo, Dragoslav Stojadinovic, Prasanthi Maddala, Ruben Mennes, **Irfan Jabandžić**, Xianjun Jiao, and Ingrid Moerman. *SCATTER PHY: An Open Source Physical Layer for the DARPA Spectrum Collaboration Challenge.* Published in MDPI Electronics, Volume 8, Issue 11, 2019.

3. Jan Bauwens, Bart Jooris, Spilios Giannoulis, **Irfan Jabandžić**, Ingrid Moerman, and Eli De Poorter. *Portability, Compatibility and Reuse of MAC Protocols across Different IoT Radio Platforms.* Published in Ad Hoc Networks, Volume 86, p. 144–153, 2019.

4. Ruben Mennes, Maxim Claeys, Felipe Augusto Pereira de Figueiredo, **Irfan Jabandžić**, Ingrid Moerman, and Steven Latré. *Deep Learning-based Spectrum Prediction Collision Avoidance for Hybrid Wireless Environments.* Published in IEEE Access, Volume 7, p. 45818–45830, 2019.

5. **Irfan Jabandžić**, Spilios Giannoulis, and Ingrid Moerman. *Enabling Generic Wireless Coexistence through Technology-Agnostic Dynamic Spectrum Access.* Published in Wireless Personal Communications, Volume 106, Issue 1, p. 151–177, 2019.

[2]The publications listed are recognized as 'A1 publications', according to the following definition used by Ghent University: A1 publications are articles listed in the Science Citation Index Expanded, the Social Science Citation Index or the Arts and Humanities Citation Index of the ISI Web of Science, restricted to contributions listed as article, review, letter, note or proceedings paper.

6. Felipe Augusto Pereira de Figueiredo, Ruben Mennes, **Irfan Jabandžić**, Xianjun Jiao, and Ingrid Moerman. *A Baseband Wireless Spectrum Hypervisor for Multiplexing Concurrent OFDM Signals.* Published in MDPI Sensors, Volume 20, Issue 4, 2020.

7. Miguel Camelo, Ruben Mennes, Adnan Shahid, Jakob Struye, Carlos Donato, **Irfan Jabandžić**, Spilios Giannoulis, Farouk Mahfoudhi, Prasanthi Maddala, Ivan Seskar, Ingrid Moerman, and Steven Latré. *An AI-Based Incumbent Protection System for Collaborative Intelligent Radio Networks.* Published in IEEE Wireless Communications, Volume 27, Issue 5, p. 16–23, 2020.

8. Ruben Mennes, Jakob Struye, Carlos Donato, Miguel Camelo, **Irfan Jabandžić**, Spilios Giannoulis, Ingrid Moerman, and Steven Latré. *Collaborative Flow Control in the DARPA Spectrum Collaboration Challenge.* Published in IEEE Transactions on Network and Service Management, Volume 17, Issue 4, p. 2024–2038, 2020.

9. **Irfan Jabandžić**, Spilios Giannoulis, Ruben Mennes, Felipe Augusto Pereira de Figueiredo, Maxim Claeys, and Ingrid Moerman. *A Dynamic Distributed Multi-Channel TDMA Slot Management Protocol for Ad Hoc Networks.* Published in IEEE Access, Volume 9, p. 61864-61886, 2021.

10. **Irfan Jabandžić**, Fadhil Firyaguna, Spilios Giannoulis, Adnan Shahid, Atri Mukhopadhyay, Marco Ruffini, and Ingrid Moerman. *The CODYSUN Approach: A Novel Distributed Paradigm for Dynamic Spectrum Sharing in Satellite Communications.* Submitted to MDPI Sensors.

1.5.2 Publications in international conferences (listed in the Science Citation Index [3])

1. Spilios Giannoulis, Carlos Donato, Ruben Mennes, Felipe Augusto Pereira de Figueiredo, **Irfan Jabandžić**, Yorick De Bock, Miguel Camelo, Jakob Struye, Prasanthi Maddala, Michael Mehari, Adnan Shahid, Dragoslav Stojadinovic, Maxime Claeys, Farouk Mahfoudhi, Wei Liu, Ivan Seskar, Steven Latré, and Ingrid Moerman. *Dynamic and Collaborative Spectrum Sharing: The SCATTER Approach* Published in proceedings of the 2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), p. 473–478, Newark, New Jersey, USA, 2019.

[3]The publications listed are recognized as ‘P1 publications’, according to the following definition used by Ghent University: P1 publications are proceedings listed in the Conference Proceedings Citation Index - Science or Conference Proceedings Citation Index - Social Science and Humanities of the ISI Web of Science, restricted to contributions listed as article, review, letter, note or proceedings paper, except for publications that are classified as A1.

2. Felipe Augusto Pereira de Figueiredo, Dragoslav Stojadinovic, Prasanthi Maddala, Ruben Mennes, **Irfan Jabandžić**, Xianjun Jiao, and Ingrid Moerman. *SCATTER PHY: A Physical Layer for the DARPA Spectrum Collaboration Challenge.* Published in proceedings of the 2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), p. 491-496, Newark, New Jersey, USA, 2019.

1.5.3 Publications in other international conferences

1. Felipe Augusto Pereira de Figueiredo, Xianjun Jiao, Wei Liu, **Irfan Jabandžić**, Spilios Giannoulis, and Ingrid Moerman. *A Framework for Intelligent Spectrum Sharing.* Published in proceedings of the 2018 IEEE 4th International Forum on Research and Technology for Society and Industry (RTSI), Palermo, Italy, 2018.

References

[1] Cisco. *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, Feb. 2019. Available from: https://s3.amazonaws.com/media.mediapost.com/uploads/CiscoForecast.pdf.

[2] D. Withers. *Radio Spectrum Management (2nd ed.)*. Institution of Engineering and Technology, London, 1999.

[3] S. Yin, D. Chen, Q. Zhang, M. Liu, and S. Li. *Mining Spectrum Usage Data: A Large-Scale Spectrum Measurement Study*. IEEE Trans. Mobile Comput., 11(6):1033–1046, 2012. doi:10.1109/TMC.2011.128.

[4] R. I. C. Chiang, G. B. Rowe, and K. W. Sowerby. *A Quantitative Analysis of Spectral Occupancy Measurements for Cognitive Radio*. In 2007 IEEE 65th Veh. Technol. Conf. - VTC2007-Spring, pages 3016–3020, 2007. doi:10.1109/VETECS.2007.618.

[5] M. Wellens, J. Wu, and P. Mahonen. *Evaluation of Spectrum Occupancy in Indoor and Outdoor Scenario in the Context of Cognitive Radio*. In 2007 2nd Int. Conf. Cogn. Radio Oriented Wirel. Netw. Commun., pages 420–427, 2007. doi:10.1109/CROWNCOM.2007.4549835.

[6] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood. *Chicago Spectrum Occupancy Measurements & Analysis and a Long-Term Studies Proposal*. In Proc. 1st Int. Workshop Technol. Policy Access. Spectr., TAPAS '06, New York, NY, USA, 2006. Association for Computing Machinery. doi:10.1145/1234388.1234389.

[7] Federal Communications Commission. *Spectrum Policy Task Force Report*, Nov. 2002. Available from: https://www.fcc.gov/document/spectrum-policy-task-force.

[8] Defense Advanced Research Project Agency (DARPA). *Spectrum Collaboration Challenge (SC2)*. Available from: https://archive.darpa.mil/sc2/.

[9] K.-H. Phung, B. Lemmens, M. Goossens, A. Nowe, L. Tran, and K. Steenhaut. *Schedule-based multi-channel communication in wireless sensor networks: A complete design and performance evaluation*. Ad Hoc Netw., 26:88 – 102, 2015. doi:https://doi.org/10.1016/j.adhoc.2014.11.008.

[10] J. Day and H. Zimmermann. *The OSI reference model*. Proc. IEEE, 71(12):1334–1340, 1983. doi:10.1109/PROC.1983.12775.

[11] W. Ye, J. Heidemann, and D. Estrin. *An energy-efficient MAC protocol for wireless sensor networks*. In Proc. 21st Annu. Joint Conf. IEEE Comput. Commun. Soc., volume 3, pages 1567–1576, June 2002. doi:10.1109/INFCOM.2002.1019408.

[12] I. Rhee, A. Warrier, M. Aia, J. Min, and M. L. Sichitiu. *Z-MAC: A Hybrid MAC for Wireless Sensor Networks*. IEEE/ACM Trans. Netw., 16(3):511–524, June 2008. doi:10.1109/TNET.2007.900704.

[13] R. A. Saeed, S. Khatun, B. Mohd Ali, and M. Khazani A. *Performance Analysis of Ultra-Wideband System in Presence of IEEE802.11a and UMT-S/WCDMA Frequency Bands*. In 2007 Int. Conf. Inf. Commun. Technol., pages 117–122, 2007. doi:10.1109/ICICT.2007.375355.

[14] A. Jayasuriya, S. Perreau, A. Dadej, and S. Gordon. *Hidden vs. exposed terminal problem in ad hoc networks*. In Proc. Aust. Telecommun. Netw. Appl. Conf., pages 52–59, Jan. 2004.

2

Enabling Generic Wireless Coexistence Through Technology-Agnostic Dynamic Spectrum Access

As previously mentioned, current fixed spectrum allocation is not a feasible approach for the future of wireless communications. If such an approach is maintained, there will come a time when no more operators and new wireless technologies will have access to the spectrum. Also, this is not a consequence of actual spectrum scarcity, but rather its inefficient utilization. This chapter presents a dynamic licensing framework to allow spatial and temporal reuse of the spectrum and solve the issue of spectrum underutilization. Moreover, another model is presented for enabling coexistence and avoid mutual interference of heterogeneous technologies in the unlicensed bands.

⋆ ⋆ ⋆

This chapter is adapted from:

I. Jabandžić, S. Giannoulis, and I. Moerman, *Enabling Generic Wireless Coexistence Through Technology-Agnostic Dynamic Spectrum Access*

Published in Wireless Personal Communications, Volume 106, Issue 1, p. 151–177, Mar. 2019.

Abstract Every year that passes, new standardized and proprietary wireless communication technologies are introduced in the market. They seek to find their place within the already highly congested spectrum. Regulatory bodies all around the globe are struggling to keep up with the continuously increasing demands for new bands by specific technologies, with some of them requiring by design an exclusive frequency band to operate efficiently. Even wireless bands offered for public or scientific usage like the Industrial, Scientific and Medical (ISM) bands are becoming the natural habitat of multiple wireless technologies that seek to use or "abuse" them to provide even more bandwidth to their offered applications. Wireless research teams targeting heterogeneous wireless communication coexistence are developing techniques for enabling one-to-one coexistence between various wireless technologies. Can such an exhaustive approach be the solution for N wireless technologies that wish to operate in the same band? We believe that a one-to-one approach is inefficient and cannot lead to a generic coexistence paradigm, applicable to every existing or new wireless communication technology that will arise in the future. Can another approach provide a more generic solution in terms of frequency reuse and coexistence compared to the one-dimensional frequency separation approach commonly used in commercial deployments today? Can such a generic approach provide a simple and easily adoptable coexistence model for existing technologies? In this chapter, we present a new generic medium sharing model that simply and efficiently solves the huge underutilization and coexistence problems observed today. Our approach is technology-agnostic and compatible with all existing wireless communication technologies and also has the capability to support emerging ones with minimum overhead.

2.1 Introduction

The increase in the number of devices that are dependent on wireless communication and the use of diverse wireless technologies will result in an explosive growth in traffic demands via wireless access services [1, 2]. In addition, with constant technological advances, new wireless technologies are being developed and are in demand for radio spectrum resources that will satisfy their needs. For these reasons, Regulatory Committees are struggling to keep up with the continuously increasing demand for additional available spectrum. The main problem is not the actual shortage of available radio spectrum, but inefficient use of the spectrum (underutilization). Issues with low spectrum utilization are best shown in Fig. 2.1 [3]. Spectrum utilization measurements are conducted in an urban area during the midday in the 0-6 GHz frequency band. In the spectrum below 3 GHz, we can see the utilization of roughly 30%, and the spectrum in the interval of 3-6 GHz is utilized only around 0.5%. Because radio spectrum has high economic value and spectrum usage efficiency is of huge importance, wasting this important resource must be avoided.

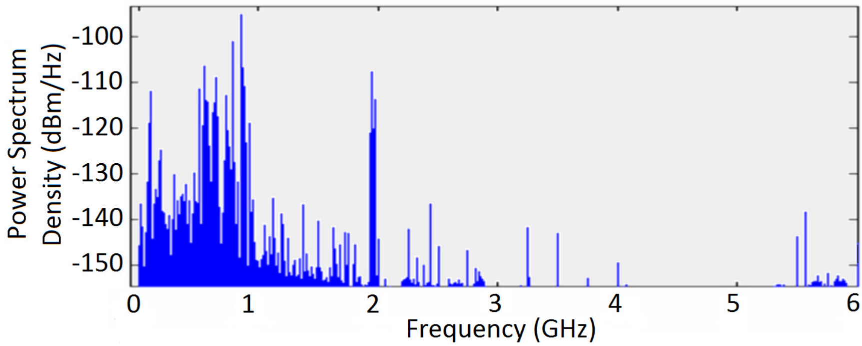


Figure 2.1: Measurement of spectrum utilization in downtown Berkeley [3].

Two main models of spectrum usage are dominant today (Table 2.1). Those are licensed access and unlicensed shared access. In the licensed access model, frequency bands are assigned to exclusive users where Regulatory Committees are guarantying interference-free usage of the assigned bands. On the other hand, the unlicensed shared access model is a license-free model. In unlicensed bands, wireless networks operating in spatially overlapping domains can experience significant mutual interference and performance degradation.

To reduce the mutual negative impact of competing heterogeneous technologies in unlicensed bands, there is a need for coexistence schemes. Many coexistence schemes are implemented until today, but with the constantly increasing number of technologies in the unlicensed radio spectrum, there is also an increased need for developing new and improved coexistence schemes, mainly between pairs of specific mainstream wireless technologies. Significant standardization effort/overhead is needed for two technologies to coexist. The number of coexistence solutions needed for N technologies is calculated based on Eq. 2.1. With the constant technological advancements and an increasing number of technologies, soon these requirements would be impossible to handle.

$$\binom{N}{k} = \frac{N!}{k!(N-k)!}, k = 2 \Longrightarrow \binom{N}{2} = \frac{N(N-1)}{2} \tag{2.1}$$

To solve the impending problems observed today in the wireless community worldwide, an alternative approach is needed. We propose an approach that provides a more generic solution in spectrum reusability and coexistence of heterogeneous wireless technologies. In our approach, instead of the typical frequency

Table 2.1: Spectrum usage models.

	Model	Description	Deployment	Technologies
TODAY	Licensed access	Exclusive assignment of frequency band	Single mobile operator	Traditional cellular networks (e.g. GSM, UMTS, LTE, 5G)
TODAY	Unlicensed shared access	License-free operation according to regional regulation	Many providers	Uncoordinated operation of many technologies (short range: IEEE802.11, IEEE802.15.4, Bluetooth, MulteFire,..., long range: LoRa, SigFox, Dash, IEEE802.11ah...)
EMERGING	Licensed spectrum sharing	Frequency band assigned to multiple providers based on sharing rules (location, spectrum)	Authorized provider (micro-operators) + mobile operators	Private LTE/5G networks (e.g. CBRS) Directive from FCC (US) : SAS Directive from RSC (EC) : LSA
EMERGING	Licensed assisted access (LAA)	Use of unlicensed band(s) in addition to licensed band to boost performance	Mobile operator + unlicensed network providers	Coexistence of cellular + unlicensed technologies (e.g. LTE-LAA)
EMERGING	Sharing in application-specific bands	Frequency band assigned to specific applications	Multiple providers	DSRC 5.9 GHz for ITS: IEEE802.11p + LTE V2X
EMERGING	Dynamic spectrum access (DSA)	Unlicensed users use temporarily unused licensed band(s)	Authorized provider + unlicensed network providers	IEEE 802.11, IEEE 802.15.4, IEEE 802.22, WiMAX

division of the radio spectrum, we divide the radio spectrum both in frequency and time. Those basic units of the radio spectrum, referred to as spectrum slices from now on, are dynamically allocated to the users based on the users' demands. The proposed approach would increase the utilization of the spectrum and reduce interference in the license-free parts of the spectrum. In addition, it offers compatibility with all existing wireless technologies and can effortlessly support any emerging wireless technology with minimum overhead. This would allow multiple heterogeneous wireless networks to coexist in the radio spectrum with no or minimal knowledge of other present technologies. The proposed models replace the one-dimensional frequency separation approach that is most commonly used today in commercial applications as well as all Medium Access Control (MAC)-related efforts to provide coexistence through MAC techniques like Listen Before Talk (LBT).

In the remainder of this chapter, we first give in Section 2.2 a short review of the related work until today. In Section 2.3, we present the proposed general architecture. Section 2.4 includes a description of system functionality with a description of protocols used for achieving reliable spectrum sharing. Section 2.4 is divided into two proposals for centralized and decentralized spectrum sharing models. Analysis of costs and benefits of the proposed spectrum sharing models is presented in Section 2.5. Finally, this chapter is concluded in Section 2.6.

2.2 Related work

A lot of research has been done on implementing coexistence schemes for heterogeneous wireless technologies in the unlicensed radio spectrum. Researchers are especially focused on the coexistence of mainstream technologies, such as LTE and Wi-Fi. Two main approaches for improving LTE technology are LTE Licensed Assisted Access (LTE-LAA) [4, 5] and LTE-Unlicensed (LTE-U) [6]. LTE-LAA and LTE-U propose the use of unlicensed bands in addition to licensed bands to boost the performance of the wireless LTE networks while trying to support coexistence with other technologies in unlicensed spectrum, especially focusing on coexistence with Wi-Fi technology.

Because of increasing demands for additional radio spectrum resources, new spectrum usage models are emerging lately, as is shown in Table 2.1. A lot of significance is given to concepts of using radio spectrum in other ways than statically allocated spectrum [7–11]. The main idea is in the dynamic allocation of spectrum to radio services instead of a fixed allocation as it takes place today. Most work is being conducted in the fields of Licensed Shared Access (LSA) and Dynamic Spectrum Access (DSA). In LSA, frequency bands (shareable spectrum bands) are assigned to multiple users based on specific sharing rules. LSA represents a regulatory approach to allow any incumbent to share its licensed spectrum with prospective users following a sharing framework, predefined by the National Regulatory Authority [12, 13]. DSA is the concept of allowing Secondary Users (SUs) (unlicensed devices) to use spectrum that is unused by Primary Users (PUs) (licensed devices). SUs utilize parts of the spectrum that belong to PUs based on the agreements and constraints imposed by PUs so that PUs will not be affected in any negative way. The main goal of DSA is to create flexibility in the spectrum usage so that unlicensed users could have access to the parts of radio spectrum temporarily unused by licensed users and this technique is expected to increase spectrum utilization [14]. The DSA aims to manage the radio spectrum by sharing it among different wireless networks over space and time to increase overall spectrum efficiency [15, 16]. LSA and DSA approaches are imposing several research challenges. These challenges are emerging because of the broad range of radio spectrum and because of the diversity of Quality of Service (QoS) requirements of heterogeneous wireless networks. In addition, increased valorization and enhanced control of spectrum usage by Regulatory Committees need to be established. At the same time, these approaches do not guarantee to satisfy QoS requirements of SUs, because PUs are prioritized in these concepts.

Architecture in DSA networks for sharing the spectrum can be centralized and distributed. Centralized spectrum sharing is discussed in [17, 18], while distributed spectrum sharing is discussed in [19–22]. In [23], an entity named Global Spectrum Controller carries out tasks related to coordinated spectrum sharing. It stores information on spectrum resource availability and ensures that the shared spectrum system operates in conformance with the licensing regime. Spectrum Manager in [24] is responsible for spectrum fragmentation and allocation of fragments to base stations (BSs). Fragments are extracted from the subcarrier grid; the shared spectrum is dynamically fragmented and allocated to the operators for ”in-band” sharing. Both [23] and [24] are based on subcarriers as the minimum fragments of spectrum used for dynamic allocation; we do not consider that to be an adequate degree of spectrum fragmentation for our proposal. In addition, spectrum sharing in [23] and [24] is solely focused on network operators’ level, while our approach is more flexible and is enabling any type of deployed wireless network to be part of the architecture and enter the spectrum, without prioritizing larger networks. Dynamic Spectrum Access Protocol (DSAP) is a centralized protocol that enables lease-based dynamic spectrum access through a coordinating central entity and allows efficient resource-sharing and utilization in wireless environments. DSAP is designed to provide spectrum leases to wireless devices in a limited geographic region at small timescales and focuses solely on unlicensed bands, thus not solving problems of underutilization in licensed bands, nor addressing spectrum management on a global scale. DSAP uses a database (RadioMap) for storing important radio spectrum information and provides dynamic allocation of the radio spectrum to network nodes [25], which is similar to our proposal, however, DSAP focuses only on specific spectrum bands. Dynamic intelligent management of spectrum for ubiquitous mobile-access network (DIMSUMnet) is a centralized mechanism based on spectrum brokering that manages large chunks of the spectrum and assigns its portions to individual domains or users. DIMSUMnet implements statistically multiplexed coordinated access to the spectrum in the Coordinated Access Band (CAB). The CAB represents a contiguous chunk of spectrum reserved by regulating authorities and it is used for controlled dynamic access [26]. DIMSUMnet uses a centralized, regional network level brokering mechanism that aims to significantly improve spectrum utilization while reducing the complexity and the agility requirements of the deployed system [27]. DIMSUMnet is only focused on dynamic spectrum access in CAB bands that are part of the radio spectrum and reserved in advance for purposes of dynamic spectrum access. In the DIMSUMnet proposal, clients decide which application service to select only once they have received the reserved spectrum information. In the second part of the proposal, clients can participate in a leasing process and request spectrum access themselves. This approach induces high complexity in the leasing process, in the case of large-scale deployments.

Besides the abovementioned centralized DSA networks, using a database to store information about available spectrum in white-space (underutilized TV bands) for DSA is proposed by many researchers [28, 29]. Temporally unused parts of the spectrum are referred to as spectrum holes or white space. The existence of an accurate spectrum database to provide spectrum maps at different locations is proposed in [30]. The proposed architecture is using common control channel (CCC) for nodes to connect to the database and exchange control messages. The concept of using a repository for storing spectrum-related information can also be found in LSA [31].

It has been noticed that the current division of spectrum resources only in the frequency domain is not optimal and that much efficient usage of the spectrum would be achieved if the spectrum is divided into more axis than one. Therefore, dividing spectrum in non-overlapping orthogonal channels, where channel division can follow the format of Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), Code Division Multiple Access (CDMA), or a combination of them, has been proposed [32–34]. In [35] radio access model is reduced from frequency division based sharing among operators to a Time Division Duplex (TDD) system and each operator is allocated a certain number of slots in the superframe. The Dynamic Radio for IP Services in Vehicular Environments (DRiVE) project [36] is another approach for the problem of dynamic spectrum allocation that uses CCC and is focused on heterogeneous networks. It proposes the use of a contiguous radio spectrum instead of fixed radio bands. The extension of the DRiVE project is the Spectrum Efficient Uni- and Multicast Over Dynamic Radio Networks in Vehicular Environments (OverDRiVE) project [37] that enhances the proposed contiguous division of spectrum to fragmented radio spectrum. In our proposal, we adopt and enhance the approach presented in the OverDRiVE project, while other proposals are not offering enough spectrum granularity to consider them suitable as a starting point.

In our proposed decentralized model, long-distance wireless technology is supposed to fulfill the role of the control channel. In addition, in the centralized model, the connection of the network's entry points (access points (APs), BSs, or gateways (GWs)) to Central entity (CE) can be a wireless connection, if no wired infrastructure is available. For successful exchange of the control packets using a wireless medium, part of the spectrum (control time-frequency slice) is used. Many spectrum sharing solutions assume the abovementioned CCC for the exchange of control messages [25, 38, 39]. A similar approach is proposed in [40], where a small amount of spectrum is allocated for control messages and called Common Spectrum Coordination Channel. The abovementioned papers are reserving whole frequency bands for control messages, while in our proposal we only need a spectrum slice, so just a part of a frequency band, for the successful exchange of control messages, while the rest can be used for leasing purposes.

The research community has recognized that for dynamic allocation of the spectrum, the current Open System Interconnection (OSI) stack model is often not sufficient. Part of our proposal is an enhanced OSI stack model with an additional sublayer for dynamic spectrum allocation. Research on the need for additional layers, in the case of dynamic spectrum allocation, has been carried out earlier and Dynamic Bandwidth Allocation (DBA) sublayer has been proposed [9], while in [41] both legacy stack and modified stack are used. Based on the specifics of every proposal, every proposed model of OSI stack offers different stack modifications and functionalities. We claim that our enhanced OSI stack model offers improved functionalities with minimum modifications of the standardized OSI stack model.

2.3 Architecture of proposed radio spectrum sharing models

In this section, the architecture of the proposed radio spectrum sharing models is presented, together with the sub-modules of the architecture and their functionality and role towards the defined goal of dynamic spectrum sharing.

2.3.1 Division of radio spectrum in frequency-time slices

Licensed and unlicensed radio spectrum is currently divided into frequency bands/channels that are the fundamental units of spectrum usage. These channels are generated as non-overlapping and usually, some guard space (Minimum Frequency Separation (MFS)) is reserved in between them [42]. Licensed spectrum bands are assigned to PUs (license holders), while technologies in unlicensed spectrum try to coexist with other present technologies. To achieve more granularity out of the available spectrum we adopt and enhance the approach proposed in the OverDRiVE project [37] where the radio spectrum is divided into fragments. Instead of solely frequency division of radio spectrum, we use the typical TDMA approach and divide frequency bands in time, so the fundamental units of radio spectrum in our proposal are frequency-time slices/slots. Typical frequency division of the radio spectrum and more granular division of radio spectrum into frequency-time slices are shown in Fig. 2.2. Each frequency-time slice can be used to facilitate an unknown communication technology for the duration of the slice. This approach is completely agnostic and there is no need for any compatibility between technologies. During the slice duration, heterogeneous wireless technologies are allowed to use any medium access protocol without experiencing any negative effects caused by other technologies. In addition, any type of channel coding/modulation scheme and Physical layer (PHY) signaling is acceptable.

The proposed fragmentation of radio spectrum allows Regulatory Committees to assign to any heterogeneous wireless network an appropriate number of slices depending on the network's requirements. It also helps to mitigate underutilization and interference issues that are occurring today. The time duration of slices can be determined based on theoretical knowledge or experimental results for present technologies and updated when new wireless technologies emerge in the market.

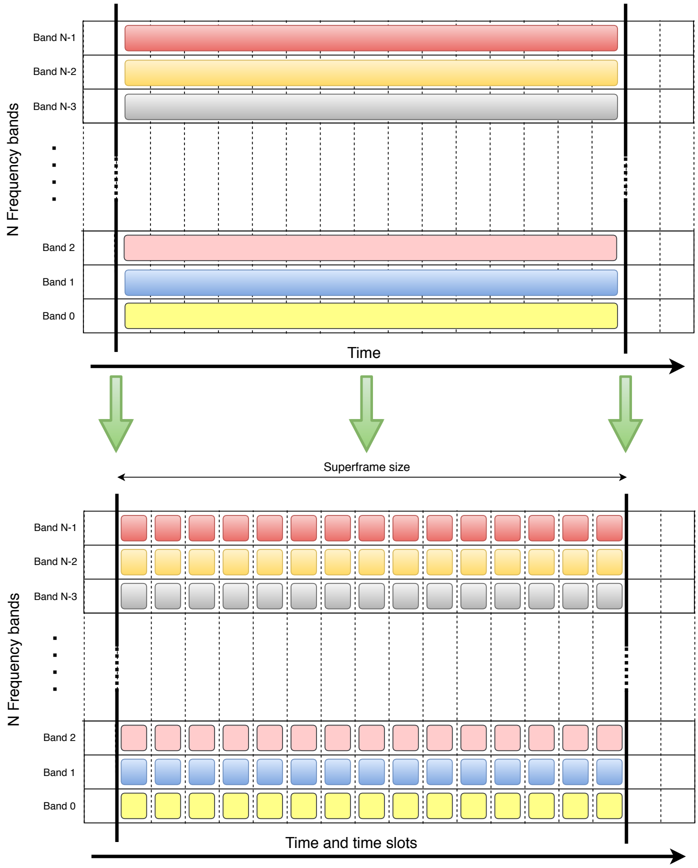


Figure 2.2: Transformation from frequency division to the frequency-time division of radio spectrum.

Additionally, the time duration of slices can vary from frequency to frequency band, based on anticipated technologies and medium access protocols employed in each frequency band.

2.3.2 MAC Spectrum Management Layer

As a part of the proposed spectrum sharing architecture, we propose an upgrade to the current OSI model, in the form of an additional layer that is responsible for providing time and frequency accessibility of radio spectrum to the link layer. In the proposed OSI model, MAC Spectrum Management Layer exists in parallel with the legacy MAC layer, as shown in Fig. 2.3. In the remainder of the chapter, we shall refer to this layer only as Spectrum Management Layer (SML).

For every managed network, SML is going to keep information about allocated time-frequency slices in local storage space. The information needed is central frequency, bandwidth, slice start and end time as well as potential periodicity of the slice. SML of every node is communicating with the network's entry point (EP) to obtain information about the allocated/deallocated spectrum slices for that network. Further, the MAC layer is informed about slice allocations by the SML. If the managed network is assigned with slices that are going to be shared with other managed networks, this information is also provided to SML. Now, the SML can notify the MAC layer of what additional features should be enabled to achieve coexistence or cooperation with other managed networks. The proposed OSI model improves conventional layer architecture and it is a core of our dynamic spectrum sharing (DSS) proposal.

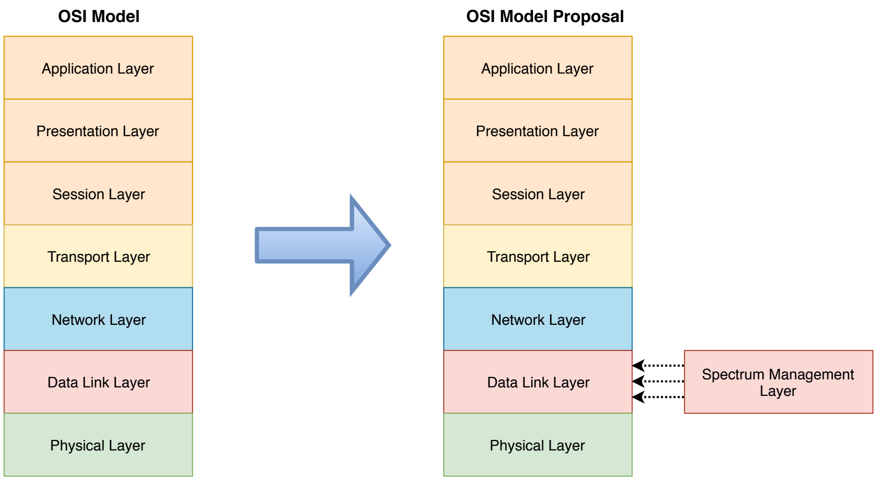


Figure 2.3: Proposed OSI model with MAC Spectrum Management Layer.

Legacy MAC layer for every network should only exploit transmission opportunities in slices that are approved for that network. It also needs to take into account the duration of the slice, so that transmission attempts will not go out of slice boundaries. By carefully designed interface of SML, this can be achieved with the minimal modifications of existing legacy MAC layers. For this to be feasible, SML's interface needs to present requirements in a transparent way to all existing legacy MAC layers.

SML is imposing constraints on the legacy MAC layer in form of Application Programming Interface (API) primitives as shown in Fig. 2.4. SML is sending Block_at and Unblock_at calls in a way that is transparent to any existing legacy MAC layer. Block and unblock calls are sent with at least millisecond precision, taking into account the overall accuracy of the chosen synchronization mecha-

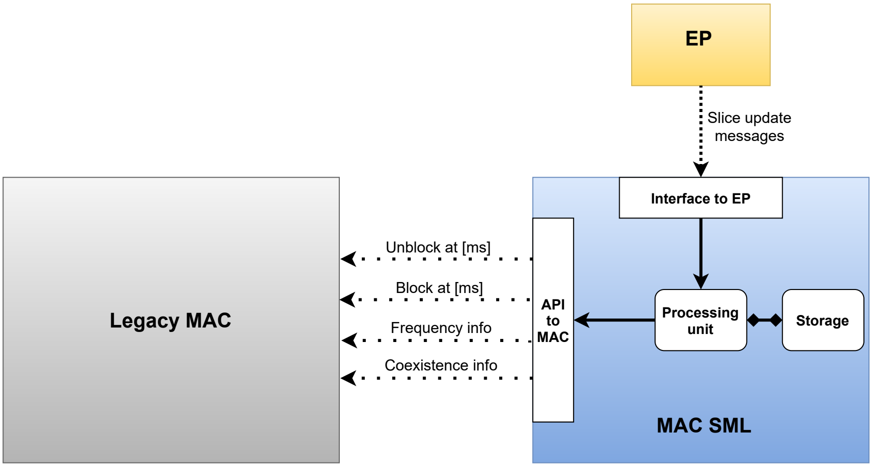


Figure 2.4: Application programming interface from SML to MAC layer.

nism across the architecture. In addition to blocking/unblocking calls, another API primitive interfacing to legacy MACs is information about the central frequency and allowed bandwidth. In case that spectrum slice is shared with another managed network, another API primitive is used to inform the legacy MAC layer about the type of technology that is already present in the slice, so that legacy MAC can enable the existing coexistence scheme during that slice. Slices, where coexistence with other managed networks is needed, would be proposed only as a last resort measure (overcrowded unlicensed bands for example) and only when CE has information that legacy MAC contains the implementation of coexistence schemes that are prerequisite for successful coexistence in the corresponding slice. Required updates of legacy MACs to comply with SML's API interface are minimal. Any legacy MAC only needs to exploit received calls to schedule execution of already implemented MAC algorithms within reserved time slices. Additional new functionality of legacy MACs is the support for runtime enabling of coexistence schemes if slices are shared. Block_at and unblock_at calls are sent adequate time in advance, so that legacy MAC has time to process received calls, do the scheduling, and perform MAC's algorithm inside the bounds of an active spectrum slice. Based on the technology requirements sent from the network's EP to CE, the duration of the slice, bounded by blocking/unblocking calls, is sufficient for legacy MAC protocol to perform typical MAC operations without the risks of exceeding the imposed time limits (see Section 2.3.5 for more information on the minimum duration of spectrum slices).

2.3.3 Central entities of Regulatory Committees

Regulatory Committees today mostly play the role of the rule-maker but most of the time they lack the mechanisms to enforce spectrum access based on the defined spectrum slices. In our proposal, Regulatory Committees take up the active role of spectrum-sharing policymakers in runtime. All radio spectrum management (radio spectrum slices allocation and deallocation) is controlled by Regulatory Committees' CEs that are responsible for controlling and maximizing spectrum utilization. These CEs are fixed infrastructure components that are connected through wireless/wired connections to all managed deployed networks and are responsible for the direct control of the allocation of available radio spectrum resources. CEs represent the main part of the regulatory system and are responsible for exchanging the necessary control messages with the deployed managed networks whose spectrum access will be regulated. Loss of control messages should be minimal, therefore a wired connection to the deployed networks is the preferred option. However, if the wireless connection is the selected or only available option, part of the radio spectrum for communication of CEs to other parts of the system should be allocated (control radio slices), only to be used for that purpose. As today the spectrum is regulated on a national, regional, and worldwide level, CEs should equally negotiate usage of radio spectrum slices on those levels. The overall behavior of CEs should enable enforcement of national and regional regulatory spectrum allocation policies while retaining flexibility on spectrum allocation. CEs should mainly control spectrum usage on a national level, but they should also be coordinated at higher levels so that there is no interference and usage of the same slices in border regions of neighboring states. National CEs should be connected to regional CEs and share information about spectrum usage via those regional CEs. The focus is especially on sharing information about spectrum usage in border regions of the countries. The main task of the national CEs is to dynamically provide spectrum slices to different wireless networks/providers/vendors while mitigating the interference and enabling efficient utilization and reuse of the spectrum. In the remainder of this work, heterogeneous wireless networks managed by CEs shall be called managed networks.

Every national CE is keeping a location-time-frequency database (repository) and should update its database with the data received from national CEs of neighboring countries and regions. CEs also contain information on the availability of spectrum resources (time-frequency slices), in accordance with each country's respective regulatory framework. Further, they keep information about: geographical features of the heterogeneous managed networks (spatial location represented with latitude, longitude, and altitude), frequency-time slices assigned to different networks, and spectrum regulations and policies. Once the duration of a slice reservation expires, CE's database is updated and the appropriate request to free the slice is sent to the associated network's EP. In the database, CE stores specific information about heterogeneous managed networks that are utilizing the spectrum, as wireless technology used, medium access protocol used, modulation schemes, etc. This way, CE can assign the same slices to different managed networks in case that spectrum is fully utilized and CE has information that two or multiple managed networks can successfully cooperate or coexist. As new wireless technologies emerge constantly, before any upcoming technology enters the market, it should provide Regulatory Committees and CEs with enough information on its capability of coexistence or cooperation with existing technologies. Provided information enables CEs to reuse parts of the spectrum if deemed necessary and under certain conditions. Every new technology that enters the market and provides the necessary information to CEs is then assigned a unique technology Identification (ID) number, and the managed networks based on that technology can participate in the future in the dynamic radio spectrum allocation process.

CE is responsible for communicating with networks' EPs for purposes of allocation/deallocation of frequency-time slices. It calculates slices to allocate/deallocate based on the proposed algorithm, described in Section 2.4. As an input, CE takes managed networks' EPs requests, extracts necessary data from the requests, and accesses the available information in the CE's database. CE is also responsible for updating its database when an allocation/deallocation procedure is finalized and it should periodically exchange radio spectrum information with other CEs in its vicinity. Therefore, CEs become the active instrument that allows regulatory authorities to monitor managed networks and to ensure that managed networks operate in conformance with the existing regulatory framework.

2.3.4 Entry points of managed networks

The link between CEs and managed networks is the EP (AP, BS, or GW) of each managed network. The EP of the managed network is responsible for the introduction of a new managed network to the radio spectrum and maintenance of the network. It keeps track of the network requirements and forwards these requirements to the related CE for dynamic adjustment of the allocated spectrum for this network. In addition, EP is responsible for handling any request it might receive from the CE to keep the network in compliance with the regulatory framework. In the case of local area heterogeneous networks, EPs communicate directly to other EPs in the area and keep a local database with information on spectrum usage in that area. On the global level, any change in the allowed spectrum access of the network is first forwarded to the EP by the CE, and the EP then forwards the information to all the nodes in the network. For our DSS proposals, we assumed the time synchronization of all network components. If the nodes in the network do not have any means for global time synchronization, EP is obliged to synchronize all the nodes in the network to global time. Information about global time can be obtained using any standardized method (Global Positioning System (GPS), Network Time Protocol (NTP), Precision Time Protocol (PTP), etc.) as long as the slices' minimum duration is not in the range of the synchronization accuracy of the employed method.

2.3.5 Minimum duration of spectrum slices

EP of every managed network has knowledge of wireless technology and MAC protocols used by the devices in the network. All existing legacy MAC protocols have specific timing requirements for the successful execution of the protocol. Timing requirements are most important for procedures of packet transmission and reception. To avoid being assigned with spectrum slices of insufficient time duration, where basic MAC operations will not fit, an important parameter to provide to the CE when requesting new spectrum slices is the minimum acceptable duration of the slice; we shall refer to it as slice base duration. The requested slice base duration for any MAC protocol should not exceed the predefined national value, otherwise, CE will discard the request and notify EP that the allocation request failed, with provided reason for failure.

As an example, in a typical Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) MAC protocol, feasible slice base duration can be calculated based on the time needed to complete one packet transmission. In the calculation, the time duration of typical basic procedures or critical defined times of Carrier Sense Multiple Access (CSMA) are included like: duration of Interframe Spaces (Short Interframe Space (SIFS) and DCF Interframe Space (DIFS)), time needed for Clear Channel Assessment (CCA), transmission period for maximum valid data packet size, and acknowledgment transmission period. In the case of CSMA/CA protocols containing Request to Send (RTS) and Clear to Send (CTS) messages, the time for transmission of these messages is included in the calculation as well. In the case of a typical TDMA MAC protocol, the slice base duration should be at least equal to the duration of TDMA superframe or a multiple of TDMA superframe duration, so that one or multiple superframes can fit into one allocated spectrum slice.

It is worth mentioning that while networks may request slices of very small time duration, it is expected that the majority of networks will request slices big enough to accommodate at least a couple of transmissions. Further, networks are allowed to request allocation of full channels, if they are based on the technologies utilizing whole channels in order to operate, e.g. stream-based technologies.

2.4 Proposed radio spectrum sharing models and protocols

2.4.1 Centralized model proposal

In the proposed centralized model for spectrum sharing in licensed bands, Regulatory Committees' CEs handle requests and approve the usage of frequency-time slices. CEs have full information on spectrum utilization stored in their database/repository following the history of approved requests of spectrum usage in their area of control. To communicate with CEs and exchange messages for allocation and deallocation of the radio spectrum, every EP must have a wired/wireless connection to the CE. In the case of employing a wireless connection between EPs and CEs, part of the radio spectrum should be reserved (control radio slices) for the exchange of control messages. In addition to communication between EPs and CEs, the control radio slices are also used for communication between different CEs forming a control channel between the regulatory components of the architecture. Managed networks are not permitted to use the control radio slices for user data in order to avoid interference and the negative impact of high data traffic on the reliability of the control traffic. The proposed architecture for a centralized dynamic sharing model is presented in Fig. 2.5. The centralized spectrum sharing model is intended for spectrum sharing in licensed bands, however, its concepts can also be applied for the spectrum sharing in unlicensed bands.

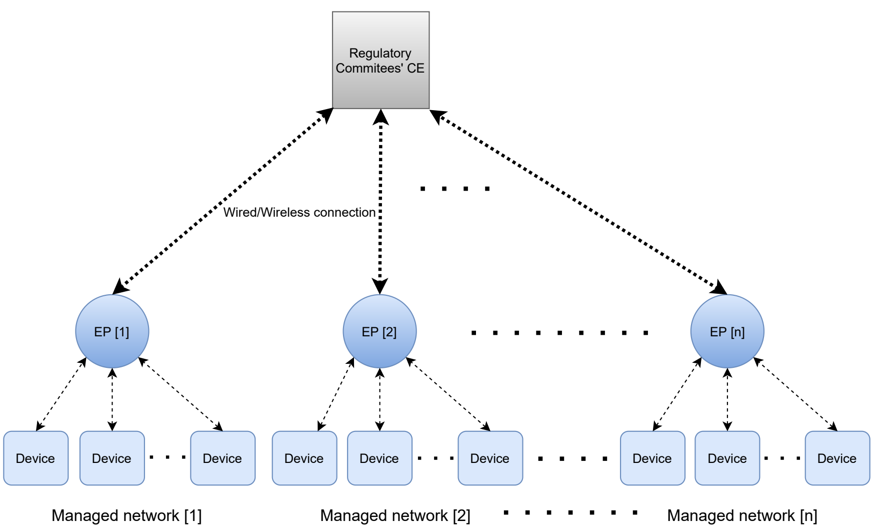


Figure 2.5: Centralized model architecture.

For efficient utilization of the spectrum, CEs need to know the location and spatial distribution of the managed networks that require to access the medium. For that to be possible, localization is needed for EPs, either through using GPS or other standardized methods. The location of the EP can also be fixed, set during deployment if the EP cannot move. To avoid collision with other managed networks in the area where multiple networks reside, time synchronization for EPs is required through additional hardware like GPS, by using established network-based techniques like NTP, or using any other acceptable method that can provide at least millisecond-level accuracy. Besides time synchronization at the EPs level, to follow the established rules of transmitting inside the assigned slices, every client node also needs to be synchronized to the global time. For the time synchronization of the network nodes, the same approach can be used as for the EPs, or the EPs can adopt the role of synchronization master of their internal network nodes based on any synchronization protocol. To avoid extreme costs of user equipment devices, GPS can be employed at the EP level for worldwide synchronization of all EPs, while network protocols like PTP or NTP can be used to synchronize the client devices to the local EP.

Slice negotiation and allocation for all heterogeneous managed networks are initiated by the network's EP based on specific requirements (data rate, user demands, QoS, etc.). EPs are performing continuous monitoring of the requested QoS. When QoS requirements change in a way that the currently assigned spectrum is insufficient or part of the spectrum becomes unused, EPs initiate appropriate action. Spectrum usage is paid per number of allocated slices in licensed bands, which as a result reduces unutilized radio spectrum, as no spectrum user is willing to pay for more spectrum than needed to satisfy its demands. To reduce the costs of using the spectrum, if the bandwidth requirements are reduced, EPs can send remove requests and deallocate part of the used spectrum. The main purpose of the centralized DSS model for licensed spectrum is efficient utilization of spectrum but it also leads to reduced costs for the users of the spectrum. All this is achieved while at the same time the authorized networks (spectrum slices holders) are being served the necessary spectrum in order to be able to provide a certain QoS to their users.

When a new managed network wants to enter the spectrum, its EP follows the specified procedure and exchanges request and response messages with the associated CE. The proposed procedures are explained later in this section. Based on their requests, managed networks are allowed access to an adequate number of frequency-time slices. Once the network's EP acquires information of radio spectrum available for it to use, EP communicates to all network nodes over the proposed SML of the OSI model, forwarding all the necessary information. Every change in the usage of radio spectrum initiated by EP or CE is forwarded to the SML of every node and being enforced on all nodes. EPs of managed networks are responsible for carrying out the successful transfer of slices information messages to internal nodes, using any approach that is preferred by the managed network's specific technology characteristics. It is assumed that every EP of a managed network can have direct or multi-hop communication with all internal nodes of their network. After the reception of slices information messages, network nodes can operate using the spectrum slices that are allocated for that managed network, regardless of the existence of other managed networks in its vicinity. An example of spectrum sharing between heterogeneous managed networks in the same vicinity is shown in Fig. 2.6.

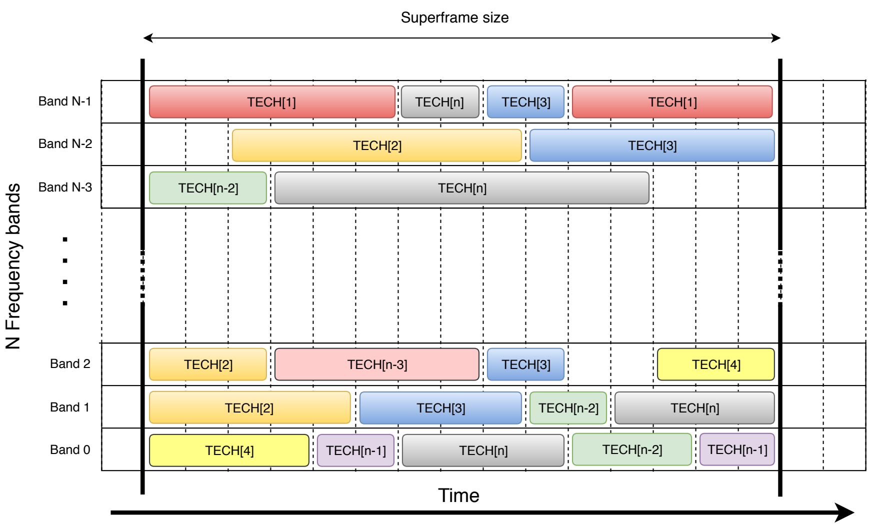


Figure 2.6: Example of centralized model spectrum allocation.

CEs can be set up on the national, regional, continental, or global level. All the lower CEs (national level) are synchronized and aligned in a global network of CEs controlled by higher-level CEs (regional, continental level). This enables the exchange of information between CEs, avoidance of the spectrum interference in the border areas that are controlled by national entities, and better spectrum utilization. Any prioritization or precedence that needs to be given to specific networks, like military or public safety networks, should be implemented within the CEs and embedded in the decision-making process of accepting or denying spectrum requests from all managed networks. The network priority parameter should be taken into account in the spectrum allocation procedure to make sure that the higher priority networks are never starved of spectrum resources.

2.4.2 Radio spectrum slices negotiation protocol (centralized model)

When a new managed network enters the spectrum, the first step is to allocate part of the spectrum based on QoS requirements. If the requirements for the spectrum increase over time, the network's EP can request more slices of the spectrum from the CE following the same procedure as in its initial entry phase. In addition, if there is a reduction in the required spectrum, EP can ask CE to deallocate some of the previously allocated slices. If there is a need to move time-frequency slices, which are used by the network, from one frequency band to another, the slice move procedure can be initiated. CE initiates procedures only when it is required to move or remove a network slice in order to decrease interference or because a more prioritized request has occurred in the vicinity of the managed network and there is no free spectrum to satisfy the prioritized request. CE initiates procedures to keep the system stable, reduce interference, and increase the number of unused spectrum slices that can later be assigned to other upcoming technologies. Four main primitives comprise the protocol for centralized spectrum sharing and are described in more detail in the following section:

1. Slice allocation
2. Slice deallocation
3. Slice move
4. Housekeeping (heartbeat)

There are four message types used for the successful negotiation of slices between network EPs and national CEs:

1. Request messages
2. Response messages
3. Slice update messages
4. Heartbeat messages

Request messages are sent from network EP to CEs. They are used whenever a managed network needs more spectrum slices to satisfy user demands or some of the slices are not needed, so deallocating them would reduce costs induced by Regulatory Committees. Allocation, deallocation, and move protocols are initiated by corresponding Request messages.

Response messages from CEs to EPs are sent in case of both possible outcomes: successful execution or failure of protocol primitive. Response failure messages are classified on response failure type A and type B messages. Response failure type A messages are response messages from CE to the EP in case that request messages are not satisfying protocol requirements, for example, request message not including all information needed for CE to do the computation and generate a proper response. Response failure type B messages are exchanged when CE calculates that request from the EP is not feasible (requesting too much spectrum share, not enough free spectrum slices available in the database, network technology not able to coexistent with any other technology already using the spectrum, etc.). Response success messages are sent to EPs from CEs if negotiation is successful and the network is approved to allocate additional spectrum, remove part of the currently allocated spectrum or reallocate its current slices to other parts of the radio spectrum.

Slice update messages are used to update the SML of every node in the managed network with newly allocated/deallocated slices.

Heartbeat messages are part of housekeeping protocol and are responsible for keeping the system stable and error-free.

1. **Slice allocation procedure:** In the Request allocate message, the network's EP needs to provide its unique technology ID number to the CE, for the CE to fetch additional information about the specific technology features from its database, like MAC protocols supported by the technology, possibility of coexistence with other technologies, etc. EP also needs to include spectrum requirements based on its user demands and QoS requirements. Furthermore, the request allocate message should contain: required bandwidth, duty cycle, slice base duration, expected usage duration of approved slices, and location information of the network. On the CE side, the location information can be used to build a spatial map of wireless coverage in order to proceed with the most appropriate spectrum slice allocation, minimizing interference to co-locating networks. CE gathers necessary information from CE's database (free spectrum slices, technology information based on received unique ID number, etc.) and calculates the minimum number of frequency-time slices that can satisfy requirements received in the Request allocate message. CE then updates the local database and notifies EP about calculation results via the Response allocate message. Response allocate message carries information about spectrum slices that are now allocated to the managed network. In the case that the newly deployed managed network should coexist with other networks in one or more of assigned slices, additional information about coexisting networks' technologies is also provided. Slice allocation protocol is presented in Fig. 2.7.

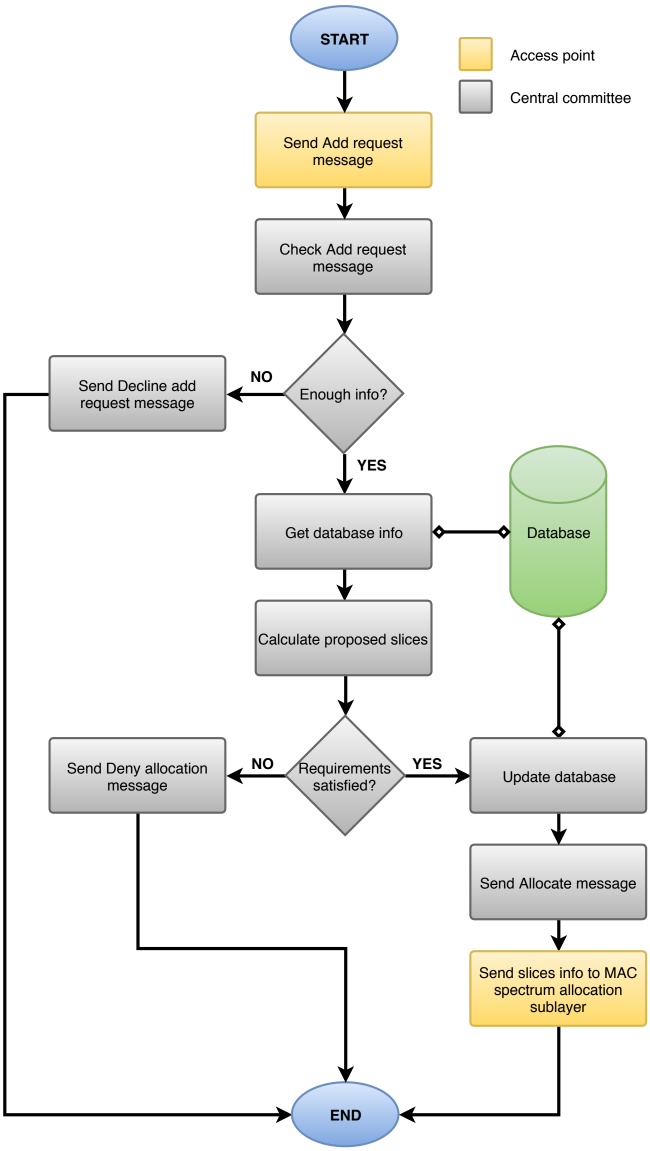


Figure 2.7: Allocation protocol of centralized model.

2. **Slice deallocation procedure:** Request deallocate message carries information about network requesting removal of a subset of its slices and its new bandwidth and duty cycle requirements. CE will do the recalculation and find a number of slices that can safely be deallocated, to keep up with new requirements of the managed network. CE is going to deallocate those slices

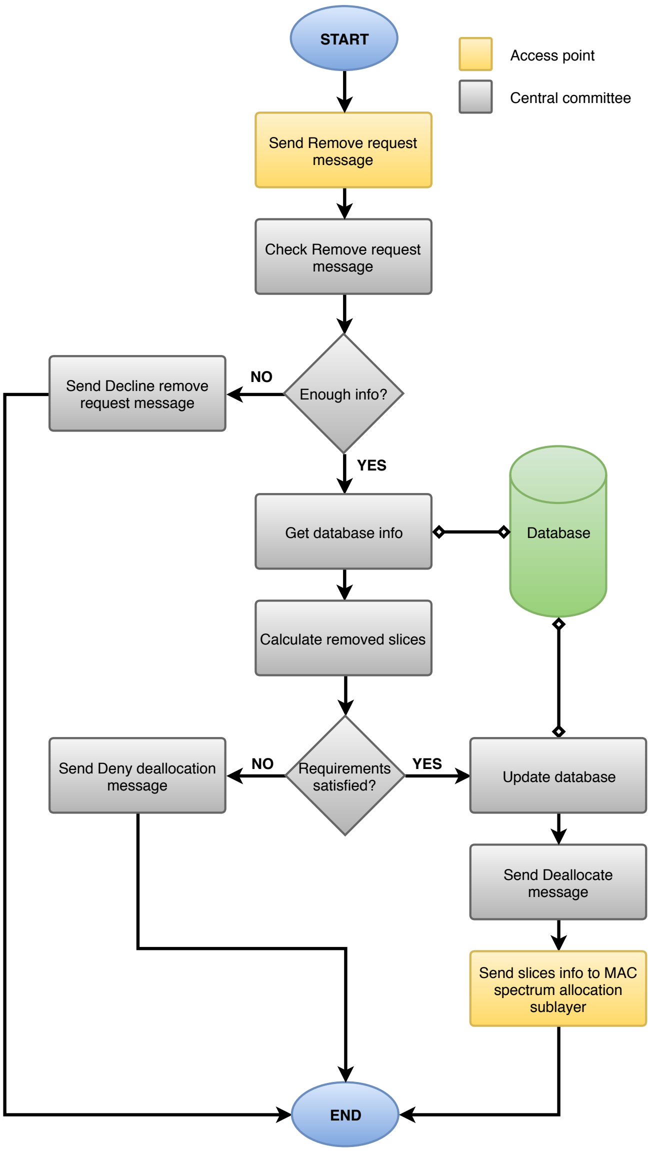


Figure 2.8: Deallocation protocol of centralized model.

from the database and notify EPs via Response deallocate message to do the same. Slice deallocation protocol is presented in Fig. 2.8.

3. **Slice move procedure:** Move procedure is similar to the two procedures mentioned above. CE receives the Request move message and it deallocates

part of the allocated frequency-time slices. Then CE responds to EP with a list of slices that should be deallocated and a list of the slices that should be used in their place from now on by EP's managed network. Slice move protocol is presented in Fig. 2.9.

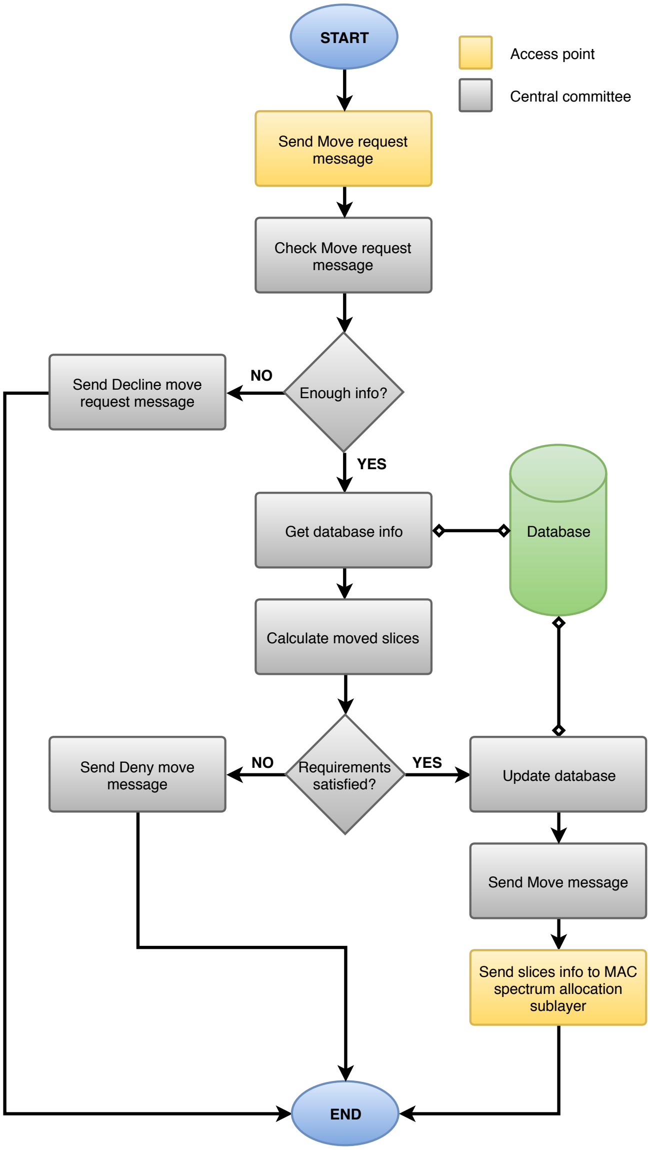


Figure 2.9: Move protocol of centralized model.

In case of the failure of any of the above procedures, CE keeps its database intact and notifies EP about the encountered issues. Based on the response, EP can send another request to the CE that would avoid the issues that occurred in the previously sent request. Once the EP successfully negotiates allocation/deallocation of slices, the next step is to inform SML about changes in spectrum usage. This is done by sending Slice update messages (Slice allocation and Slice deallocation messages). Slice allocation message contains a list of slices that EP acquired, while Slice deallocation message contains a list of removed slices. Slices are represented as frequency-time pairs with the reserved duration of the slice. According to that information, SML updates a local list of slices that can be used for transmission/reception. If there are slices where coexistence with other technologies should be in place, the local list of the slices contains appropriate additional information, so that the related coexistence schemes can be enabled in the MAC layer of the internal network nodes.

4. **Housekeeping (heartbeat) procedure:**

   For housekeeping of the whole system, CEs periodically send probe messages (Heartbeat request messages) to every EP that they have stored in the database. After receiving the probe message, EP answers with the Heartbeat message that contains spectrum usage information in form of used frequency-time slices. After the reception of the Heartbeat message, CEs are aware that the network previously approved with part of the spectrum is still active and can compare information from heartbeat message with the information stored in CE's database. If data from the heartbeat message and database differs, so that message contains a reduced or in any way altered number of slices, CE deallocates slices that are not part of the heartbeat message and is going to reuse them for other purposes. If the heartbeat message reports usage of slices that are not stored as used by the corresponding network, CE sends Response deallocate message with slices that need to be released by the network. Therefore, with the execution of periodic housekeeping protocol, besides checking for inactive networks, CEs check if any systematic errors occurred during the process of slice allocation/deallocation and make sure that those issues are fixed promptly. In case of an unanswered Heartbeat request message, after the $Twait$ delay period has passed, CE sends another probe message and this procedure continues until the $counterMax$ value is exceeded. Then, CE concludes that probed managed network is not operating anymore in the radio spectrum and that it has vacated the spectrum without sending the Request deallocate message. CE can safely deallocate from the database all the slices previously used by that network and assign them to other networks that request part of the spectrum in the future. Values for $Twait$ and $counterMax$ can be determined per network deployment type (static long-term deployment or opportunistic short-term local deployments) based on experimental results and these values can differ between different technology types. The housekeeping procedure is presented in Fig. 2.10.

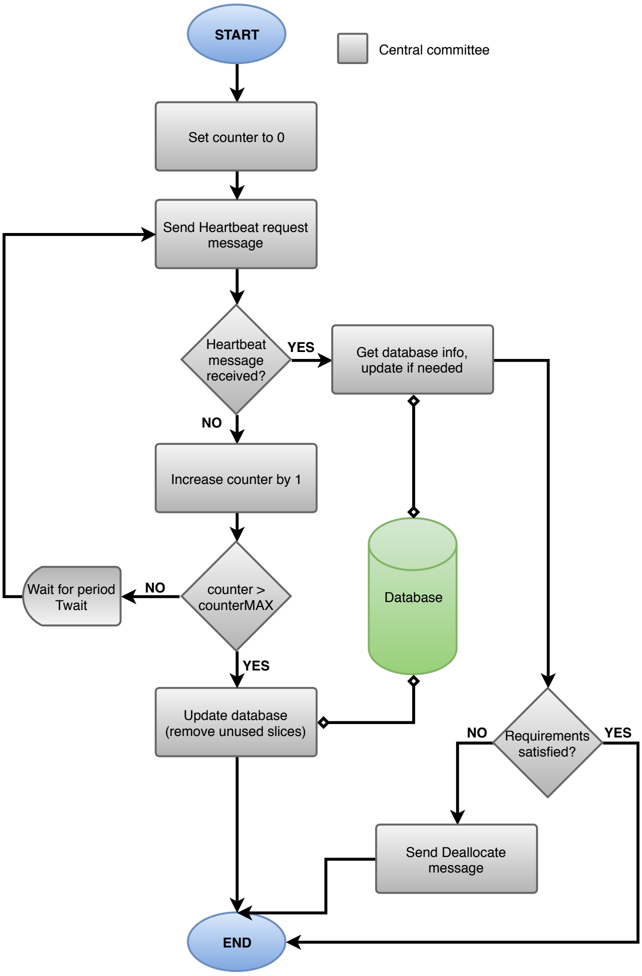


Figure 2.10: Housekeeping protocol of centralized model.

2.4.3 Decentralized model proposal

For dynamic spectrum sharing of unlicensed bands in the case of local area networks, we propose a decentralized spectrum sharing model that can operate in the absence of a CE, as there may not be any reachable CEs to control usage of the spectrum. The main focus of this model is on improved utilization and interference avoidance within the unlicensed spectrum. Because heterogeneous wireless networks present in remote areas are not managed by CEs, instead of managed networks, we shall refer to them as self-managed networks.

For successful negotiation of the available unlicensed spectrum in local area networks, EPs should take the responsibility of exchanging information on spectrum usage. For that to be possible, all EPs of the local networks need to have integrated a common long-range wireless technology besides the primary wireless technology used for data transmission. The architecture of a decentralized DSS model is presented in Fig. 2.11.

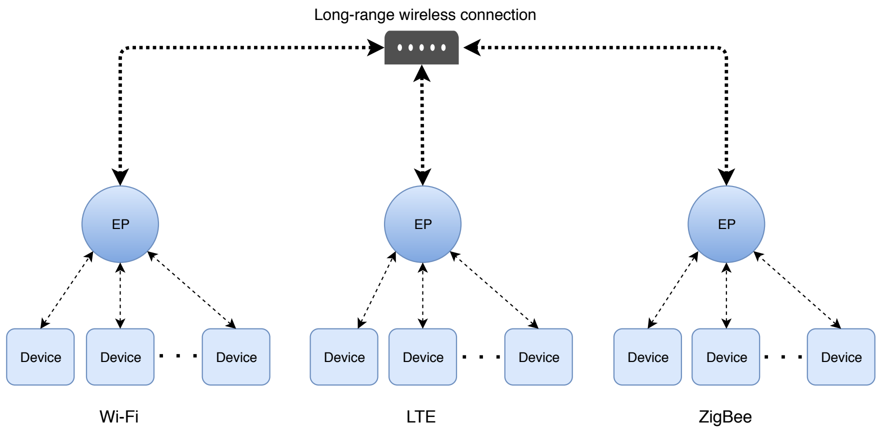


Figure 2.11: Decentralized model architecture.

For communication between the EPs in a local area, a single time-frequency slice (control radio slice) in the unlicensed radio spectrum is reserved and this slice is not available to local self-managed networks for data transmission. If an EP of a local self-managed network does not have integrated long-range wireless technology, another possibility for entering the unlicensed spectrum is the employment of a local Artificial Intelligence (AI) module and radio spectrum sensing capabilities. Spectrum sensing is used to inform the AI module about the state of the spectrum. Based on the sensing information, the AI module can choose radio spectrum slices where no transmission activities are recorded for some predefined period. A distributed approach for accessing radio spectrum that relies on spectrum sensing capabilities is heavily researched and present in form of Cognitive Radios (CRs) [43–45]. An example of the coexistence of the three most common wireless technologies LTE, Wi-Fi, and ZigBee in unlicensed radio spectrum is shown in Fig. 2.12. The proposed sharing model can be employed to offer a complete solution for decentralized spectrum access while increasing spectrum utilization and minimizing possible interference between competing networks on the spectrum slice level.

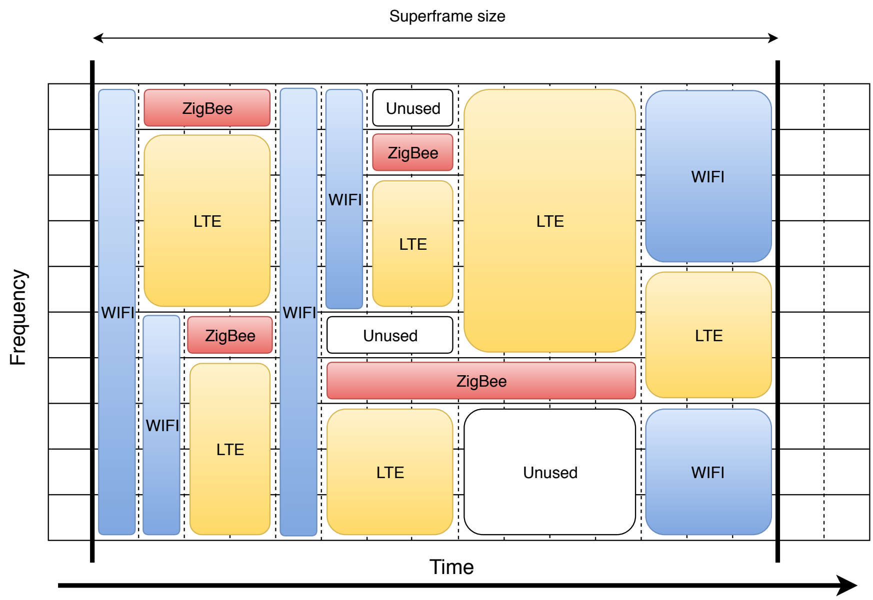


Figure 2.12: Example of decentralized model spectrum allocation.

If a new network is trying to establish itself in a local area where many other networks are operating, there might not be enough free slices available in the unlicensed bands. In that case, EP can choose slices that are used by the networks with technologies that can coexist with its wireless technology. This approach must not drastically reduce the performance of the other networks already using spectrum slices. In any case, even if two self-managed networks use wireless technologies that are perfectly cooperative with each other, the same slices should be used by different self-managed networks only as a last resort.

Similar to the centralized model proposal, EPs in the decentralized model proposal, which use long-distance wireless technology for the exchange of control messages, need to have a method of obtaining global location and global time information. Moreover, global time synchronization needs to be achieved on the node level, directly by network nodes or by synchronization with EPs. If the AI and CR approach is used, time synchronization is not mandatory but it would improve interference mitigation, while localization, in this case, is not needed.

2.4.4 Radio spectrum slices negotiation protocol (decentralized model)

In this section, the decentralized procedure for the allocation of spectrum slices in a local area is briefly described and its flow diagram presented in Fig. 2.13.

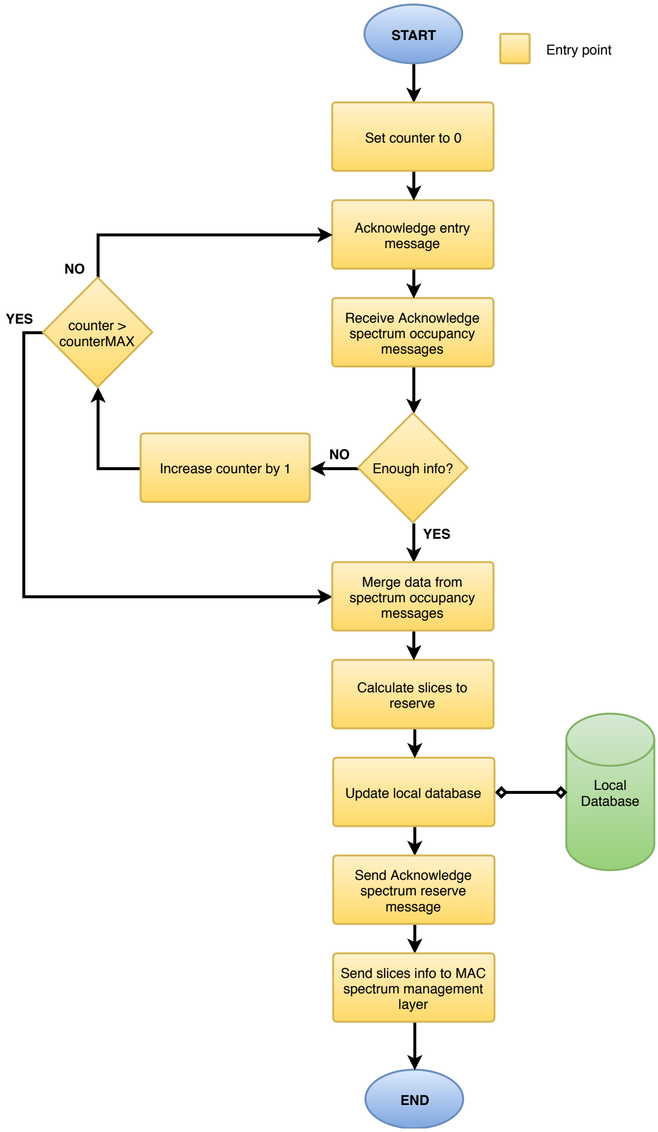


Figure 2.13: Decentralized model protocol for accessing unlicensed radio spectrum in the local area.

For mutual coordination of self-managed networks in the local area, five different types of Acknowledge messages are used: entry, spectrum occupancy, spectrum reserve, presence, and departure message type.

When a new self-managed network wants to access the unlicensed spectrum in a local area, its EP first sends an Acknowledge entry message. Upon reception of the entry message, other EPs in the local area respond with Acknowledge spectrum occupancy messages containing data about the used spectrum, spatial distribution of the networks, and additional data about wireless technology used by each network. If EP does not receive any response from other EPs in the local area or if the received information is not sufficient, then EP retries sending an entry message. After the maximum number of retries $counterMax$ is reached or after enough data about the state of the unlicensed spectrum in the local area is acquired, EP merges all received information and creates a table of used and unused spectrum slices. Then EP calculates the number of slices needed to satisfy its network's requirements and it chooses which time-frequency slices to allocate until a sufficient number of slices is reached. The local database is updated with all slices allocated by other networks present in the local area and slices allocated by that network itself. An Acknowledge reserve spectrum message is sent to all other EPs to acknowledge spectrum slices that the new network would use in the future. Other EPs update their local databases with newly allocated slices. The next step for the EP of the new network is to inform SMLs of all internal nodes in the network about the slices to be used for the transmission/reception. If the local area is too congested with many other networks already present, new networks may allocate the slices, where cooperation or coexistence with other technologies is required. All networks that are sharing slices with other neighboring networks, need to inform SMLs to enable coexistence schemes in the MAC layer. SMLs receive the same Slice update messages as in centralized proposal and after the reception of these messages, they can inform MAC layers about acquired slices.

Local self-managed networks' EPs should periodically send Acknowledge presence messages to inform other networks that their network is still active. If after some time one self-managed network appears idle, other EPs should update their local databases by removing slices previously used by the idle self-managed network. Now, these slices can be reused in the future by existing self-managed networks or by any new self-managed network that attempts to enter the local area. If a local self-managed network stops operating, it should inform other local EPs about its intention to leave the unlicensed spectrum by sending Acknowledge departure message. If that procedure fails, unused slices would be freed in any case by idle timeout during which no Acknowledge presence message is transmitted by the local self-managed network.

2.5 Costs and benefits of proposed spectrum sharing models

Licensing regime that is used today, called command and control, allocates full non-overlapping frequency bands to licensees. Licensees are granted access to the spectrum for long predefined periods, with limited rules and imposed requirements regarding the use of the assigned spectrum. This model of the licensing regime is the main cause for spectrum scarcity and low spectrum utilization. The main benefits of the spectrum sharing models, which are proposed in this chapter, are increased efficiency in the utilization of radio spectrum and better interference mitigation in unlicensed bands. By allocating smaller chunks of the spectrum and thus reducing initial costs, new technologies and new products will be able to more easily enter the market. Reduced barriers for entering the market would allow better adjustment of wireless technologies to consumer trends and more innovation freedom in the field of telecommunications. Market demands would be the main driving force behind rational spectrum allocation and usage; more popular and dense deployed wireless technologies would be given more spectrum slices.

For the proposed sharing models, EPs need to be equipped with additional hardware for time synchronization, location acquisition, and wireless control channel support. In addition, for the centralized model proposal, static infrastructure would need to be installed and coordinated at least on the country level. However, with constant technological advances, hardware components are becoming cheaper and thus more accessible. Extra costs of new infrastructure and required hardware to accomplish the goals of our spectrum-sharing framework would be compensated by increased revenue from denser networks and higher efficiency use of the spectrum by all the new players that would be allowed to enter the market. New players would be able to more easily deploy new advanced networks with low costs for accessing and using the spectrum. These reduced costs would come from the fact that every wireless network would only pay for the exact part of the radio spectrum that is needed to satisfy the demands and only for the time that the network is active. Leasing of full frequency bands for a long period of time and for high prices, which constitutes a high cost-wall for every new player, would be replaced with dynamic leasing of smaller chunks of the spectrum. This would enable smaller providers of wireless services to enter the licensed spectrum and thus terminate the monopoly present today in telecommunication services. As a result, users of the telecommunication services would be able to choose between a wider range of different options, which would push the providers of the wireless services to be more competitive in terms of reduced consumer prices and provided user services. The wireless providers, offering the best value for money, would attract more users and have better revenues. As a result, users would be offered enhanced services and better QoS. At the same time, increased revenues would give better financial conditions for providers to request more slices in the radio spectrum. This would allow the wireless technologies and wireless providers that provide the best user experience to be in a better position to be granted bigger chunks of the priceless resource that is the radio spectrum, making it impossible for oligopolies to be formed in the wireless market of a country.

2.6 Conclusion

We have presented a generic, technology-agnostic DSS framework that, if adopted, can revolutionize the way the spectrum is allocated, shared and paid for. The proposed framework would give the ability to all national and regional Regulatory Committees to monitor and enforce spectrum usage policies in real time across their area of control while enabling them to charge per accurate spectrum use of each spectrum user. All stakeholders of the spectrum would be able to lease at any given time the spectrum they require, further leading to reduced costs since they would release the assigned spectrum when their traffic load is low. Coexistence problems between heterogeneous technologies would be defacto solved since there would be no chance of two incompatible technologies sharing the same spectrum slice. Dynamic sharing of the same spectrum slice based on MAC characteristics, which is the most common approach today towards coexistence, would still be possible for compatible technologies if the need arises. The proposed framework will present no conflict with any future technology, as there are no requirements towards any emerging technology other than to be able to perform a single spectrum access procedure within a definitive period of time. The proposed framework also supports the option of decentralized self-managed local network deployments, permitting the typical application scenario of local area network deployments, like a Wi-Fi network, with or without central control. Local deployments without central control support, can employ spectrum sensing and take intelligent decisions on spectrum usage to minimize local interference. Adopting the proposed framework would allow the research community to turn its focus from defining and prototyping one-to-one coexistence solutions with huge costs in time and effort, releasing a big research potential and allowing to focus on new wireless technologies, protocols, and PHY implementations. To the best of the authors' knowledge, this is the first study up to date proposing a generic, technology-agnostic, DSS/leasing framework for all available regulated radio spectrum, not aiming just for sub-parts of the radio spectrum, like the ISM bands, and not focused or applicable only for specific existing technologies.

Acknowledgment

The work was supported by FWO-SB project No 140048 SAMURAI (Software Architecture and Modules for Unified RAdIo control), the FWO EOS project No 30452698 MUSE-WINET (MUlti-SErvice WIreless NETwork) and the UGent-BOF-GOA project Disposable and biodegradable wireless networks for extreme conditions.

References

[1] J. G. Andrews et al., ”What will 5G be?”, IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1065-1082, Jul. 2014.

[2] S. Chen, and J. Zhao, ”The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication”, IEEE Commun. Mag., vol. 52, no. 5, pp. 36-43, May 2014.

[3] D. B. Cabric, R. W. Brodersen, ”Cognitive Radios: System Design Perspective”, Technical Report No. UCB/EECS-2007-156, Dec. 2007.

[4] ”Feasibility Study on Licensed-Assisted Access to Unlicensed Spectrum”, 3GPP, 2015. Available from: www.3gpp.org.

[5] ”New Work Item on Licensed-Assisted Access to Unlicensed Spectrum”, 3GPP, 2015. Available from: www.3gpp.org.

[6] ”Overview of 3GPP Release 13”, 3GPP, 2015. Available from: http://www.3gpp.org/release-13.

[7] ”The Book of Visions 2001 - Visions of a Wireless World”, Wireless World Research Forum, Dec. 2001.

[8] Q. Bi, G. I. Zysman, and H. Menkes, ”Wireless Mobile Communications at the Start of the 21st Century”, IEEE Commun. Mag., vol. 39, no. 1, pp. 110-116, Jan. 2001.

[9] W. W. Lu, ”Compact Multidimensional Broadband Wireless: The Convergence of Wireless Mobile and Access”, IEEE Commun. Mag., vol. 38, no. 11, pp. 119-123, Nov. 2000.

[10] G. I. Zysman et al., ”Technology evolution for mobile and personal communications”, Bell Labs Tech. J., vol. 5, no. 1, pp. 107-129, Jan.-Mar. 2000.

[11] M. Haardt, and W. Mohr, ”The Complete Solution for Third Generation Wireless Communications: Two Modes on Air, One Winning Strategy”, IEEE Pers. Commun., vol. 7, no. 6, pp. 18-24, Dec. 2000.

[12] ”RSPG Opinion on Licensed Shared Access, RSPG13-538”, Radio Spectrum Policy Group, Nov. 2013.

[13] ”ECC Report 205, Licensed Shared Access (LSA)”, ECC, Feb. 2014.

[14] A. A. Ayeni, N. Faruk, N. T. Surajudeen-Bakinde, R. A. Okanlawon, and Y. A. Adediran, ”Spatial Spectrum Utilization Efficiency Metric for Spectrum Sharing System”, International Journal of Digital Information and Wireless Communications, vol. 5, no. 1, pp. 44-51, 2015.

[15] P. Leaves, K. Moessner, R. Tafazolli, D. Grandblaise, D. Bourse, R. Tonjes, and M. Breveglieri, ”Dynamic spectrum allocation in composite reconfigurable wireless networks”, IEEE Communications Magazine, vol. 42, no. 5, pp. 72-81, May 2004.

[16] D. Thilakawardana, K. Moessner, and R. Tafazolli, ”Darwinian approach for dynamic spectrum allocation in next generation systems”, IET Communications, vol. 2, no. 6, pp. 827-836, Jul. 2008.

[17] C. Raman, R.D. Yates, and N.B. Mandayam, ”Scheduling variable rate links via a spectrum server”, Proc. IEEE DySPAN 2005, pp. 110-118, Nov. 2005.

[18] S. A. Zekavat, and X. Li, ”User-central wireless system: ultimate dynamic channel allocation”, Proc. IEEE DySPAN 2005, pp. 82-87, Nov. 2005.

[19] L. Cao, and H. Zheng, ”Distributed spectrum allocation via local bargaining”, Proc. IEEE Sensor and Ad Hoc Communications and Networks (SECON) 2005, pp. 475-486, Sep. 2005.

[20] J. Huang, R.A. Berry, and M.L. Honig, ”Spectrum sharing with distributed interference compensation”, Proc. IEEE DySPAN 2005, pp. 88-93, Nov. 2005.

[21] L. Ma, X. Han, and C.-C. Shen, ”Dynamic open spectrum sharing MAC protocol for wireless ad hoc network”, Proc. IEEE DySPAN 2005, pp. 203-213, Nov. 2005.

[22] S. Sankaranarayanan, P. Papadimitratos, A. Mishra, and S. Hershey, ”A bandwidth sharing approach to improve licensed spectrum utilization”, Proc. IEEE DySPAN 2005, pp. 279-288, Nov. 2005.

[23] M. G. Kibria, G. P. Villardi, K. Ishizu, F. Kojima, and H. Yano, ”Resource Allocation in Shared Spectrum Access Communications for Operators with Diverse Service Requirements”, EURASIP Journal on Advances in Signal Processing, Dec. 2016.

[24] J. Luo, J. Eichinger, Z. Zhao, and E. Schulz, ”Multi-Carrier Waveform Based Flexible Inter-Operator Spectrum Sharing for 5G Systems”, 2014 IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN), Apr. 2014.

[25] V. Brik, E. Rozner, S. Banerjee, and P. Bahl, ”DSAP: A Protocol for Coordinated Spectrum Access”, Proc. IEEE DySPAN 2005, pp. 611-614, Nov. 2005.

[26] T. Kamakaris, M. M. Buddhikot, R. Iyer, ”A Case for Coordinated Dynamic Spectrum Access in Cellular Networks”, Proc. IEEE DySPAN 2005, Nov. 2005.

[27] M. M. Buddhikot, P. Kolodzy, S. Miller, K. Ryan, J. Evans, ”DIMSUMNet: New Directions in Wireless Networking Using Coordinated Dynamic Spectrum Access”, Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks, Jun. 2005.

[28] H. Zhou, B. Liu, F. Hou, N. Zhang, L. Gui , J. Chen, and X. Shen, ”Database-assisted dynamic spectrum access with QoS guarantees: A double-phase auction approach”, China Communications, vol. 12, no. 1, pp. 66-77, Jan. 2015.

[29] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl, ”SenseLess: A database-driven white spaces network”, 2011 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), May 2011.

[30] Y. Dai, J. Wu, and X. Du, ”Hierarchical and Hybrid: Mobility-compatible Database-assisted Framework For Dynamic Spectrum Access”, IEEE Transactions on Network Science and Engineering, pp. 1-1, Apr. 2018.

[31] M. D. M., S. Srikanteswara, and B. Badic, ”Spectrum Sharing: Licensed Shared Access (LSA) and Spectrum Access Systems (SAS)”, Intel, Oct. 2015.

[32] R. Berezdivin, R. Breinig, and R. Topp, ”Next-Generation Wireless Communications Concepts and Technologies”, IEEE Communications Magazine, vol. 40, no. 3, pp. 108-116, Mar. 2002.

[33] C. Peng, H. Zheng, and B. Y. Zhao, ”Utilization and Fairness in Spectrum Assignment for Opportunistic Spectrum Access”, Mobile Networks and Applications, vol. 11, no. 4, pp. 555-576, Aug. 2006.

[34] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, ”NeXt generation/-dynamic spectrum access/cognitive radio wireless networks: A survey”, Computer Networks, vol. 50, no. 13, pp. 2127-2159, Sep. 2006.

[35] G. Middleton, K. Hooli, A. Tolli, and J. Lilleberg, ”Inter-Operator Spectrum Sharing in a Broadband Cellular Network”, 2006 IEEE Ninth International Symposium on Spread Spectrum Techniques and Applications, Aug. 2006.

[36] L. Xu, R. Tonjes, T. Paila, W. Hansmann, M. Frank, and M. Albrecht, ”DRiVE-ing to the internet: dynamic radio for ip services in vehicular environments”, Proc. of 25th Annual IEEE Conference on Local Computer Networks, pp. 281-289, Nov. 2000.

[37] D. Grandblaise, D. Bourse, K. Moessner, and P. Leaves, ”Dynamic spectrum allocation (DSA) and reconfigurability”, Proc. Software-Defined Radio (SDR) Forum, Nov. 2002.

[38] L. Ma, X. Han, and C.-C. Shen, ”Dynamic open spectrum sharing MAC protocol for wireless ad hoc network”, Proc. IEEE DySPAN 2005, pp. 203-213, Nov. 2005.

[39] S. Sankaranarayanan, P. Papadimitratos, A. Mishra, and S. Hershey, ”A bandwidth sharing approach to improve licensed spectrum utilization”, Proc. IEEE DySPAN 2005, pp. 279-288, Nov. 2005.

[40] X. Jing, and D. Raychaudhuri, ”Spectrum co-existence of IEEE 802.11b and 802.16a networks using CSCC etiquette protocol”, Proc. IEEE DySPAN 2005, pp. 243-250, Nov. 2005.

[41] J. Zhao, H. Zheng, and G.-H. Yang, ”Distributed coordination in dynamic spectrum allocation networks”, Proc. IEEE DySPAN 2005, Nov. 2005.

[42] P. Seidenberg, and M. Lott, ”Analysis of the inter-system interference with respect to the required minimum frequency separation.”, 4th Cost259 Meeting, 1998.

[43] A. Ahmad, S. Ahmad, M. H. Rehmani, N. U. Hassan, ”A Survey on Radio Resource Allocation in Cognitive Radio Sensor Networks”, IEEE Communications Surveys & Tutorials, vol. 17, no. 2, pp. 888-917, Feb. 2015.

[44] G. I. Tsiropoulos, O. A. Dobre, M. H. Ahmed, K. E. Baddour, ”Radio Resource Allocation Techniques for Efficient Spectrum Access in Cognitive Radio Networks”, IEEE Communications Surveys & Tutorials, vol. 18, no. 1, pp. 824-847, Oct. 2014.

[45] D. M. Alias, G. K. Ragesh, ”Cognitive Radio networks: A survey”, 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 1981-1986, Mar. 2016.

3

Dynamic and Collaborative Spectrum Sharing: The SCATTER Approach

In this chapter, another approach for solving underutilization and coexistence problems in wireless communications is presented. It is based on a full-stack system developed by the SCATTER team that took part in the Defense Advanced Research Projects Agency (DARPA) Spectrum Collaboration Challenge (SC2) competition. The Medium Access Control (MAC) layer of the system plays the main role in enabling dynamic spectrum sharing (DSS) and collaboration with other networks operating in the same frequency bands within the same collision domain. This chapter contains a brief description of the MAC layer, but also a description of other layers and their interconnections. A detailed description of the MAC layer and its DSS functionality is introduced later in Chapter 4.

⋆ ⋆ ⋆

This chapter is adapted from:

S. Giannoulis, C. Donato, R. Mennes, F. A. P. de Figueiredo, I. Jabandžić, Y. De Bock, M. Camelo, J. Struye, P. Maddala, M. Mehari, A. Shahid, D. Stojadinovic, M. Claeys, F. Mahfoudhi, W. Liu, I. Seskar, S. Latré, and I. Moerman, *Dynamic and Collaborative Spectrum Sharing: The SCATTER Approach*

Published in proceedings of the 2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), p. 473–478, Newark, New Jersey, USA, Nov. 2019.

The contribution of the author of this thesis is the implementation and testing of the MAC layer and DSS functionalities of the SCATTER wireless communication system, together with their connections towards other layers of the system.

Abstract This chapter presents the architecture and the basic principles behind the design and implementation of the SCATTER system, a wireless end-to-end communication system that participated in the Defense Advanced Research Projects Agency (DARPA) Spectrum Collaboration Challenge (SC2). The focus is mainly on presenting the architecture and the supported interactions between the different components of the system in order to deliver a truly dynamic collaborative spectrum allocation and usage while coexisting with numerous unknown heterogeneous wireless technologies.

3.1 Introduction

Certain parts of the spectrum, in particular the license-free Industrial, Scientific and Medical (ISM) bands, are overcrowded, while other parts, mostly licensed bands, may be significantly underutilized. As such, there is a need to introduce more advanced techniques to access and share the wireless medium, either to improve the coordination within a given band or to explore the possibilities of intelligently using unused spectrum in underutilized (e.g. licensed) bands.

Numerous efforts have already taken place in order to enable coexistence and/or cooperation between resident technologies, especially focusing on the ISM bands. Most of these efforts focus on delivering techniques for coexistence based on the inherent characteristics of specific technologies like Wi-Fi, Bluetooth, and ZigBee that reside in the ISM bands as well as LTE Licensed Assisted Access (LTE-LAA), LTE-Unlicensed (LTE-U), and MulteFire that also aspire to enter the ISM bands [1]. None of those techniques is generic enough to be generalized, even worse, most of them are simple sense-and-avoid techniques [2] or even fixed duty cycle based techniques that proclaim some technology as the main user of the spectrum band while the rest compete for the empty leftover parts.

The SC2 aims to deliver a new way of dynamically sharing the spectrum without the need for technology-specific knowledge, providing a framework of cooperative sharing of the spectrum while offering fairness and true coexistence in any possible spectrum range, aspiring to even break the barriers between free and licensed bands. With the recent launch of Citizen Broadband Radio Service (CBRS) in the United States, the dynamic spectrum sharing (DSS) era has just started [3].

3.2 SCATTER System-Level Architecture

3.2.1 General Architecture

SCATTER system has been designed to split every functionality within different modules, all connected through a common data bus where all information is exchanged. Fig. 3.1 shows the general architecture of our software design and summarizes the interaction across modules. The main purpose of this approach is i) to provide an abstraction layer that allows developers to code modules using different programming languages, choosing the most convenient one depending on the feature: while most of the modules are written in C/C++, the Machine Learning (ML) modules are written in python using third party frameworks such as TensorFlow[1] and CuPy[2] to offload matrix computations to the Graphics Processing Unit (GPU); ii) to offer a fail-safe system: as each module is an individual system process, in case that one of them crashes, SCATTER system can just restart that process during runtime; and iii) to follow a plug-in approach: replacing, deleting, or adding certain functionality, is just a matter of attaching another process to the data bus. Therefore, the Application Programming Interface (API) for any information exchange or supported functionality is defined as a message template that two or more modules can exchange.

As shown in Fig. 3.1, SCATTER has three main blocks: the Data and Control plane, as well as an RF Monitor (RFMON) module. The Data plane is composed of the User Data Management (UDM), Medium Access Control (MAC), and Physical layer (PHY) modules. The Control plane is composed of two main modules: the rule-based and the Artificial Intelligence (AI)/ML-based Control Layers (CLs). The RFMON module offers real-time spectrum monitoring in the form of a continuous stream of Fast Fourier Transform (FFT) samples. A system time module provides synchronization to all modules based on the system clock. Finally, some additional blocks are included to support specific requirements in the context of the SC2: the Collaborative Intelligent Radio Networks (CIRNs) Interaction Language (CIL) module, used to interact with other networks, and the environment and traffic flows Quality of Service (QoS) parser (we shall call a flow with specific QoS requirements a mandate from now on), responsible to accept input on required settings of the Radio Frequency (RF) front-end and data traffic types along with their QoS characteristics.

[1]https://www.tensorflow.org
[2]https://cupy.chainer.org

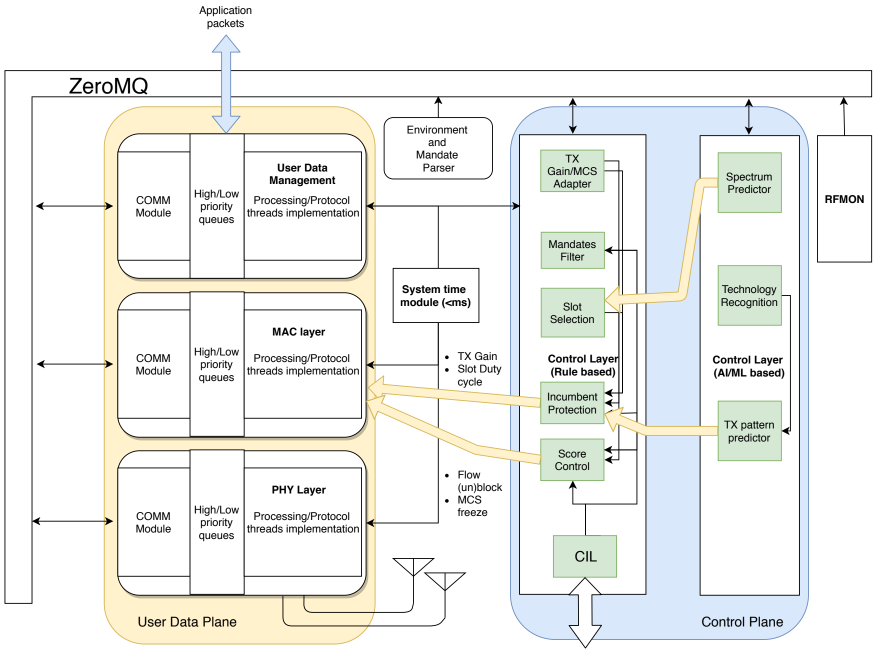


Figure 3.1: SCATTER system architecture.

3.2.2 Layers description

3.2.2.1 SCATTER PHY description

The high-level architecture of the SCATTER PHY is depicted in Fig. 3.2. The figure illustrates the different software/hardware layers composing the SCATTER PHY and the threads within each one of them. Red dashed arrows indicate data paths while black arrows indicate control/information interaction between threads.

The SCATTER PHY is implemented as Software Defined Radio (SDR) and is built upon the srsLTE library [8], evolving beyond the existing LTE features. It communicates to a Universal Software Radio Peripheral (USRP) X family of SDR devices including NI's RIO platforms[3] [5] for pass-band conversion and transmission using the USRP Hardware Driver (UHD) software API [6, 7]. As it can be seen in the figure, the individual PHY modules are connected to the ZeroMQ (Data/Control) module, also known as 0MQ, which interconnects the SCATTER PHY with the MAC layer through the ZeroMQ bus [4]. This module manages the exchange of control and statistics messages between the SCATTER PHY and MAC layer.

Communication with the SCATTER PHY is implemented through a well-

[3] https://www.ettus.com/product/category/USRP-X-Series

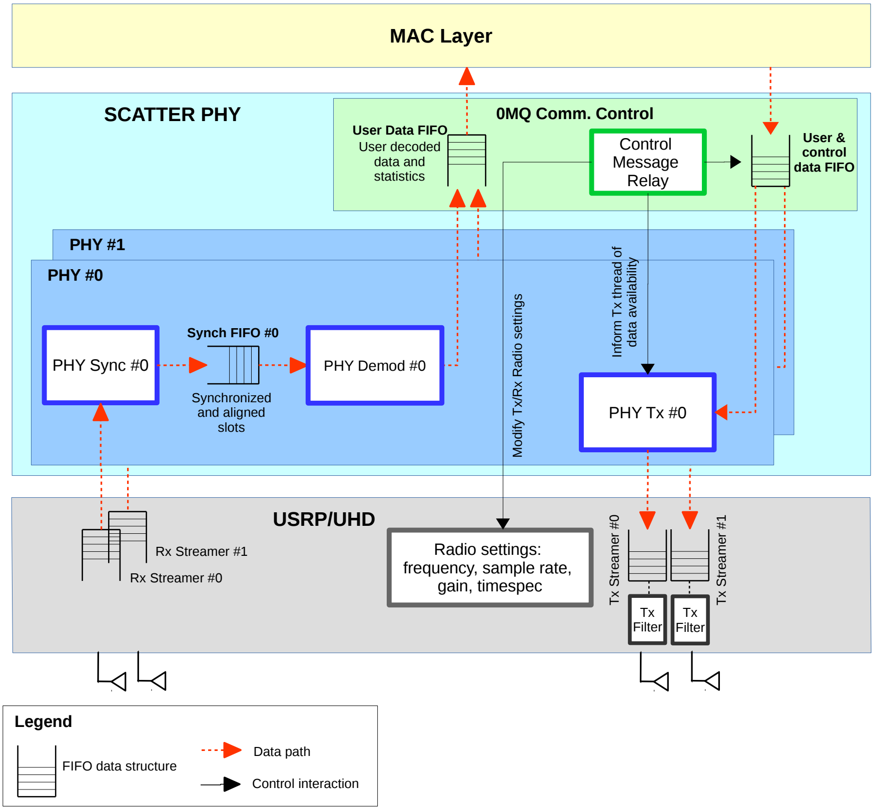


Figure 3.2: High-level architecture of the SCATTER PHY.

defined interface designed with Google's Protocol Buffers (protobuf)[4] for data serialization, coupled with the ZeroMQ messaging library [4] for distributed exchange of control, statistics and data messages. Implementing the ZeroMQ push-pull pattern allows local or remote MAC layer's real-time configuration of several parameters and reading of several types of information/statistics provided by the SCATTER PHY. Based on the ZeroMQ logic, PHY and MAC layers are able to exchange control and data messages following a non-blocking communication paradigm. The SCATTER PHY was designed to be completely decoupled and independent of the MAC layer module, not posing any constraints on hardware, software, and/or programming language adopted by it. The SCATTER PHY contains the following set of main features:

- OFDM waveform: We adopt Orthogonal Frequency Division Multiplexing (OFDM) as the SCATTER PHY waveform. OFDM is a mature technol-

[4]http://code.google.com/p/protobuf/

ogy, which is implemented in a wide range of products due to its several advantages such as robustness to severe multi-path fading, low implementation complexity, easy integration with Multiple-Input Multiple-Output (MIMO), simple channel estimation, etc. [11].

- Bursty transmissions: with discontinuous transmissions, it is possible to improve the use of the available spectrum and to coordinate its usage with other networks/radios in an opportunistic/intelligent/collaborative way.

- Dual-Concurrent PHYs: having two physical interfaces simultaneously transmitting and receiving at independent frequencies enables the Multi-Concurrent-Frequency Time Division Multiple Access (McF-TDMA) scheme to be implemented by the MAC layer. The ability to allocate concurrent slots allows for more flexible spectrum utilization, as vacant disjoint frequency chunks can be concurrently used.

- Field Programmable Gate Array (FPGA)-based filtered transmissions: filtering the transmitted signal effectively minimizes out-of-band emissions, allowing better spectrum utilization by enabling radios to have their transmissions closer to each other in the frequency domain.

- Frequency Division Duplexing (FDD) full-duplex operation: both PHYs operate completely independently, meaning that Tx and Rx modules are able to transmit and receive at different channels, set different gains and use different PHY bandwidths.

- Timed-commands: this feature allows the configuration of the exact time in the future to (i) start transmission and (ii) change Tx/Rx frequencies/gains.

3.2.2.2 RFMON description

The RFMON module is very important to the whole system as it gives the CL a local insight into the spectrum usage, by enabling CL to access spectrum sensing measurements. These measurements are used to train ML algorithms, which are employed to better understand the environment, optimize the spectrum usage, and cooperatively work with other networks while being completely agnostic to other network's characteristics. RFMON continuously monitors the whole competition bandwidth which can go up to 40MHz in an SC2 scenario. Performing this compute-intensive task on the FPGA reduces Central Processing Unit (CPU) load. It also reduces the amount of data to be transferred from USRP to host, as only periodic snapshots (time-averaged spectral energy) are sent to the host. This custom FPGA module along with transmission Finite Impulse Response (FIR) filters was built and integrated within the SCATTER system using the Ettus Research RFNoC

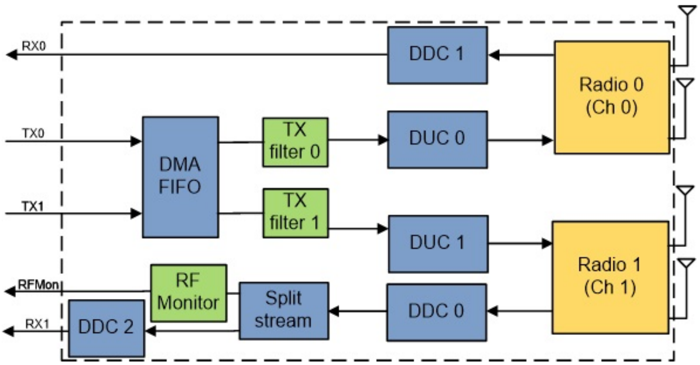


Figure 3.3: FPGA configuration with dual PHY and RFMON.

framework as shown in Fig. 3.3. As SCATTER uses dual-concurrent PHY, samples from the second radio are split into two streams, with one stream feeding RFMON and the other stream feeding Rx decoding pipeline.

3.2.2.3 MAC description

The SCATTER MAC protocol is based on an enhanced Multi-Frequency Time Division Multiple Access (MF-TDMA) scheme, taking advantage of our dual concurrent PHY support and the separated Rx and Tx channels offered per PHY. Since two slots can be active at the same time using PHY0 and PHY1, we can support the McF-TDMA table where two slots can be allocated for Tx and two slots for Rx at any given timeslot per node. When the node allocates Tx and Rx slots at the same timeslot, self-interference at adjacent channels needs to be taken into account. One channel separation between Tx and Rx channels has proven to be sufficient. By employing the McF-TDMA scheme, our CIRN can utilize the entire offered spectrum if needed. MAC layer maintains the McF-TDMA table per node and updates it with every successful slot allocation/removal procedure. An example of the McF-TDMA table is shown in Fig. 3.4.

MAC layer is also responsible for exchanging slot allocation/removal control messages and for notifying neighbor nodes about any newly allocated/released time-frequency slots. Based on the latency and throughput requirements of an incoming flow, the MAC layer initiates one/multiple slot allocation procedures towards the destination node in order to serve the incoming traffic. As a protection mechanism against failures in slot allocation/removal procedures, the MAC layer periodically broadcasts table status to neighboring nodes, in order to align all McF-TDMA tables.

The MAC protocol is based on three generic operations:

- Support of a distributed slot setup protocol with any neighboring node that

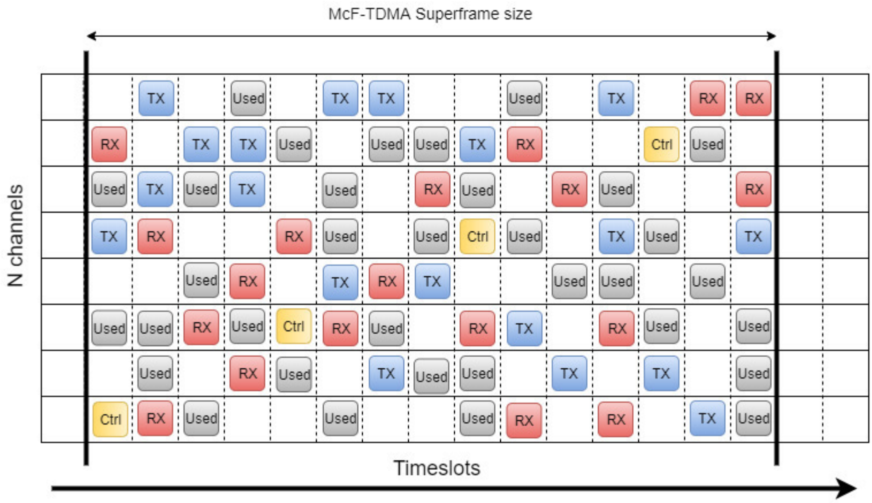


Figure 3.4: McF-TDMA slot table structure and possible states.

can perform allocation, removal, and move of any [channelX-timeslotY] slot. Each allocated slot serves traffic of a single link and all flows belonging to it.

- Maintain a scheduling MF-TDMA table [slot, channel, type of slot, node] that keeps track of the assignment of channel-time slot tuple to transmit or receive to/from a specific node as well as the locations of Control-Broadcast (CB) slots.

- Support a semi-static CB slot allocation scheme. CBs are based on slotted aloha medium access if there is a large number of nodes in the network or are divided into mini-slots, where each mini-slot is used by one node, offering a Time Division Multiple Access (TDMA) like medium access. The CB scheme adapts to the channelization of the available bandwidth during boot time to avoid the possibility of CB slots interfering all across the available spectrum.

Apart from the basic MAC operation, several other features exist in the MAC layer to support QoS requirements and link robustness. A double layer of buffering exists to support aggregation/fragmentation of incoming network layer Protocol Data Units (PDUs), in order to fit in the PHY PDU size (based on the running Modulation and Coding Scheme (MCS) on the link) and avoid wasted space. Also, retransmissions are supported based on aggregated acknowledgments (ACKs) sent from the receiver, informing the sender about successfully receiving a packet. Retransmission maximum retries are dynamically calculated per packet to ensure that the latency requirements of every packet will not be violated. Data packets are dropped proactively if it is found that they cannot be delivered in time to the destination.

3.2.2.4 Description of the User Data Management module

Our decision engine is designed to constantly keep track of the radio performance. To achieve this, UDM reports to CL several metrics per flow such as the number of packets per second and average packet size. These metrics are crucial for CL to understand how close the node is to fulfill the flow QoS requirements, and therefore, correctly quantify the success of our system. UDM monitors in runtime all incoming traffic flows and reports to several submodules of CL the required monitoring information. UDM also performs buffering when specific bursty types of traffic are injected into our system, taking into account the latency characteristics of the bursty incoming traffic, and reshaping the traffic flow to a Constant Bit Rate (CBR)-like flow.

3.2.2.5 Description of the CL and decision engine

The CL of the SCATTER solution holds the intelligence of the system, making use of all the knobs exposed by the MAC layer in addition to the ground truth vision provided by the RFMON. Those represent the main enablers of our decision engine. Any required decision is taken by the submodules that constitute our decision engine: from MCS and Tx gain adaptation, scheduling of the slots, Incumbent Protection (INCP) up to traffic prioritization and score control.

Individual modules of CL Here we present the individual modules constituting the CL.

- Link adaptation module

In the SCATTER system, link adaptation is controlling two main aspects of a link (e.g. Node A to node B), the MCS and the Tx gain that are used. We have kept the Rx gain static at 7 dBm, following the recommendation of NI and the DARPA team. The link adaptation plays a two-folded role: i) finding optimal settings for initializing and keeping links stable with high Packet Success Rate (PSR) while not interfered; ii) providing the first level of interference countering when other teams' radios would interfere with a specific link. This means that the link adaptation algorithm must be able to adapt Tx gain and MCS fast enough when interference is detected through PSR and link statistics like Received Signal Strength Indicator (RSSI) and Channel Quality Indicator (CQI). As all data packets are acknowledged from the receiver to the sender, the ACKs were also used to provide the sender with the receiver-based statistics, thus closing a fast control loop between the sender and the receiver. This control loop is the core of our link adaptation algorithm.

- Strategy module

We understand collaboration as helping all coexisting communication systems to reach an equal and acceptable level of performance. In the DARPA SC2 context, this translates into helping the ensemble reach the Bonus Threshold (BT) and protect potential incumbents. Consequently, we define two sorts of policies that depend on the status of a match: score policies and environmental policies. When the ensemble score is below BT, our score policy sets a target score according to other teams' scores from CIL. The target score is calculated based on the second weakest team's reported score plus an appropriate margin. Since the score of the ensemble may change rapidly, we have defined an aggressiveness factor that enables the margin of score to grow as fast as the ensemble score does, maximizing our achieved score at the moment the ensemble surpasses the BT. These policies have been implemented following a hybrid approach: the gateway (GW) sets the target score while the nodes dynamically choose the most beneficial mandates, by employing the mandate prioritization module. As soon as the ensemble crosses the BT, the target score is temporarily disabled and our strategy tries to maximize our score by enabling new mandates and wait for them to stabilize into a success state before enabling additional mandates.

On the other hand, a scenario may include incumbents that must be protected. Our strategy module implements two policies that handle passive and active incumbents. In the case of passive incumbents, the policy notifies the Tx gain adapter to minimize the transmission power for a given link while prioritizing flows with a smaller spectrum footprint. For the active incumbent, the policy learns the incumbent's Tx pattern, utilizing the information from the Technology Recognition (TR) module in order to dynamically enable/disable the overlapping regions of the superframe. In addition, our system is also capable to detect the presence of a jammer. The environmental policy limits the upper boundary of the MCS adapter to increase the robustness against interference while the slot predictor helps avoid the regions where the jammer is present.

- Traffic prioritization module

This module is responsible to sort the offered mandates, aiming to maximize the efficiency of our system (spectrum footprint vs achieved score). We have defined, for each type of mandate, a set of benefit-cost functions to calculate the benefit of a mandate. The most beneficial mandates are selected to be enabled first in order to reach our target score. For every type of mandate, our traffic prioritization module keeps track of its status and evaluates its success. In case that due to bad channel conditions, the system is not able to stabilize a mandate, the node blocks the flow and picks the next most beneficial mandate on the list.

- Slot allocation and spectrum management

The CL selects a slot to be used for communication between two nodes by combining the information from multiple input sources. To combine this information, the input data must be normalized. To this end, we have designed a slot selection system using multiple filters, where each input source is represented as a filter. For each filter output, the value is normalized between 0 and 1 as well as the final ”Goodness” value of a slot. The slot selection system supports two classes of filters: i) MUL-filters, where the values are multiplied by a factor that increases the impact of the filter during the slot allocation procedure, e.g. the INCP-presence filter downgrades the slots in the superframe that are overlapping with the incumbent's spectrum region or the External McF-TDMA filter, which makes sure two nodes do not select the same slot; ii) ADD-filters, with a summation effect to the slot Goodness. Such filters are the Slot prediction filter detailed in Section 3.2.2.5, the Channelizing filter, concentrating slot allocations in one or more frequency channels, and the Historic voxel filter that uses the voxel information from other teams provided via CIL to lower the Goodness value of parts of the spectrum that are used by other teams to reach the BT.

During a typical slot allocation procedure between two nodes, the Tx node, after applying all filters, selects slots with the highest Goodness from available slots and then proposes a subset of these slots to the Rx node. The Rx node then selects the best slot out of the proposed set based on its own filters and reports the selected slot back to the Tx node.

- Inter-node Reporting/Communication module

Since our strategy module follows a hybrid approach, every node reports the mandate performance metrics and node's spectrum usage to the GW. The central decision agent uses this information for two main purposes: to calculate the target score and to disseminate these metrics through CIL. To reduce control overhead, the size of these messages has been highly optimized. We have also implemented mechanisms to keep the nodes' individual states synchronized, e.g. the strategy module in the GW announces the policy to be executed locally in the nodes, the nodes report back to the GW the current policy running, target score, and mandate performance metrics.

- CIL support module

A node acting as the GW is the unique entity of a CIRN that is connected to the CIL network. This node is in charge of collecting mandate performance reports and spectrum usage from each node, packing this information in a single report and sending it to other networks. When other networks share any related information, the CIL module parses the messages and passes relevant values to the submodules of the decision engine for further processing.

Description of the ML/AI-based modules Spectrum sharing is about understanding and predicting the environment in real time. In SCATTER, this is possible using RFMON data providing ground truth vision to the decision engine and allowing spectrum state prediction.

- Slot prediction

Pinpointing and predicting holes in the spectrum is the key feature for our system. As detailed in [9], we have designed and implemented a Deep Convolutional Neural Network (CNN) to learn and predict the usage of the spectrum. As a preliminary step, the model was trained offline with spectrum data that was collected during the second and third phases of SC2. In addition to the acquired knowledge during the offline training, the model is fed with RFMON data for online training during matches to quickly learn, recognize, adapt and predict the (possibly new) behavior of other networks, aiming to avoid interference with other transmissions. Notice that after each match, the weights of the CNN were stored and used in new matches. Once the CNN has learned the right features to provide good performance in most of the scenarios, we keep using the online learning but the resulting weights are not used anymore in future matches. The outcome of the model is a matrix of values with the same dimensions as our superframe. These values are used as a filter into our slot selection module to help with the decision-making process of our nodes when they select, negotiate, and allocate slots.

- Technology recognition

Identifying what is in the spectrum is key for making better decisions on how to access the spectrum. The TR module uses RFMON data, which is framed according to our superframe size, to distinguish the following five types of radio signals signature: radar, jammer, SCATTER, other teams, and noise. The TR module was implemented following the same Deep Neural Network architecture as [10, 12, 13] but modified to support real-time processing with limited computing resources: instead of using raw samples, the TR module creates spectrograms with no overlapping by using the 32-averaged 512-point FFT samples collected by RFMON at 23.04 Msps (or 46.08 Msps if the available bandwidth is $>$ 20 Mhz).

3.3 Cross-layer and Data-Control plane interactions

In this section, the interactions between the main layers and modules of our system are presented, in order to make clear how every layer/module contributes towards achieving true dynamic collaborative spectrum usage.

3.3.1 RFMON - CL interactions

RFMON provides to the CL a local view of spectrum usage in the competition band. Implemented as an FPGA module, it computes the FFT magnitude square of the received samples and averages a configurable number of these blocks before sending them to the host. These snapshots of spectral energy are sent to the CL over ZeroMQ as vectors of 32-bit integers using an FFT size of 512 and averaging size of 32.

3.3.2 PHY - MAC interactions

The communication between the SCATTER PHY and the MAC layer is carried out through the exchange of four predefined messages. The first two, namely, Tx and Rx control messages, are used to manage subframe transmission and reception, respectively. The parameters carried by these two messages can be configured and sent by the MAC layer to the individual PHYs before the transmission of every subframe, hence allowing run-time configuration. The other two messages, namely, Tx and Rx statistics messages, are used to provide real-time feedback from each PHY to the MAC layer, yielding vital information necessary for upper layers to take action.

3.3.3 MAC - CL interactions

MAC layer feeds CL with all data necessary to make optimal decisions. It forwards to CL the Tx and Rx statistics received from PHY, it extracts statistics incorporated in ACKs and sends them to CL as well as slot usage statistics and flow statistics. CL calculates and feeds the MAC layer with link-level parameters like MCS and Tx gain per destination. To increase/decrease or just alter the spectrum signature of the network, it can inform the MAC layer to enable/block specific flows or increase/decrease the duty cycle of already allocated slots. MAC layer and CL have also bidirectional interaction in the context of a slot allocation/removal procedure. MAC requests new Tx slots based on incoming user data traffic, while CL decides which slots are the best slots to allocate. MAC layer also feeds CL with information about slot usage received from other nodes in the network. Since the CL is aware of the current link status and the performance of any traffic flow, it can inform MAC to enable/disable flow-specific robustness mechanisms in order to ensure that all flows abide by requested QoS characteristics.

3.3.4 UDM layer - CL interactions

UDM layer's main interaction with the CL is to forward information about the incoming traffic during run-time. This information is packet size per flow, packets/second per flow, and burst detection (size of burst in terms of packets and bytes).

3.4 Enhanced features of our Radio

The main overall characteristics that make our system unique, to the best of our knowledge, are as follows:
-Active/passive incumbent protection: Detecting and protecting incumbents is performed via a combination of CIL information and TR. For the active incumbent, TR provides the time and frequencies where the incumbent is detected in case of low interference, and CIL violation reports are used to enhance this information. The combined information is fed to pattern recognition and prediction algorithm, to learn and predict the time slots and frequencies where the active incumbent will transmit in the near future with respect to our superframe. For the passive incumbent, our nodes limit the transmission power jointly with smart traffic prioritization based on spectrum footprint. This way, our system allows others to transmit without crossing the interference threshold of all incumbent variants.
-Spatial reuse of frequency/time slots: By design, the SCATTER system can spatially reuse spectrum, as the slot allocation protocol provides the ability to solve the exposed and hidden node problem, exploiting the full potential of spectrum reuse in a multi-hop environment.
-Decentralized operation: All decisions for dynamic spectrum allocation and usage are taken locally at each node, with no need for centralized infrastructure or a special node to command and control our system.
-Spectrum usage prediction: Every Cognitive Radio (CR) can detect and react to what is happening in the spectrum today. SCATTER system, going beyond the CR state-of-the-art, can react to past events, but can also predict and react proactively to future spectrum usage events, taking into account patterns and predicting the future usage of the spectrum in run-time when other users are present.

3.5 Conclusions and Future work

In this work, we have presented a general overview of the SCATTER system, providing a detailed view of the system architecture and the main modules while pinpointing the unique features of our system design and radio capabilities. DARPA SC2 is the first step towards autonomous dynamic spectrum management, proving that collaboration across multiple radio systems is a mechanism that can facilitate coexistence in the same spectrum band. Our solution presents a state-of-the-art system that can solve many of the compelling scenarios where traditional radios are not able to cope. Our future work will include porting SCATTER outside of the context of the SC2 Colosseum emulator to enable experimentation with real wireless environments and assess the performance of our system in real-life scenarios.

Acknowledgment

The project leading to this publication has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 732174 (ORCA project) and from USA NSF under award #1836901 (ICE-T).

References

[1] V. Maglogiannis, D. Naudts, A. Shahid, S. Giannoulis, E. Laermans, and I. Moerman, *Cooperation techniques between LTE in unlicensed spectrum and Wi-Fi towards fair spectral efficiency*, SENSORS, vol. 17, no. 9, 2017

[2] A. Garcia-Saavedra, P. Patras, V. Valls, X. Costa-Perez, D.J. Leith, *ORLA/OLAA: Orthogonal coexistence of LAA and WiFi in unlicensed spectrum.* IEEE/ACM Transactions on Networking, 26(6), 2665-2678.

[3] M. M. Sohul, M. Yao, T. Yang and J. H. Reed, "Spectrum access system for the citizen broadband radio service," in IEEE Communications Magazine, vol. 53, no. 7, pp. 18-25, July 2015. doi: 10.1109/MCOM.2015.7158261

[4] P. Hintjens, *ZeroMQ Messaging for Many Applications*, O'Reilly Media, March 2013.

[5] National Instruments, *Overview of the NI USRP RIO Software-Defined Radio*, White Paper, December 2015. Available from: http://www.ni.com/white-paper/52119/en/

[6] Ettus Research LLC, *USRP Hardware Driver*, Available from: http://files.ettus.com/manual/

[7] Ettus Research LLC, *Products*, Available from: https://www.ettus.com/product

[8] Ismael Gomez-Miguelez, Andres Garcia-Saavedra, Paul D. Sutton, Pablo Serrano, Cristina Cano, and Doug J. Leith. 2016. srsLTE: an open-source platform for LTE evolution and experimentation. ACM, New York, NY, USA, 25-32. DOI: https://doi.org/10.1145/2980159.2980163

[9] Mennes, R., Claeys, M., De Figueiredo, F. A., Jabandžić, I., Moerman, I., & Latré, S. (2019). Deep Learning-Based Spectrum Prediction Collision Avoidance for Hybrid Wireless Environments. IEEE Access, 7, 45818-45830.

[10] Camelo, M., Shahid A., Fontaine, J., De Figueiredo, F., De Poorter, E., Moerman, I., & Latré, S. (2019). A Semi-Supervised Learning Approach Towards Automatic Wireless Technology Recognition, IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), Newark, NJ, USA, 2019.

[11] Sravanthi Kanchi, Shubhrika Sandilya, Deesha Bhosale, Adwait Pitkar, and Mayur Gondhalekar, *Overview of LTE-A technology*, IEEE Global High Tech Congress on Electronics (GHTCE), March 2014.

[12] A. Shahid, J. Fontaine, M. Camelo, J. Haxhibeqiri, M. Saelens, Z. Khan, I. Moerman, E. De Poorter, A Convolutional Neural Network Approach for Classification of LPWAN Technologies: Sigfox, LoRA and IEEE 802.15. 4g, 2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). IEEE, 2019.

[13] V. Maglogiannis, A. Shahid, D. Naudts, E. De Poorter, and I. Moerman, "Enhancing the Coexistence of LTE and Wi-Fi in Unlicensed Spectrum Through Convolutional Neural Networks," IEEE Access, vol. 7, pp. 28464-28477, 2019.

4

A Dynamic Distributed Multi-Channel TDMA Slot Management Protocol for Ad Hoc Networks

While Chapter 3 gives a general introduction of the Medium Access Control (MAC) layer of the SCATTER team's wireless system, in this chapter design and implementation of distributed scheduling functionality of the MAC layer is presented in more detail. Further, in this chapter it is proved that the proposed scheduling protocol can be applied to more dense and larger wireless networks operating in various wireless scenarios. Results show that the proposed protocol is a viable solution for problems of spectrum underutilization and interference mitigation in licensed and/or unlicensed radio bands.

⋆ ⋆ ⋆

This chapter is adapted from:

I. Jabandžić, S. Giannoulis, R. Mennes, F. A. P. de Figueiredo, M. Claeys, and I. Moerman, *A Dynamic Distributed Multi-Channel TDMA Slot Management Protocol for Ad Hoc Networks*

Published in IEEE Access, Volume 9, p. 61864-61886, Apr. 2021.

The contribution of the author of this thesis is the conceptualization, design, implementation, and validation of the majority of features presented in this chapter.

Abstract With the emergence of new technologies and standards for wireless communications and an increase in application and user requirements, the number and density of deployed wireless ad hoc networks is increasing. For deterministic ad hoc networks, Time Division Multiple Access (TDMA) is a popular medium access scheme, with many distributed TDMA scheduling algorithms being proposed. However, with increasing traffic demands and the number of wireless devices, proposed protocols are facing scalability issues. Besides, these protocols are achieving suboptimal spatial spectrum reuse as a result of the unsolved exposed node problem. Due to a shortage of available spectrum, a shift from fixed spectrum allocation to more dynamic spectrum sharing is anticipated. For dynamic spectrum sharing (DSS), improved distributed scheduling protocols are needed to increase spectral efficiency and support the coexistence of multiple co-located networks. Hence, in this chapter, we propose a Dynamic Distributed Multi-Channel TDMA (DDMC-TDMA) slot management protocol based on control messages exchanged between one-hop network neighbors and execution of slot allocation and removal procedures between sender and receiver nodes. DDMC-TDMA is a topology-agnostic slot management protocol suitable for large-scale and high-density ad hoc networks. The performance of DDMC-TDMA has been evaluated for various topologies and scenarios in the ns-3 simulator. Simulation results indicate that DDMC-TDMA offers near-optimal spectrum utilization by solving both hidden and exposed node problems. Moreover, it proves to be a highly scalable protocol, showing no performance degradation for large-scale and high-density networks and achieving coexistence with unknown wireless networks operating in the same wireless domain.

4.1 Introduction

Wireless ad hoc networks are networks without fixed infrastructure where nodes are often inexpensive devices with wireless transceivers, which could form different network topologies anywhere at any time [1, 2]. These networks operate autonomously in a self-organized manner with no central device managing the network. Typically, wireless ad hoc networks are multi-hop networks, where each node participates in forwarding data for other nodes in order to realize communication between two nodes out of direct communication range. Optionally, the nodes of an ad hoc network may be mobile. Due to low cost, easy deployment and maintenance, ad hoc wireless technology is gaining more and more attention recently with an increasing number of widespread applications being based on ad hoc technology, in particular for monitoring, surveillance, mission-critical and vehicular applications [3]. Wireless Sensor Networks (WSNs) are a specific class of wireless ad hoc networks, forming the basis for the Internet of Things (IoT). However, despite its widespread adoption and the advantages it offers, there are still challenges that need to be solved for ad hoc technology. According to [4], there are two main challenges for ad hoc networks: Quality of Service (QoS) guarantees and scalability.

For accessing the shared wireless medium in ad hoc networks, two families of Medium Access Control (MAC) protocols are dominant. The first family is contention-based protocols, usually using the Carrier Sense Multiple Access (CSMA) technique. Such MAC protocols use available bandwidth on demand and are very flexible and efficient for low traffic load conditions and small network sizes [5]. When network size increases and network traffic is high, CSMA-based protocols are not able to satisfy QoS requirements, implying that CSMA-based protocols are not scalable. A second family of MAC protocols for ad hoc networks are contention-free protocols, usually based on the TDMA mechanism. TDMA-based medium access is one of the most common medium access methods where the wireless medium is time-shared by all nodes. Channel bandwidth in the network is divided into time frames, called superframes, with every superframe further partitioned into time slots. Multi-Frequency Time Division Multiple Access (MF-TDMA) extends the basic TDMA medium access method, which uses only one frequency channel, to multiple channels. Slots in an MF-TDMA superframe are represented as time-frequency tuples. In TDMA-based protocols each node transmits only during slots allocated to it, avoiding any contention for accessing the shared medium [6]. Compared to CSMA-based protocols, TDMA-based protocols mitigate internal collisions and thus improve delivered QoS for large-scale networks with high traffic demands. Due to its favorable properties in terms of scalability, TDMA scheduling techniques have gained attention for larger ad hoc networks in recent years [7]. However, the reliability and throughput of networks with TDMA access schemes may still be impacted by external interference or the occurrence of exposed/hidden nodes.

For the function of slot allocation in TDMA schemes, there are static and dynamic algorithms. As ad hoc networks need to support constant changes in traffic demands and network topology, dynamic scheduling algorithms are known to outperform static scheduling algorithms [8]. There exist two main models for handling dynamic TDMA scheduling: centralized and distributed. Centralized models consist of one or more control nodes that gather information about the network state and make scheduling decisions that are advertised to each node. In distributed models, decision-making is done at the node level based on local information on the network without requiring any centralized control; nodes exchange information about slot usage with their neighbors in order to take distributed decisions on slot allocation.

Even though centralized scheduling protocols can offer close to optimal solutions for some use cases as they have global knowledge of network topology and traffic patterns [9], they are not suitable for networks with frequently changing topology and traffic demands over time. Changes in network topology or traffic patterns result in continuous schedule recalculations and increased control overhead, thus leading to degraded network performance. Moreover, centralized scheduling protocols are not scalable as they incur high control overhead for large-scale wireless networks. In dynamic and large ad hoc networks, distributed slot allocation algorithms are preferred to cope with scalability and changes in the network topology [10]. Also, distributed algorithms are more fault-tolerant, as a major problem in centralized algorithms is the existence of a single point of failure; if the central control node fails or disconnects, slot scheduling cannot be executed anymore. In any case, whereas many distributed scheduling protocols are proposed so far, an increase in size and/or density of wireless networks still induces scalability issues for existing protocols. The most common reason for the scalability issues in large-scale ad hoc networks is their multi-hop nature, which highly depends on network size and packet forwarding capabilities [11].

In large multi-hop ad hoc networks, hidden and exposed node problems represent a serious issue [12]. The hidden node problem may significantly increase the number of collisions and degrade network performance. On the other hand, the exposed node problem leads to the underutilization of the available radio spectrum and also results in reduced network throughput. In [12], it is concluded that for ad hoc networks with node density greater than 4 per collision domain, which is the case for most real-life wireless networks, the number of exposed nodes is higher than the number of hidden nodes. Hence, there is no doubt that for topologies of large-scale ad hoc networks, performance degradation is more influenced by the exposed node problem. Whilst most existing distributed TDMA protocols recognize and tackle the hidden node problem, only a minority of protocols are capable to detect the presence of exposed nodes and to minimize the huge impact in multi-hop ad hoc networks. Besides, existing protocols are mainly focusing on solving the dynamic TDMA scheduling problem for a single-frequency channel. Whereas it is possible to enhance most protocols with multi-channel TDMA scheduling capabilities, the associated overhead and scalability issues may degrade protocol performance.

To overcome the drawbacks of existing TDMA scheduling solutions, this chapter presents a dynamic distributed multi-channel TDMA slot management protocol called DDMC-TDMA. Its unique design makes it a generic and scalable solution suitable for large-scale high-density networks that support dynamic traffic demands and node mobility. DDMC-TDMA mitigates the exposed/hidden node problems and is inherently robust against external interference, leading to improved wireless QoS guarantees in terms of reliability, throughput, and latency.

In this work, DDMC-TDMA is presented and analyzed for ad hoc networks, but DDMC-TDMA is also more generally applicable to other types of wireless networks like infrastructure-based networks or satellite networks.

DDMC-TDMA was initially designed in the context of the Defense Advanced Research Projects Agency (DARPA) Spectrum Collaboration Challenge (SC2) [13]. This chapter discloses the key DDMC-TDMA features and performance, as successfully validated during SC2. DDMC-TDMA was integrated as the MAC solution for the wireless system of team SCATTER (see Chapter 3). The SCATTER system proved to be one of the top-performing systems in the competition, and as an integral component of the SCATTER system, DDMC-TDMA was experimentally validated throughout the 3 years of the SC2 competition. It was experimentally validated (i) how DDMC-TDMA copes with external interference, in various challenging dynamic wireless scenarios including mobility, changing traffic patterns, single-hop or multi-hop scenarios, and (ii) how DDMC-TDMA solves hidden and exposed node problems. The main scope of DARPA SC2 was to demonstrate the feasibility of DSS in the presence of other unknown wireless technologies and interferers. However, the maximum size of wireless networks in SC2 was limited to only 10 nodes. To prove that the DDMC-TDMA protocol, as developed for DARPA SC2, can scale and adapt to a wider range of network sizes and dynamic scenarios, simulations using an ns-3 model have been executed and analyzed in this chapter. Generating DDMC-TDMA models for the ns-3 environment [14] was a straightforward task as the DDMC-TDMA solution for the SCATTER system has been developed with scalability in mind.

The main contributions of this chapter can be summarized as follows:

1. The proposed DDMC-TDMA protocol offers distributed slot allocation and removal functionality. It is a topology-independent solution that only relies on MAC layer control messages exchanged between neighboring nodes.

2. TDMA scheduling with low control overhead is provided for both single-channel and multi-channel environments.

3. By design, DDMC-TDMA solves both hidden and exposed node problems that may occur in large multi-hop wireless networks.

4. DDMC-TDMA represents a robust solution against interference from other networks.

5. DDMC-TDMA automatically adapts its schedule to traffic, allocating and removing slots based on changing traffic demands.

6. In-depth simulations based on an ns-3 model are performed to validate the behavior of the protocol in multiple realistic wireless scenarios with various network topologies of different sizes and diverse system requirements.

For the remainder of the chapter, we present in Section 4.2 the state-of-the-art of distributed TDMA scheduling protocols applicable to ad hoc networks. Design goals and architecture of DDMC-TDMA are explained in Section 4.3, with a detailed description of the slot allocation and removal procedures. Validation of the proposed protocol, based on multiple simulations executed, is presented and analyzed in Section 4.4. The chapter ends with a summary of the main conclusions in Section 4.5.

4.2 Related work

Several distributed algorithms for dynamic TDMA scheduling have been proposed in the existing studies for different classes of ad hoc networks.

Distributed TDMA scheduling protocol called Five-Phase Reservation Protocol (FPRP) is presented in [15]. FPRP is a heuristic TDMA slot allocation protocol that can reasonably allocate spectrum resources according to the requirements [16]. The five-phase reservation process dynamically decides the winner of each data slot, where every data slot has its own reservation slot. During reservation slots, nodes contend for the associated data slots. All the one-hop and two-hop neighbors need to approve the request for slot allocation, which leads to non-conflicting slot allocations. As all the one-hop and two-hop nodes need to approve slot allocation request, control overhead increases significantly with the increase of network size and network density. Non-conflicting slot allocation solves the hidden node problem, but FPRP only supports the reuse of the same slots by the nodes beyond two hops, which results in the occurrence of the exposed node problem. Other drawbacks of FPRP are that it is only focused on broadcast scheduling and during FPRP execution it is required that network topology does not change.

In [17] distributed randomized time slot scheduling algorithm called DRAND is proposed. DRAND represents an efficient and practical distributed scheduling algorithm [10]. It is an extension of the graph coloring scheme and it is based on heuristic centralized solution RAND [18]. RAND generates efficient slot schedules but does not offer an optimal solution and it carries all the drawbacks of centralized solutions. DRAND represents distributed implementation of RAND and thus, achieves the same channel efficiency but with increased message complexity and convergence time. This solution solves the hidden node problem, but the exposed node problem is not avoided and achieved spectrum utilization is not optimal. Wireless nodes in the DRAND algorithm work in cycles, during which control messages are exchanged between neighbor nodes. The main drawback of the DRAND algorithm is the possible occurrence of unsuccessful cycles, resulting in extra control message overhead and algorithm running time. An increase in network size and network density results in increased control overhead and increased probability of failed cycles, leading to performance degradation of the DRAND algorithm. Therefore, it can be concluded that the DRAND algorithm is not scalable.

An extension of the DRAND algorithm called Localized-DRAND (L-DRAND) is proposed in [19]. L-DRAND is a distributed slot allocation algorithm that enhances DRAND characteristics by adding features for localization and exchanges position information between network nodes. As it is a position-based algorithm, it requires that every node possess localization capability. This algorithm is based on Lamport's bakery algorithm and gives slot allocation priority to nodes close to the center of the wireless network, as it is assumed that these nodes have the largest number of neighbor nodes. L-DRAND reduces the run time of the DRAND algorithm but at the expense of an increased number of control messages with increased message complexity.

Energy-Topology DRAND (E-T-DRAND) [9] is an amelioration of the DRAND algorithm where an Energy-Topology (E-T) factor is taken into account. Definition of the E-T factor is based on the influence of residual energy and topology on slot allocations. In E-T-DRAND, nodes require residual energy information from their neighbors, which is exchanged via control messages and stored in an appropriate table. During a slot allocation procedure, nodes with the lowest energy level and the highest degree are given priority. This algorithm is useful in the case of WSNs, where it reduces the energy consumption of nodes with low residual energy. The performance of the E-T-DRAND algorithm is better compared to the DRAND algorithm; it improves slot allocation efficiency and reduces control overhead and energy consumption. However, as noted in [16], the message complexity, slot allocation time, and energy consumption are still too high, with a significant increase in slot allocation time for large-scale networks and the possibility of non-convergence of slot allocation time and the number of rounds.

Hence, in [16] authors proposed another amelioration of DRAND called Exponential Backoff and Energy-Topology DRAND (EB-ET-DRAND). This algorithm is based on both the E-T factor and Lamport's bakery algorithm. It uses exponential backoff to adjust slot allocation priorities. These features allow the algorithm to reduce the collision probability of control messages and message complexity compared to the DRAND algorithm. However, despite these improvements, it possesses all other shortcomings of the DRAND algorithm. On top of these drawbacks, one control slot is assigned to every data slot within the proposed frame of the EB-ET-DRAND algorithm, which results in a large portion of the spectrum allocated for the exchange of control messages. The size of the proposed frame depends on network size, with EB-ET-DRAND assuming that the upper limit of network nodes is 32 and the upper limit of neighbor nodes for each node is 8, which does not provide a scalable solution for larger networks.

In [20] a novel graph coloring technique is presented, called the Color Constraint Heuristic (CCH). Based on CCH, both Centralized Slot Assignment CCH

(CSA-CCH) and Distributed Slot Assignment CCH (DSA-CCH) algorithms are proposed. Compared to DRAND, DSA-CCH requires fewer colors or TDMA slots, fewer control message collisions are expected, but the execution time of the algorithm is longer. DSA-CCH is not scalable as it uses a spreading approach instead of parallel algorithm execution. Execution of the algorithm at the node level is sequential; it starts at the sink/gateway and propagates through the network. Thus, the convergence of the algorithm for large network deployments becomes slow.

DRAND algorithm and its ameliorations are single-channel TDMA scheduling algorithms that store and update a superframe table per node and exchange control messages between network nodes. Ameliorated DRAND algorithms solve some of the shortcomings of the basic DRAND algorithm, but high control overhead and algorithm execution times are still present, especially in the case of large-scale networks. As such, they are more suitable for wireless networks where network topology does not change for a long period of time. The DRAND algorithm, its derivatives, and the FPRP algorithm use a greedy graph coloring approach, which is inherently sequential. This leads to inefficient and slower algorithm execution compared to synchronous algorithms. Besides, these algorithms require two-hop neighbor information, which increases execution time, decelerates scheduling convergence, and increases the number of control messages. They also try to reuse slots as much as possible, but without achieving optimal spectrum utilization.

Another example of a distributed single-channel TDMA scheme that stores slot usage information per node is disclosed in [21]. The proposed algorithm referred to as DICSA, stores a list of forbidden slots per node and enables nodes to participate concurrently in slot reservation procedures. Both aspects contribute to a more efficient slot allocation procedure. However, the proposed method does not provide a solution for the exposed node problem, leading to suboptimal utilization of the available spectrum. Also, similar to DRAND and its ameliorations, DICSA requires an affirmative response from all the neighboring nodes to reserve a requested slot. Therefore, it requires a high number of control messages for the execution of a slot allocation, resulting in scalability issues in the case of large-scale multi-hop networks.

In [22], a distributed node scheduling algorithm for multi-hop wireless networks, called Local Voting, is proposed. It is shown that the proposed algorithm achieves better performance than the other distributed algorithms in terms of average delay, maximum delay, and fairness. The Local Voting algorithm consists of two functions: requesting and releasing free time slots, and load balancing. If a node has no traffic, all its time slots are released. If a node requires new slots, it examines all the slots sequentially and the first available slots are allocated. If all slots have been allocated to one-hop or two-hop neighbors of the examined node, then no new slot is allocated to the node. This conservative approach leads to the occurrence of exposed node problem. The load balancing function is invoked to keep the load balanced between the nodes. The message exchanges for requesting and releasing slots are considered equivalent to message exchanges in the DRAND algorithm. Furthermore, same as the DRAND algorithm, the Local Voting algorithm requires the exchange of two-hop neighbor information, which leads to higher control overhead than algorithms based on one-hop neighbor information dissemination.

All algorithms presented so far are single-channel TDMA scheduling algorithms. From the respective papers, it is not clear if these algorithms can be extended to cover multi-channel environments. Consequently, if such enhancement were possible, it is unknown to which performance degradation it would lead in these more complex conditions. In the IEEE 802.15.4 standard, multi-channel communication is applied with nodes being able to switch channels quickly, or more specifically, to perform channel hoping. Sixteen non-overlapping channels are used in the 2.4 GHz band and 10 channels in the 915 MHz band [23]. This helps to avoid external interference and allows multiple simultaneous transmissions on different channels in order to increase network throughput. In recent years, the IEEE 802.15.4e standard [24], an amendment of IEEE 802.15.4 standard, proposes a redesign of the MAC layer in order to provide a framework for schedule-based communication for Wireless Personal Area Networks (WPANs) in multi-channel environments. Anyway, in [24] there is no guidance on how to assign time-frequency slots to each link, only how to apply it.

The Internet Engineering Task Force (IETF) 6TiSCH working group integrated the IPv6-based upper protocol stack with IEEE 802.15.4e standard [25] and introduced the 6TiSCH operation sublayer (6top) [26], which implements and terminates the 6top Protocol (6P) [27] and runs one or more 6top Scheduling Functions (SFs). 6P allows a node to communicate with neighboring nodes in order to add/remove cells, whereas the SFs define the rules when to add/remove cells, monitor performance, and collect statistics. The 6TiSCH working group has introduced the Scheduling Function Zero (SF0) [28], which is later replaced by Experimental Scheduling Function (SFX) [29]. Recently, 6tops is enhanced with the Minimal Scheduling Function (MSF) [30]. Proposed SFs aim to provide a minimal set of scheduling functionalities to be usable in a wide range of applications. However, provided functionalities are basic and based on restrictive assumptions. For example, MSF is optimized for applications with regular upstream traffic from the nodes to the sink, whereas downward traffic is considered as sporadic and its management is not specified. We further differ from 6TiSCH technology by employing multi-hop MF-TDMA scheduling without random channel hopping. Any scheduling algorithm, different from those proposed by the 6TiSCH working group, can be implemented in the 6top. A detailed literature overview of scheduling algorithms in IEEE 802.15.4e can be found in [31].

Distributed multi-hop scheduling algorithm GALLOP [32] is one example of distributed multi-hop scheduling algorithm that can be integrated with IEEE 802.15.4e standard. It has been designed to address the challenges of multi-hop closed-loop control in wireless networks with a tree topology. GALLOP supports bi-directional scheduling for cyclic information exchange and it is based on low-overhead signaling that leads to scalable operation. GALLOP also supports frequency/channel hopping for mitigating external interference. Execution of algorithm is sequential; uplink scheduling is going through ranks in tree topology from leaf nodes to sink node, whereas downlink scheduling starts from sink node and propagates through ranks of tree topology to leaf nodes. However, uplink and downlink scheduling in the GALLOP algorithm does not offer protection against exposed node problem.

A multi-channel Dynamic TDMA Slot Reservation (DTSR) protocol is proposed in [33]. It represents an example of a protocol based on Cognitive Radio (CR), where it is assumed that every node is equipped with a reconfigurable transceiver and spectrum scanner. This protocol divides the spectrum into time-frequency pairs and proposes a superframe, which in addition to data and control slots, incorporates sensing periods. During the sensing periods, nodes listen for transmissions of other nodes and acquire knowledge of free slots. DTSR requires constant traffic, as otherwise, nodes may reuse slots allocated by idle nodes, resulting in later internal collisions. Control slots are used for the exchange of control messages and every node keeps track of slot usage by its neighbors. In the case of collided control messages, DTSR does not provide a protection mechanism and it may result in superframe tables with obsolete slot usage information. Utilization of available spectrum by DTSR protocol is suboptimal as it does not protect against exposed node problem.

DSAT-MAC [34] is a multi-channel distributed TDMA protocol where slots are dynamically distributed between CR users. This protocol allows CR users to opportunistically use unused licensed spectrum without harming the primary users that own the licensed spectrum. Same as for DTSR protocol, a prerequisite for DSAT-MAC is the spectrum-monitoring capability of every node in the network. In DSAT-MAC one control slot per node is assigned, leading to inefficient spectrum utilization, especially for large-scale networks. To acknowledge its existence, every node needs to transmit at least one control message during every single superframe, leading to an increase in control overhead. Furthermore, the execution time of protocol is high and multiple specific wireless problems may occur as hidden, exposed, and deaf node problems. Deaf node problem occurs when the transmitter and receiver are tuned on different frequencies [35].

Self-Organizing Medium Access Control for Sensor Networks (SMACS) is another multi-channel distributed protocol. It is a hybrid TDMA protocol proposed for WSNs [36]. The main advantage of this protocol is that it does not require the synchronization of nodes to global time. However, the proposed protocol requires many frequency channels in the available radio spectrum, assigning randomly different frequency channels to different nodes. If there are only a few radio frequency bands and multiple nodes are assigned to operate on the same frequency bands, the possibility of overlapping slots and internal collisions is high. On another hand, due to lack of synchronization, SMACS offers non-optimal spectrum utilization with unused spectrum left between allocated slots.

Whilst the majority of distributed TDMA scheduling algorithms solve the hidden node problem, the exposed node problem and spatial spectrum reuse are not the focus of these algorithms, resulting in suboptimal spectrum utilization. Spatial TDMA (STDMA) presented in [37] is one algorithm that takes into account the exposed node problem and the fact that network nodes are usually spread out geographically. In case of sufficient spatial separation, STDMA enables nodes to reuse the same slots within the TDMA schedule. In this way, both the capacity and spectrum utilization of multi-hop networks are increased. However, STDMA is a single-channel TDMA scheduling algorithm for networks with known topologies and with a fixed number of static nodes, thus not offering flexibility to cope with the network dynamics as observed in many network deployments.

Table 4.1: Qualitative comparison of different scheduling algorithms/protocols.

Algorithm/protocol	Multi-channel scheduling	Hidden node avoidance	Exposed node avoidance	Supported wireless networks/topologies	Evaluation method	Sequential or synchronous execution
FPRP [15]	No	Yes	No	Static topologies	Simulation	Sequential
DRAND family [9, 16, 17, 19, 20]	No	Yes	No	Network topologies that do not change for a long period of time	Analytical [9, 17] Simulation [all] Experimental [17]	Sequential
DICSA [21]	No	Yes	No	WSNs	Simulation	Synchronous
Local Voting [22]	No	Yes	No	Multi-hop networks with static nodes	Analytical Simulation	Sequential
SFs [28–30]	Yes	Yes	No	Multi-hop networks	None	Synchronous
GALLOP [32]	Yes	Yes	No	Tree topology	Analytical Simulation Experimental	Sequential
DTSR [33]	Yes	Yes	No	Ad hoc networks with CR nodes	Simulation	Synchronous
DSAT-MAC [34]	Yes	No	No	Ad hoc networks with CR nodes	Simulation	Synchronous
SMACS [36]	Yes	Yes	No	WSNs	Simulation	Synchronous
STDMA [37]	No	Yes	Yes	Known topologies with a fixed number of static nodes	Analytical Simulation	Sequential
DDMC-TDMA	Yes	Yes	Yes	Topology-agnostic	Simulation[1]	Synchronous

Based on the literature analysis, all existing distributed TDMA algorithms possess certain shortcomings, mainly focusing on solving a limited set of problems arising in ad hoc networks and generally disregarding scalability and spectrum efficiency. For this reason, we propose a new protocol targeting multiple dynamic use cases and solving most of the known problems occurring in distributed TDMA scheduling for ad hoc networks. Key differences between the existing state-of-the-art and proposed algorithm are shown in Table 4.1. As it can be seen, our proposed solution is the only one that is topology-agnostic, while being highly

[1]Previously experimentally evaluated in DARPA SC2 competition [13]

scalable, solving exposed/hidden node problems, and supporting multi-channel scheduling. It is worth noting that execution times of algorithms could not be compared directly, therefore in the last column of Table 4.1 we present if algorithms are sequential or synchronous. In general, sequential algorithms are converging more slowly to the steady state than synchronous algorithms.

4.3 Design goals and architecture of DDMC-TDMA

First, the main design goals of the proposed DDMC-TDMA scheduling approach are presented. Afterward, DDMC-TDMA architecture and its main features and procedures are described in more detail.

During the design and development of DDMC-TDMA, besides satisfying requirements for SC2, the main goal was to provide a generic solution for TDMA scheduling that would maximize spectrum efficiency while minimizing the control overhead. As centralized solutions are not scalable and cannot cope with dynamic network and traffic conditions, it was decided to go for a distributed scheduling solution. By opting for a distributed solution, single point-of-failures, like in centralized solutions, can be avoided. Furthermore, in a distributed solution, control messages are only exchanged between neighboring nodes, limiting control overhead for spectrum management. This leads to improved performance in terms of overall system capacity and individual node throughput and latency. It was decided to further reduce control overhead by minimizing the number of control messages exchanged during a slot allocation procedure and by limiting the number of nodes participating in the procedure to only the two nodes that want to establish a link.

In addition to a slot allocation procedure, also a dynamic slot removal procedure is required to maximize coexistence with other unknown wireless networks (such as competing networks in the DARPA SC2 competition). We hereby assume that future wireless networks do not have exclusive access to the spectrum, but have to share the spectrum with other unknown networks. To mitigate interference, removal and reallocation of affected slots may be needed. Moreover, traffic demands are often varying over time and it is not spectrum efficient to keep all allocated slots when traffic demands drop.

As existing distributed solutions are mostly focused on specific use cases with limited applicability elsewhere, another goal was to create a generic solution that covers a wide range of application scenarios and network topologies. The slot management protocol should support both infrastructure-based and infrastructure-less wireless networks, static as well as mobile nodes. To maximize the applicability of the protocol, it should support both single-hop and multi-hop networks operating either in a single-frequency band or multi-channel environment. The proposed MAC should also be Physical layer (PHY) agnostic, meaning that it can be transparently integrated on top of any existing or future wireless PHY standard.

Another design goal was to design a protocol that maximizes throughput and reliability performance by solving the typical hidden and exposed node problems that may occur in various wireless network topologies. In other words, the protocol should guarantee full spectrum utilization with d when traffic load is high and with no overlapping slots allocated.

A Dynamic Distributed Multi-Channel TDMA (DDMC-TDMA) slot management protocol has been developed in line with the abovementioned design goals. In this chapter, DDMC-TDMA is presented and described as a scheduling solution for ad hoc networks. However, as already pointed out, the proposed protocol is also applicable for infrastructure-based networks without any loss of generality. DDMC-TDMA can be applied to both single-channel TDMA schemes and Frequency Division Multiple Access (FDMA) schemes. If not indicated otherwise, in the remainder of this work DDMC-TDMA is referring to MF-TDMA access schemes, as MF-TDMA represents the most complex case of TDMA-based spectrum sharing.

The main focus of this chapter is on the MAC layer and more specifically on the TDMA slot management protocol, its applicability and scalability in different use cases. For this reason and since DDMC-TDMA has already been evaluated experimentally during SC2 competition with realistic PHY and wireless medium (see Chapter 3), in this chapter we decided to adopt a simplified PHY with symmetric links between two nodes. We further do not take into account propagation delay and transmission error models.

For TDMA access schemes to be feasible, all nodes' clocks need to be synchronized. Based on observations from the real-time SCATTER system developed for SC2, it is determined that submillisecond (sub-ms) accuracy is sufficient and this can easily be maintained with traditional synchronization methods. Therefore, it is assumed that node clocks are synchronized with this accuracy. Guard spaces are introduced to TDMA slots, which guarantees stable operation of the system as long as the clock drifts between the nodes are below 1 ms. Every node of the ad hoc network has one or more transceivers, which may work in Frequency Division Duplexing (FDD) full-duplex or half-duplex mode. Thus, DDMC-TDMA supports distributed scheduling for networks consisting of single-radio and/or multi-radio devices.

The distributed slot management protocol is executed on all nodes of the ad hoc network. For local execution of DDMC-TDMA, a node does not need any knowledge regarding its position in network topology or the network size. Dynamic slot allocation/removal procedures of a node and access of a node to the shared medium are managed by an entity that is schematically presented in Fig. 4.1. Every node of the ad hoc network is running its own instance of DDMC-TDMA entity, allowing the node to establish communication with other nodes of the network, and therefore managing spectrum utilization in a distributed way.

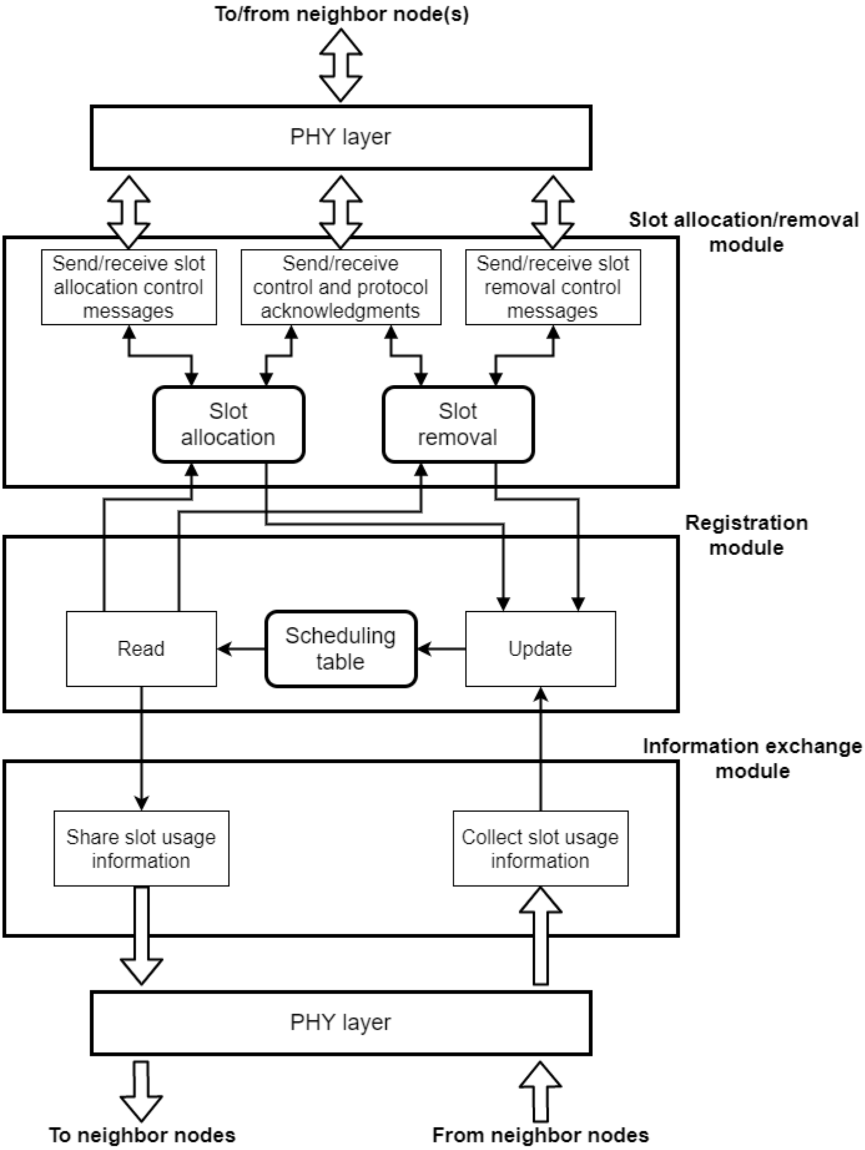


Figure 4.1: Local DDMC-TDMA entity running on each node for a distributed slot management.

The DDMC-TDMA entity for distributed slot allocation/removal consists of three modules:

1. A registration module, maintaining a scheduling table of the slots, that represents time slots and/or frequency channels.

2. An information exchange module, (i) collecting slot usage information indicative for the use of slots by neighbor nodes, (ii) updating the scheduling table based on the received slot usage information, and (iii) sharing its own local slot usage information with neighbor nodes.

3. A slot allocation/removal module, allocating or removing a slot used for communication with a neighbor node. During the allocation/removal procedures, it consults the scheduling table and exchanges control messages with the neighbor node. In addition, it updates the scheduling table based on the outcome of the executed slot allocation/removal procedures.

In the following sections, the modules of the DDMC-TDMA entity are described in more detail.

4.3.1 Registration module

The registration module consists of a scheduling table and primitives for accessing the scheduling table. Other modules of the DDMC-TDMA entity can retrieve information from the scheduling table using the 'Read' primitive or can change the content of the table using the 'Update' primitive. The registration module is responsible for storing the most recent information of spectrum utilization as seen from a single node perspective.

4.3.1.1 Scheduling table

The scheduling table can be one-dimensional or two-dimensional. For single-channel TDMA schemes, the scheduling table is one-dimensional and only consists of time slots in a single channel. An FDMA scheme is another example of a one-dimensional table, where the radio spectrum is divided into multiple channels and no time division is applied. If both time and frequency division is enabled, the scheduling table is two-dimensional and represents an MF-TDMA medium access scheme. An example of an MF-TDMA scheduling table kept per node is presented in Fig. 4.2.

Information stored in the scheduling table specifies how the slots are allocated for communication within the ad hoc network. It keeps the information about slots assigned for its own communication needs, either for transmission (blue slots) or reception (red slots), and slots assigned for communication of neighbor nodes (grey slots). Some slots (yellow slots) are reserved for exchanging control messages. Only one-hop slot usage information is needed for maintaining the table, as is explained further in this chapter. Scheduling tables are kept locally for each node, hence each node has its local view of the network, and scheduling tables in the same ad hoc network differ from node to node. The scheduling table is updated with every successful slot allocation/removal procedure of the node itself

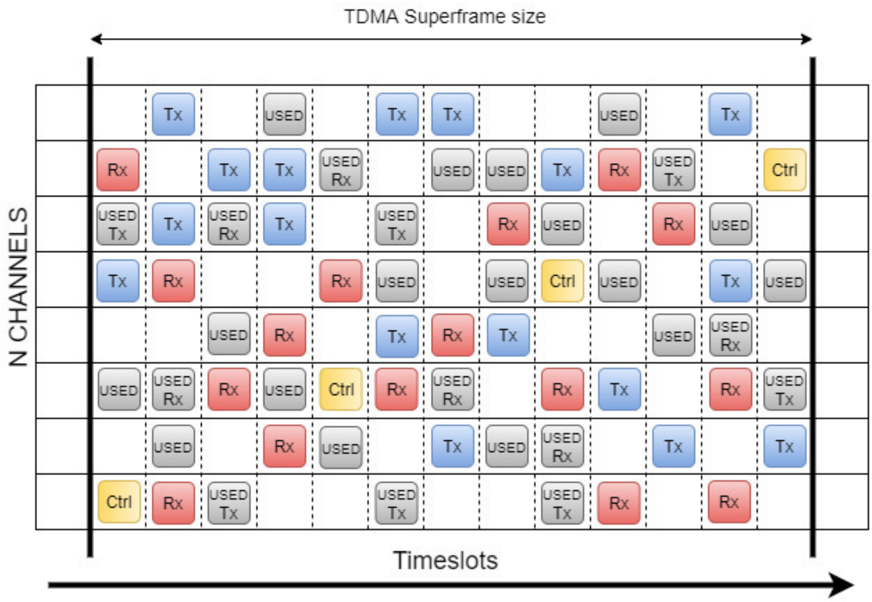


Figure 4.2: Example of a scheduling table stored per node.

or any node in its vicinity. It is also updated by collecting and parsing periodic *slot usage* messages from neighbor nodes. The allocation/removal module and the information exchange module inform the registration module about any recent changes of slot usage via the 'Update' primitive of the registration module. During slot allocation/removal procedures, the slot allocation/removal module consults the scheduling table for the usage information of time-frequency slots using the 'Read' primitive of the registration module. The same primitive is used by the information exchange module for periodic reporting of the scheduling table status to neighbor nodes.

4.3.1.2 Types of time-frequency slots

As shown in Fig. 4.2, every time-frequency slot is assigned an appropriate slot type, based on the utilization of slots by the node itself or neighbor nodes. Empty cells in the scheduling table represent unused time-frequency slots, in the remainder of the chapter referred to as *Empty* slots. These slots represent time-frequency slots that are not allocated by the node itself or by any node in its vicinity. Thus, these slots are free for future use in case that a new slot needs to be allocated for communication with any of the neighbor nodes.

Used slots are divided into two classes, internally and externally used slots. Internally used slots are slots allocated for communication links belonging to the specific node with its neighbors. Three types of internally used slots exist:

1. Transmission (*Tx*) slots
2. Reception (*Rx*) slots
3. Control (*Ctrl*) slots

The scheduling table also stores information about slot usage from one-hop neighbors as externally used slots. Information about the usage of slots by neighbor nodes is either received during execution of distributed slot management procedures or by collecting periodical messages containing slot usage information and then is stored in the scheduling table as one of three following slot types:

1. *USED Tx* slots
2. *USED Rx* slots
3. *USED* slots

For a more comprehensive explanation of different slot types stored within scheduling tables, let us analyze Fig. 4.3. The left section of Fig. 4.3 represents an arbitrary wireless network topology with initially established communication links between the nodes. The superframe state is presented in the upper right section of Fig. 4.3. The superframe comprises of 5 time slots 1, 2, 3, 4, and 5 in a single-channel TDMA scheme with corresponding transmitter-receiver pairs in slots 1-4, corresponding to the four established links. Time slot 5 is used by all the nodes of the network for control messages. Nodes A and D and their scheduling tables are selected as explanatory examples of previously defined slot types. Therefore, collision domains of nodes A and D are illustrated in the network topology.

First of all, from the perspective of node D, time slots 2 and 3 are Empty, as all the nodes communicating within these slots are not neighbors of node D. Tx slots are slots that the node uses for transmission to a neighboring node, whereas Rx slots are allocated for the reception of transmissions from a neighboring node. In the scheduling table of node A, time slot 1 is marked as Tx since node A is using this slot itself for transmission to node B. Node D is using time slot 4 itself for reception from node E, so this slot is marked as Rx in its scheduling table. Slots, marked in the scheduling table as Ctrl, are reserved for the exchange of control messages during distributed slot management protocol procedures. Time slot 5 is a common Ctrl slot in scheduling tables of nodes A and D, as well as all other nodes in the network.

Slots marked as USED Tx are allocated by a neighbor node for transmission, but none of the neighbor nodes is using them for reception. The receiver node in this case is out of range of the node in consideration. There are two slots stored

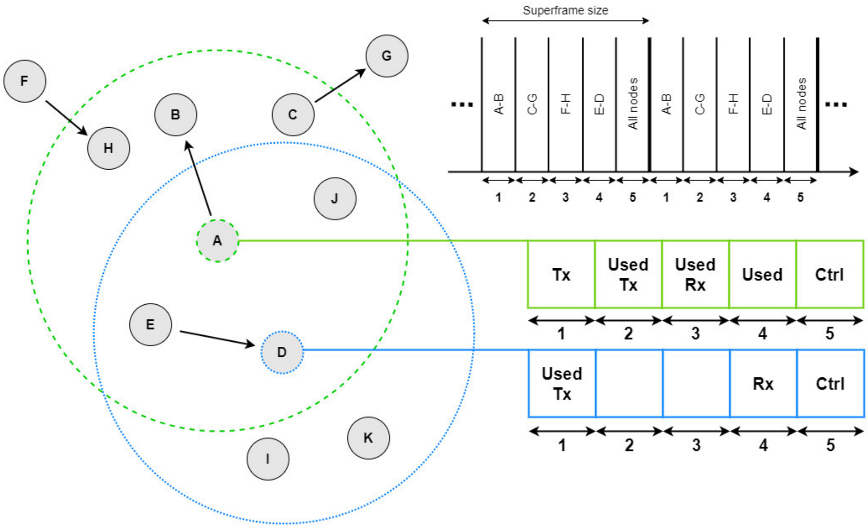


Figure 4.3: Arbitrary network topology, superframe state and scheduling tables of the network nodes.

as USED Tx in the scheduling tables of nodes A and D. First is time slot 2 in the scheduling table of node A, as this time slot is used for the transmission from C to G, where only node C is a neighbor of node A. As node G is not a neighbor of node A, none of the slot usage information collected by node A from nodes in its vicinity will indicate this slot being used for reception. The second slot marked as USED Tx is time slot 1 in the scheduling table maintained by the registration module of node D. Node D is only aware of the transmitter node A utilizing the slot, whereas the receiver node B is not in the vicinity of node D. USED Rx slots represent slots that are allocated by a neighbor node for reception, but the transmitter is not a neighbor of the node in consideration. Node A is storing time slot 3 as USED Rx because this time slot is used for the transmission from F to H, where the receiver node is a neighbor of node A, whereas the transmitter node is outside the range of node A. Therefore, node A cannot receive slot usage messages transmitted in a control slot by node F.

If neighboring nodes are using a time-frequency slot for both transmission and reception, that slot is labeled as USED slot. The simplest example of USED slot is a slot in which two neighbor nodes have established a communication link. This example we can see in the scheduling table of node A. Time slot 4 is marked as USED because this time slot is allocated for transmission from node E to node D and both nodes E and D are neighbors of node A. A slot is also considered as USED if one neighbor node transmits to a node out of range of node in consideration and another neighbor node is receiving at the same slot from the node out of range of node in consideration. USED Tx and USED Rx slots may transform to USED slots if subsequently reported by any neighbor node as being used for reception and transmission, respectively.

4.3.2 Slot allocation/removal module

The slot allocation/removal module supports distributed slot management on a node level with dynamic establishment and removal of communication links between any of the neighboring nodes in the ad hoc network. This module consists of two main units, the slot allocation unit and the slot removal unit.

The slot allocation unit allocates a time-frequency slot for communication with a neighbor node by executing a slot allocation procedure. The slot removal unit releases a slot allocated for communication with the neighboring node by executing a slot removal procedure. Both procedures are based on the exchange of specific control messages with the neighboring nodes. The detailed procedure for allocating or removing slots is explained in the following sections.

Slot allocation/removal operations are event-based, reacting to a difference in the required capacity of the link compared to existing link capacity and reacting to detected interference. A slot allocation operation is triggered if the required capacity is higher than the existing capacity and it results in an increase in allocated bandwidth if there is available spectrum left. A slot removal operation is triggered if the required capacity is lower than the existing capacity and it results in a decrease in allocated bandwidth. In both cases, any ongoing communication is unaffected, with enough spectrum capacity provided by the scheduling algorithm to satisfy application demands. In case of detected interference, allocated bandwidth stays the same, with interfered slots reallocated to non-interfered parts of the spectrum. As an inevitable consequence of interference impact, application flows utilizing interfered slots may experience a drop in throughput, until reallocation of slots is executed. Slot allocation/removal operations do not have any negative impact on application flows themselves, with only increasing or decreasing utilized bandwidth in order to incorporate all application traffic while minimizing spectrum footprint.

4.3.2.1 Slot allocation procedure

The slot allocation unit allocates a slot for communication with a neighboring node. During a slot allocation procedure, both the initiating node and neighbor node are involved. To agree on a slot to be allocated, their slot allocation units exchange slot allocation control messages over control slots. The procedure is initiated by the transmitter node upon demand for increasing the application throughput. The slot allocation unit of this node consults its scheduling table using the 'Read' primitive of its registration module in order to select available slots that can be proposed for establishing a communication link with the receiving neighbor node. The receiver node's slot allocation unit consults its respective scheduling table to determine which slot, from the list of slots proposed by the transmitter, may be selected for reception. Once both nodes agree on a slot, the communication link is established and the scheduling tables are updated using the 'Update' primitive of registration modules. The type of the allocated slot is Tx at the transmitter side and Rx at the receiver side. In this way, each node keeps track of the slots that have been allocated for active communication links with neighboring nodes.

Fig. 4.4 illustrates the flow diagram of a slot allocation procedure and its control messages that are exchanged between the medium access entities of two neighbor nodes A and B.

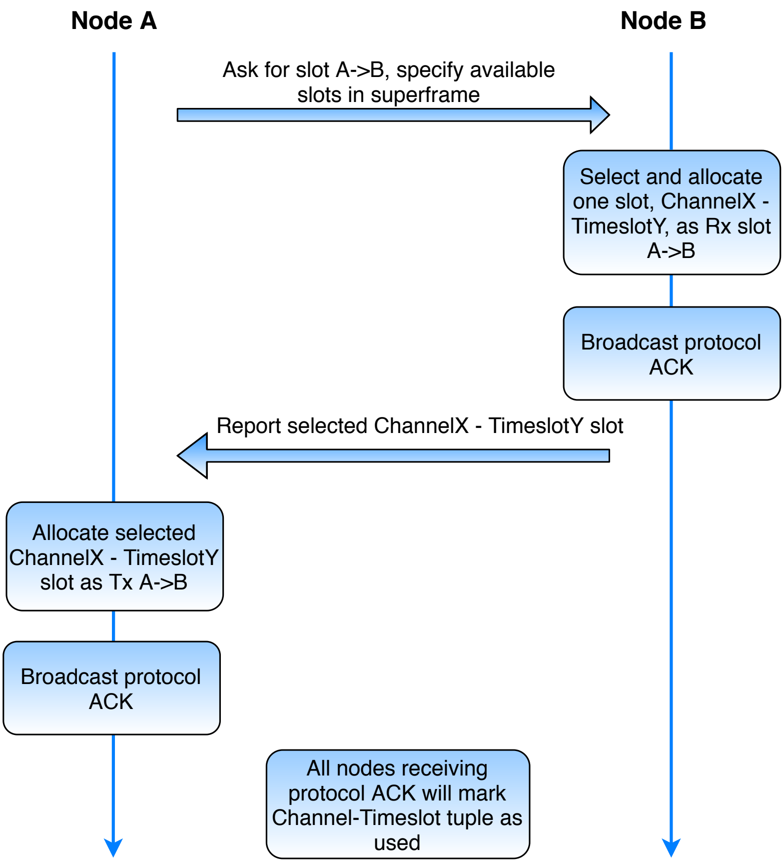


Figure 4.4: Flow diagram of a DDMC-TDMA slot allocation procedure.

If node A has a frame(s) to send to node B and there are not enough allocated slots for transmission to B, node A should initiate a slot allocation procedure in order to satisfy application data rate demands. To get an overview of the slots available to be allocated as transmission slots, node A consults its scheduling table. A list of available slots are filtered based on the following rules:

- Empty slots in the scheduling table are not used by the node itself or any of the neighbor nodes and can all be proposed for a new wireless connection.
- USED Tx slots are allocated by one neighbor node of node A for transmission, but no other neighbor node of node A is using this slot for reception. This implies that the receiver for this particular communication link is outside the neighborhood of node A and will not be affected by transmissions of node A. For this reason, such a slot may be a part of the proposed slots list.
- All other slot types cannot be proposed.

Next, node A selects one or more slots from the list of available slots. These slots are proposed to node B through a *proposed slots* control message. The maximum number of proposed slots that can be embedded in the proposed slots message depends on the transport block size supported by the PHY.

Upon reception of the proposed slots message, node B consults its scheduling table and selects one slot, further referred to as the allocated slot, from the list of the proposed slots. For the slot selection, node B applies the following rules:

- Empty slots are not used by node B or any of the neighbor nodes and are therefore available for selection.
- USED Rx slots are allocated by one neighbor node of node B for reception, but no other neighbor node of node B has allocated it for transmission. This implies that the transmitter for this particular communication link is outside the neighborhood of node B. Therefore, a USED Rx slot is available for reception by node B, as there is no risk for interference by other transmitters using the same slot.
- Other slot types are considered as unavailable for selection by the receiver node.

After selecting one slot from the filtered list, node B updates its scheduling table, so that the corresponding time-frequency tuple is marked as Rx. Node B reports the allocated slot by transmitting a *selected slot* control message to node A. Upon reception of the selected slot message, node A updates its scheduling table accordingly by marking the newly allocated slot as Tx. At the end of a slot allocation procedure, both nodes A and B have allocated the same slot and their scheduling tables are updated and aligned. As a result, they have successfully set up a new communication link or have increased the capacity of an already existing link.

The final step of a slot allocation procedure is to notify all neighbor nodes about the allocated time-frequency slot. Neighbor nodes are notified by control messages called *protocol acknowledgment* (ACK), broadcasted by the two nodes participating in the slot allocation. Protocol ACK messages are carrying information related to the outcome of the executed slot allocation procedure and upon reception of such a message, other neighbor nodes can update their scheduling tables accordingly.

By following the described steps for the proposal and selection of adequate slots, DDMC-TDMA inherently avoids the hidden node problem. On the other hand, storing the USED Tx and USED Rx slot types in the scheduling table and distinguishing them from the USED slot type, prevents these types of slots from being unduly considered as unavailable for allocation. Exploiting these slot types during a slot allocation procedure enables DDMC-TDMA to solve the exposed node problem.

Other general rules may apply for slot selection narrowing down the number of slots that may be selected, such as:

- The number of simultaneous transmit-receive pairs using the same time slot is limited by the number of transceivers supported by the node and the capability to support the full-duplex operation.

- If only one transceiver is supported by the node, the same time slot of a Ctrl slot cannot be allocated for other transmit-receive pairs.

If there are multiple appropriate slots for slot allocation, different strategies can be applied for slot selection:

- The most simple strategy is when the receiver node randomly selects one slot from the list of proposed slots.

- The receiver node can select a slot based on its position in the scheduling table, for example by first filling unused frequencies within a certain time slot.

- The receiver node can select a slot based on the weight assigned to the slot. An example of such a weight parameter could be the presence of external interference measured by spectrum monitoring.

- In order to optimize spectrum utilization, the strategy can be driven by prioritizing already used slots, marked as USED Tx and USED Rx in scheduling tables. This strategy is highly recommended when multiple wireless networks have to share the same spectrum and it is further applied in this work.

4.3.2.2 Slot removal procedure

By releasing allocated slots that are underperforming or unused, medium access control can dynamically respond to changes in application layer demands or changes in the network that affect the reliability of established links. Both transmitter and receiver node participating in a link can initiate a slot removal procedure, where an established communication link may consist of one or more time-frequency slots. The reliability and efficiency of every allocated slot are monitored over consecutive superframes. If the performance of a specific slot is not satisfying the application requirements for a predefined number of superframes, then this unreliable slot needs to be released, i.e. making it available again for another allocation by other nodes. Another trigger for initiating a slot removal procedure is when a slot is idle for a while (idle period) due to decreased application demands or because an application has ended.

The receiver node monitors the presence of network traffic within allocated Rx slots, whereas the transmitter node monitors the quality in terms of successfully delivered frames in its allocated Tx slots. If the Packet Error Rate (PER) value for any Tx slot is above a predefined threshold (PER removal initiation threshold) for a predefined number of consecutive superframes (poor quality period), the transmitter node initiates a slot removal procedure for that slot. A high PER could be caused by interference from an unknown external network. Slots might also suffer from internal interference, which can be caused by collisions between control messages or by the distributed nature of the slot management protocol. By removing unreliable slots, both external or internal collisions may be resolved.

Fig. 4.5 illustrates the flow diagram of a slot removal procedure executed between two neighbor nodes A and B that have established a link for communication. The flow diagram of a slot removal procedure is the same regardless of the initiator being the transmitter or the receiver. Slot removal units of nodes A and B will remove the previously allocated slot for communication with a neighbor node by consulting their scheduling tables, detecting unreliable or idle slots, and exchanging slot removal control messages.

After node A consults its scheduling table and detects a slot that is unreliable or idle for a predefined number of consecutive superframes, it releases the slot and notifies node B participating in communication over the detected slot. By utilizing shared control slots, node A transmits *remove slot* message to node B, advertising the time-frequency slot to be removed. Upon reception of the remove slot message, node B also releases the underperforming slot. After the successful release of the slot, the corresponding slot type Rx information in the scheduling table of the receiving node is deleted. Similarly, the slot type Tx information in the scheduling table of the transmitting node is deleted. As such, after the execution of the removal procedure, the previously allocated slot is transformed into an Empty slot in the scheduling tables of nodes A and B. Both nodes A and B broadcast proto-

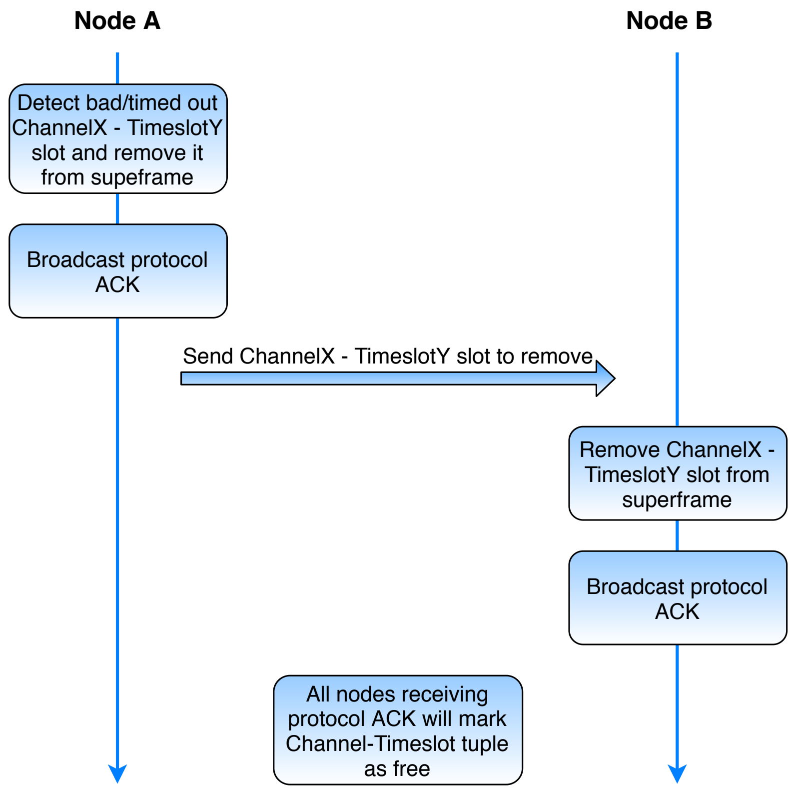


Figure 4.5: Flow diagram of a DDMC-TDMA slot removal procedure.

col ACK message with information about the slot removed, allowing all neighbor nodes of nodes A and B that can overhear broadcasted protocol ACKs to update their scheduling tables accordingly.

4.3.3 Auxiliary control messages in slot allocation/removal procedures

In addition to primary control messages for slot allocation/removal procedures, slot allocation and slot removal units of MAC entity for distributed slot management use two auxiliary types of control messages. Those control messages are *control ACKs* and protocol ACKs. Control ACKs are exchanged between the nodes participating in a slot allocation/removal procedure to detect if primary control messages are interfered and should be retransmitted. As previously explained, protocol ACKs are used to notify neighboring nodes about the outcome of executed procedures.

4.3.3.1 Control ACKs

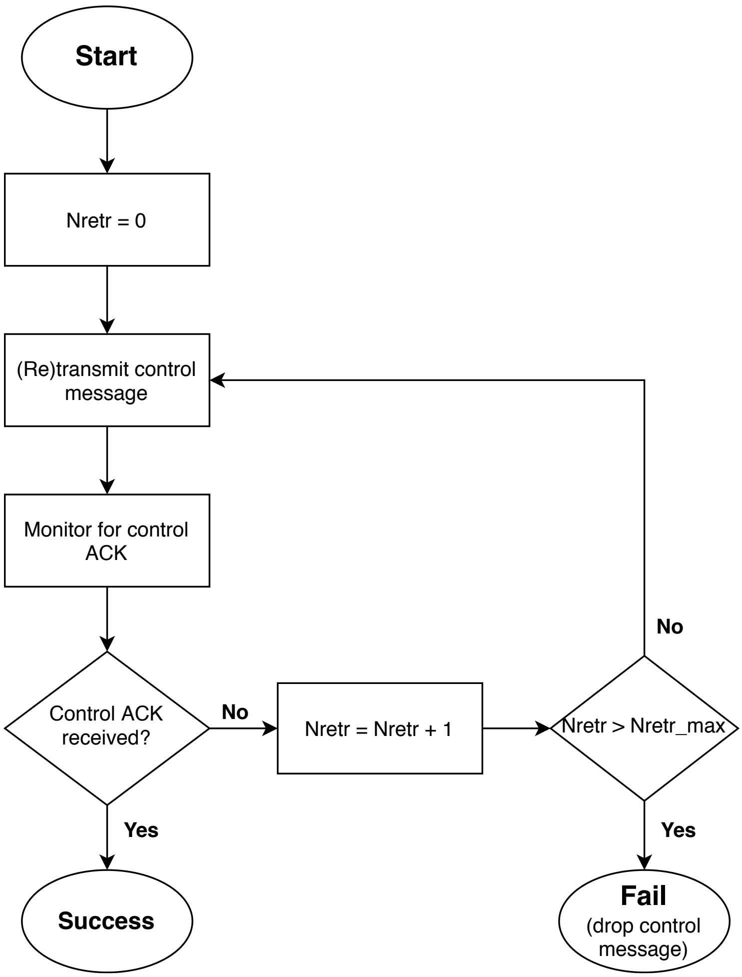


Figure 4.6: Flow diagram of a DDMC-TDMA control message transmission procedure.

Control ACKs have the same purpose for control messages as data ACKs have for data messages. Usage of control ACKs increases the reliability of the control messages exchanged during the execution of the protocol's procedures. Just like other control messages, control ACKs are transmitted in the reserved control slots. Every successfully received unicast control message is followed by the transmission of a control ACK message from the involved receiver node. If a unicast control message is not successfully acknowledged, then the transmitter will retransmit the control message, increase its retransmission counter *Nretr* and continue monitoring for a corresponding control ACK. The same procedure is repeated until a control ACK message is successfully received or the maximum number of control message retransmissions *Nretr_max* is reached, as illustrated in Fig. 4.6. The value for *Nretr_max* is predefined and may be different for different types of control messages. The introduction of control ACKs in the distributed slot management protocol increases the load in the control slots, but at the same time increases the reliability of control message transmissions and the success rate of slot allocation/removal procedures.

4.3.3.2 Protocol ACKs

Information about time-frequency resource usage of neighbor nodes is collected in two ways. The first way is based on the control messages exchanged by neighbor nodes during slot allocation/removal procedures and is event-based. Each time a slot allocation/removal is executed, each involved node broadcasts one special protocol ACK message, to inform other nodes in their communication range about the outcome of the procedure. A broadcasted protocol ACK consists of the time-frequency tuple and type of slot. In the case of slot allocation procedure, the protocol ACK includes the type of allocated slot, either Tx or Rx type of slot. For slot removal procedures, the protocol ACK includes information on the type of slot that is removed, either Tx or Rx type of slot.

By using protocol ACKs in the slot allocation/removal algorithm, neighboring nodes in the network are immediately notified about the most recent changes in slot usage, allowing them to update their scheduling tables with the latest usage information, which has a positive impact on the reliability and effectiveness of the slot management protocol. However, in ad hoc networks with mobile nodes, protocol ACKs may not be sufficient to cope with mobile nodes that were initially out of range and later enter the communication range. It may also happen that protocol ACK messages are lost due to interference, preventing the update of time-frequency resource usage information in the scheduling tables. Therefore, a second way to inform neighbor nodes about time-frequency resource usage is introduced. This approach is based on the information exchange modules periodically broadcasting scheduling tables of the network nodes, which is explained in more detail in the following section.

4.3.4 Information exchange module

The information exchange module is responsible for periodically sharing information on scheduling tables between nodes and is supplementary to the slot allocation/removal algorithm in case protocol ACKs are missed. Scheduling tables are embedded into slot usage messages exchanged between the information exchange modules of the MAC entities in order to support distributed slot management for more dynamic situations where the slot allocation/removal algorithm does not suffice. Based on the slot usage information collected by the 'Collect slot usage information' unit of the information exchange module, the scheduling table is updated. The information exchange module further comprises a sharing unit, sharing the content of the scheduling table with neighbor nodes. The 'Share slot usage information' unit accesses the scheduling table and retrieves information about time-frequency slots used by a node itself, either for transmission or for reception. It further embeds this information into a slot usage message and broadcasts it. This allows neighbor nodes to acquire knowledge about the slot usage of nodes in the one-hop neighborhood. If information about slot usage carried by a slot usage message differs from the status in the scheduling tables or if usage information is missing, the receiving nodes update their scheduling tables accordingly.

Slot usage messages are being broadcasted periodically, with a calculated period of scheduling table broadcasting (*Tsh*). In order to avoid multiple broadcasts of slot usage messages by different nodes simultaneously and to avoid the congestion of control slots, *Tsh* consists of a fixed scheduling table broadcasting delay (*Tshf*) and a random scheduling table broadcasting delay (*Tshr*) as given in Eq. 4.1.

$$Tsh = Tshf + Tshr \tag{4.1}$$

The fixed part introduces a minimum period between two consecutive slot usage messages transmitted by a node, whereas the random part introduces randomness in the time interval between slot usage transmissions.

Periodic broadcasting of slot usage messages ensures that all nodes in the network acquire an up-to-date view on occupied slots and slots available for allocation, in particular in dynamic network topologies with mobile nodes. Slot usage information collected from periodic broadcast reports complements slot usage information retrieved during event-based slot allocation/removal procedures. Such a combined approach leads to a more reliable slot management protocol that converges faster to a steady state.

4.3.5 Control slots

In DDMC-TDMA, a predefined number of control slots, marked as Ctrl in the scheduling tables, are reserved for exchanging control messages. These slots have the same time-frequency positions for every node of the ad hoc network. Control slots are used for broadcasting periodic slot usage information and for the execution of slot allocation/removal procedures. During the execution of slot allocation/removal procedures, both primary and auxiliary control messages are exchanged via control slots.

4.3.5.1 Allocation scheme of control slots

In DDMC-TDMA, initial control slots are allocated following a static scheme during boot time. The number of control slots depends on the total number of time-frequency slots present in the available radio spectrum; it depends on the channelization of the spectrum bandwidth and on the number of time slots in the superframe. The more time-frequency slots in the superframe, the more control slots are needed to offer sufficient opportunities for slot allocation/removal procedures to maintain reliable operation and to achieve fast convergence of the slot management protocol. However, a higher number of control slots also reduces the time-frequency resources available for the allocation of data slots, thus reducing the overall capacity of the network. This implies that fine-grained superframes (with many narrow frequency channels and small time slots) will be more spectrum efficient than course-grained superframes (with a limited number of channels and large time slots), as in the former case the ratio of control slots to data slots can be lower.

A minimal number of control slots, nicely distributed over the superframe, is required to minimize the impact of interference. The control slot allocation scheme is configured for the maximal distribution of the slots across time and frequency axes. If a limited number of control slots is interfered, other non-interfered control slots will still allow proper execution of slot allocation/removal procedures.

All nodes in the same ad hoc network adopt the same scheme of common control slots and update their scheduling tables accordingly. This allows neighboring nodes within the same ad hoc network to exchange control messages during the same time-frequency slots. Multiple nodes can concurrently execute slot allocation/removal procedures during the same control slot, limiting execution times and allowing them to react swiftly to changing and new traffic demands in the network.

4.3.5.2 Medium access scheme for control slots

To manage the access of multiple nodes to a common control slot, an appropriate medium access protocol is required. For example, random access protocols may be used like pure aloha, slotted aloha, or Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). In this research, we have adopted an enhanced slotted aloha protocol. Other medium access schemes may be more efficient, but optimizing the medium access scheme for control slots was not the focus of this research.

Slotted aloha requires time synchronization of the nodes. To support TDMA access, it is already assumed that nodes are synchronized with sub-ms accuracy. As it is shown in [38], to achieve maximum efficiency of slotted aloha protocol, nodes are also supposed to know the number of neighbor nodes which can be different in different parts of the network and can further change when mobile nodes are present in the network. In DDMC-TDMA, nodes acquire knowledge of their neighbor nodes by exchanging and parsing periodic slot usage messages and protocol ACK messages. Therefore, the slotted aloha protocol has been enhanced with a random slot selection that dynamically adapts to the number of neighbor nodes. This protocol enhancement is also beneficial for supporting mobility. Please note that the slot size employed in the slotted aloha scheme is a fraction of the TDMA slot size, as the slotted aloha scheme is executed within control TDMA slots.

4.3.5.3 Reliability enhancement of control slots

To prevent control slots from being saturated and becoming unreliable, the number of nodes that simultaneously execute slot allocation/removal procedures should be limited. To cope with this problem, slot allocation procedure timeout (*Talloc*) and slot allocation procedure delay (*Twait*) are introduced. *Talloc* is a fixed value, whereas *Twait* is a random value between predefined *Twait_min* and *Twait_max* values, i.e. *Twait* is calculated as:

$$Twait = \text{rand}(Twait\_min, Twait\_max) \tag{4.2}$$

where rand(a, b) is a function that returns a pseudo-random number in the range between a and b.

The concept of timeout *Talloc* and delay *Twait* is going to be explained in the example presented in Fig. 4.4. Upon the start of a slot allocation procedure, node A sets a timeout *Talloc* in which the procedure is expected to finish. If the procedure is executed successfully within this time frame, timeout *Talloc* is discarded and node A may initiate subsequent slot allocation/removal procedure after the *Twait* period. If the procedure is not finalized in time, this indicates a heavy load on control slots; control messages between nodes A and B are colliding with control messages from other network nodes. In this case, node A waits until timeout *Talloc* expires and consequently backs off from using control slots while they are used heavily by other nodes in the network. Besides, node A has to wait for the random delay *Twait* before starting the next slot allocation/removal procedure. Imposed randomness reduces the probability of the large number of nodes starting simultaneous slot allocation/removal procedures and increases the reliability of control message transmissions in control slots.

4.4 Validation and evaluation

To evaluate the correct operation and performance of the proposed DDMC-TDMA protocol, sets of simulation runs are executed for various ad hoc network configurations. Firstly, the basic functionality of slot allocation/removal procedures is evaluated in the case of a single-hop wireless network, operating both isolated and in presence of unknown external interference. Furthermore, the scalability of the DDMC-TDMA protocol is verified by increasing the number of devices in single-hop and multi-hop networks, as well as the network density of multi-hop networks. Network density represents the number of nodes in a single collision domain. Besides verifying that the protocol is scalable, the size and density of single-hop or multi-hop ad hoc networks are varied to prove that protocol is also able to achieve full spectrum utilization with an increasing number of network nodes. Multi-hop network topologies are simulated to prove the efficiency of protocol in the case of large networks covering multiple collision domains, with a focus on spectrum reuse. Finally, multi-hop topologies are analyzed to validate how the proposed protocol mitigates the hidden and exposed node problems occurring in such network topologies.

The DDMC-TDMA protocol is simulated and analyzed in the ns-3 simulator environment. The implementation of the DDMC-TDMA protocol in ns-3 is available online [39]. For every use case and its subset, 20 independent ns-3 simulations were conducted. The mean value of convergence times to steady state (*Mconv*) and confidence interval for an adopted confidence level of 95% were calculated for each set of simulation runs. Simulation runs with convergence times outside of calculated confidence interval $Mconv \pm 3.5\%$ were discarded and remaining results averaged. To avoid seed-related artifacts in simulation results, a random seed is generated from the simulation number and every node is booted at a random time at the start of the simulation. Please note that all the results presented in this section are analyzed in line with the start of the simulation and not after simulation warm-up time has passed.

Table 4.2: Default parameters used for simulations.

Simulation parameter	**Value**
Simulation duration	820 s
Superframe duration	1000 ms
Slot duration	50 ms
Slot's slice for data frames	43 ms
Slot's slice for ACKs	4 ms
Slot's slice for guard spaces	3 ms
Number of time slots in superframe	20
Number of data time slots in superframe	16
Number of control time slots in superframe	4
Maximum Tx/Rx slots allocated per node	16/16
Application data rates	400 and 800 packets/s
Maximum achievable frame rate	688 frames/s

Default parameter values for the executed simulations are presented in Table 4.2. It is worth noting that most of the simulation parameters are selected based on the implementation of the SCATTER system. For example, PHY-related parameters as transmission duration of a single frame and minimum single frame size are selected based on LTE-based SCATTER PHY characteristics [40]. The number of data and control slots and their durations are selected in order to satisfy QoS demands imposed during the DARPA SC2 competition.

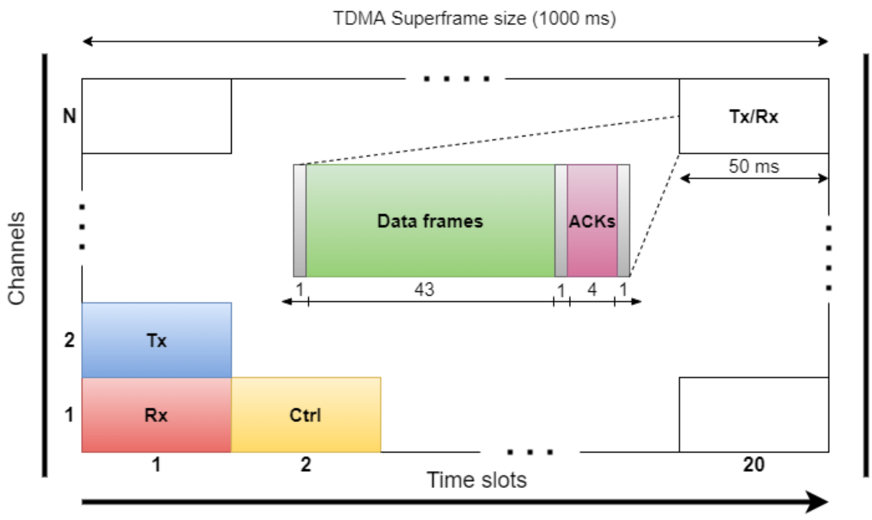


Figure 4.7: MF-TDMA superframe state for the performed simulations.

Based on the parameters from Table 4.2, the MF-TDMA superframe state for the performed simulations is presented in Fig. 4.7. Whereas the number of time slots is fixed, the number of channels varies and is specified later for every use case. It is assumed that all nodes in the network are comprised of a single transceiver supporting FDD full-duplex operations, so all the devices can simultaneously allocate Tx and Rx slots in the superframe at adjacent channels. As nodes both listen and transmit during control slots, for the duration of control slots no data transmission can be performed at adjacent channels. Four control slots are present within the superframe, and as a result, every node can allocate up to 16 Tx and 16 Rx slots. As presented in Fig. 4.7, data slots within the superframe are supporting fast ACKs, used for determining the PER value of slots. Slots may also be configured in any other way, as long as there is a reliable way to calculate slot quality. Within the adopted slot structure 43 ms are used for data transfer. With assumed LTE-based PHY, this leads to a maximum of 43 data frames transmitted per slot. The receiver reports the successful reception of data frames by transmitting ACKs. To conform with the minimum single frame size, every ACK carries a maximum of up to 11 confirmations of successful frame receptions. Therefore, the receiver transmits 4 ACKs in total, occupying 4 ms per slot. The remaining 3 ms within a slot are used as guard space, allowing sub-ms time drift between two nodes participating in the communication, thus simulating time drifts probable in real-time systems.

The maximum transmitting capacity of a node is calculated as the maximum number of Tx slots multiplied by the maximum number of frames supported per slot. As a result, the maximum achievable data rate of a node is 688 frames/s. For analyzing system behavior with different application loads, two application data rates are introduced, with one application packet fitting into a single MAC frame. The data rate of 800 packets/s, referred to as a high data rate, is the data rate that saturates the capacity of a transmitting node. The data rate of 400 packets/s is below the maximum supported data rate and in the remainder of the section is called low data rate.

Table 4.3: Default parameters for DDMC-TDMA protocol.

Protocol parameter	**Value**
Maximum number of proposed slots	10
Idle period	5 superframes =>5000 ms
Poor quality period	2 (fixed) or rand(2,5) superframes =>2000 or 2000-5000 ms
PER removal initiation threshold	20%
Maximum number of control message retransmissions (*Nretr_max*)	3
Fixed scheduling table broadcasting delay (*Tshf*)	4000 ms
Random scheduling table broadcasting delay (*Tshr*)	rand(0,4000ms)
Period of scheduling table broadcasting (*Tsh*)	4000 ms + rand(0, 4000) ms
Slot allocation procedure timeout (*Talloc*)	12000 ms
Twait_min	2500 ms
Twait_max	3500 ms
Slot allocation procedure delay (*Twait*)	rand(2500, 3500) ms

Adopted values of parameters described in Section 4.3 for the proposed DDMC-TDMA protocol are given in Table 4.3. Optimal values for default parameters of DDMC-TDMA protocol were deduced by performing multiple simulations in different wireless network topologies consisting of a different number of nodes. Surely, different values would result in better performance for specific use cases, but values given in Table 4.3 are generic for all use cases presented in this section. It is worth noting that a high value for PER removal initiation threshold is adopted, as the system also needs to work with scenarios with high network load and interference conditions.

4.4.1 Use case 1. Single-hop collision domain networks and scaling examination

In the first use case, simulations are executed to validate the basic functionality and scalability of the proposed protocol for wireless networks operating in a single-hop wireless domain. The number of nodes is gradually increased from 10 to 50 in steps of 10. Every node in the network establishes two communication links with other nodes of the network. In one of the links, the node is acting as a transmitter, whereas in the other link as a receiver. A high application data rate is assumed, thus every node tries to allocate the maximum number of transmitting slots within the superframe.

Three different sets of simulations were conducted, with the number of executed slot allocation/removal procedures and convergence time to steady state analyzed and compared. In the first and second set of simulations, the number of available frequency channels is set to 16, whereas in the third simulation, the number of channels is 50% higher than the number of nodes in the network. Therefore, in the first two sets of simulations, full spectrum utilization is expected as nodes have to compete for a limited spectrum. In the third set of simulations, two-thirds of the available spectrum is expected to be allocated for network communications. As there is enough available spectrum to satisfy network throughput requirements, every node is guaranteed enough data slots to satisfy its application demands. Another difference between the sets of simulations is in the adopted value for a poor quality period after which slot removal operations are initiated. In the first and third sets of simulations, a fixed value from Table 4.3 is adopted. For every allocated slot in the second set of simulations, the poor quality period is randomly calculated as an integer number between 2 and 5 consecutive superframes. In case that multiple communication links synchronously allocate the same slot, a randomized poor quality period increases the probability that one communication link keeps the slot while other links vacate the slot. In comparison, a fixed poor quality period may lead to all links releasing the slot. In summary, the first set of simulations is conducted with basic DDMC-TDMA, the second set with DDMC-TDMA with randomized initiation of high PER removal operations, and the third set with basic DDMC-TDMA in an environment with more frequency channels available.

Convergence times, being the times needed by the networks to achieve a steady state, are presented in Fig. 4.8 for the three sets of simulations. A number of executed slot allocation (A) and slot removal (R) procedures, averaged over 20

Table 4.4: Averaged numbers of the executed slot allocation (A)/removal (R) procedures until a steady state is reached.

Number of nodes	Basic DDMC-TDMA	Basic DDMC-TDMA + random PER removal	Basic DDMC-TDMA + more channels
20	A: 304.80 R: 60.70	A: 299.80 R: 54.0	A: 332.50 R: 25.10
30	A: 405.20 R: 184.90	A: 378.40 R: 154.50	A: 503.40 R: 70.0
40	A: 565.0 R: 382.30	A: 545.0 R: 352.60	A: 681.10 R: 131.80
50	A: 739.10 R: 613.80	A: 706.0 R: 544.70	A: 851.30 R: 210.60

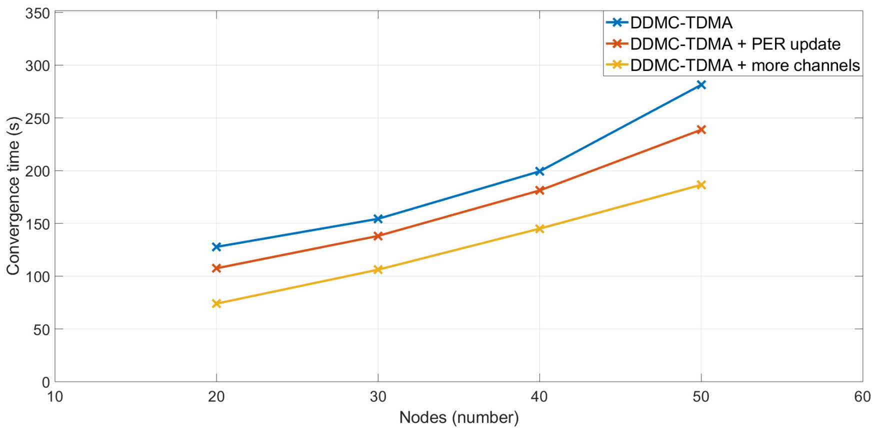


Figure 4.8: Convergence times for 3 sets of single-hop simulations.

independent ns-3 simulations, are shown in Table 4.4. From Fig. 4.8 and Table 4.4 it can be concluded that for the three sets of simulations, convergence time increases with the number of nodes and the number of executed slot allocation/removal procedures. Nevertheless, a steady state with maximum achievable spectrum utilization is reached in all three sets of simulations regardless of network density.

There are a couple of factors leading to increased convergence times with increasing network density. More communicating network nodes lead to higher congestion of available control slots. More control messages need to be exchanged and, hence, there is a higher probability of control messages colliding. As such, slot allocation/removal procedures are less reliable and multiple retransmissions of control messages might be required to complete initiated procedures, resulting in increased execution times of procedures and consequently increased convergence times. The increase in control overhead with the increase in network density is shown in Fig. 4.9. If control slots get saturated, it might even lead to failure in finalizing some of the initiated slot allocation/removal procedures, causing them to timeout. Heavy utilization of control slots might also lead to nodes missing slot usage reports, i.e. protocol ACKs or slot usage messages. Missed slot usage reports, synchronous execution of the protocol in the different nodes of the network, and the randomness of slot selection might result in two or more nodes allocating the same slot simultaneously. As a result, communication links within overlapping slots will interfere, and a high PER removal procedure is triggered and executed by one or more nodes. The probability of overlapping slot allocations is proportional to the active communication links and inverse proportional to the available radio spectrum, which is limited for the first two sets of conducted simulations.

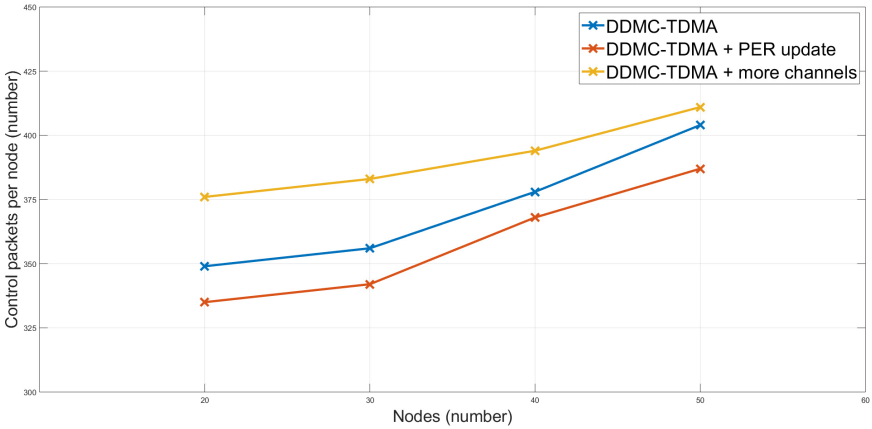


Figure 4.9: Per node control overhead of single-hop networks.

In Fig. 4.10 Tx slot allocation graphs are presented. Vertical dashed lines represent a point in time after which the number of allocated Tx slots within the network stays above 95% of its maximum achievable value. It can be seen that for the first two sets of simulations, approximately above 95% slot allocation is achieved relatively quickly. Afterward, network nodes struggle to successfully allocate and preserve the remaining unallocated slots. All nodes in the network are trying to synchronously allocate the last couple of slots to satisfy their throughput demands. For the last 5% of the slots, a limited number of nodes are allocating overlapping slots followed by initiating slot removal procedures. This process continues until the nodes finally allocate slots that do not overlap with slots allocated by other nodes. A prolonged period of slot allocations and removals, when only a limited number of free slots (or unutilized spectrum resources) is available, influences the convergence time significantly.

With the random calculation of the poor quality period instead of the fixed poor quality period, the probability that one link remains active within a slot allocated by several links is increased, while other links vacate the slot, thus solving internal interference. This adjustment of basic DDMC-TDMA leads to reduced

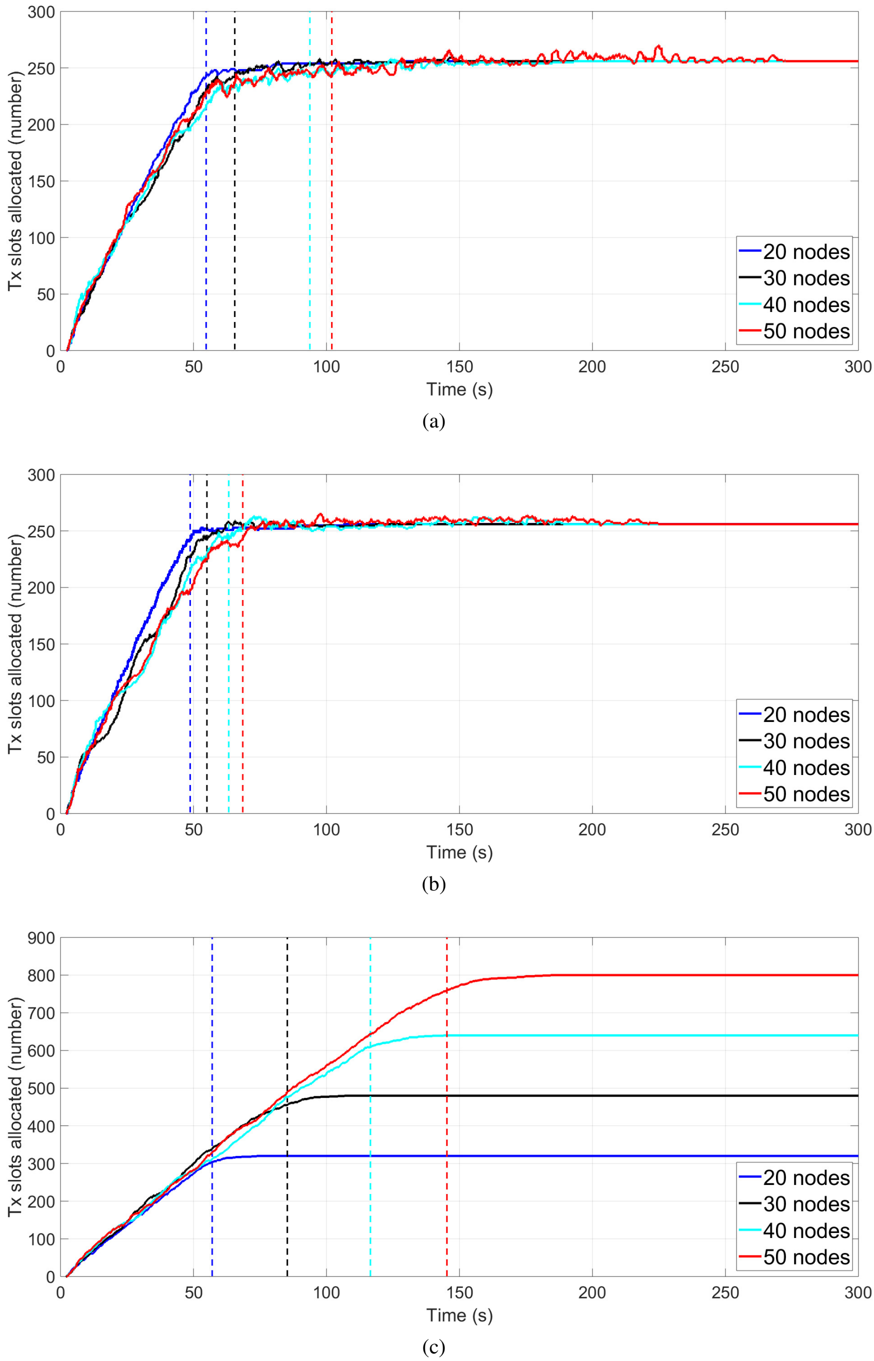


Figure 4.10: Tx slot allocation graphs. (a) Basic DDMC-TDMA. (b) Basic DDMC-TDMA + random PER removal. (c) Basic DDMC-TDMA + more channels.

convergence time and the number of allocation/removal procedures for all network densities (see Fig. 4.8 and Table 4.4).

Despite using a fixed value for poor quality period threshold, the third set of simulations with more available channels offers the fastest convergence time, as seen in Fig. 4.8. As more radio spectrum is available, overlapping slot allocations are less probable. In Fig. 4.10 we can indeed see that the number of allocated Tx slots is almost linearly increasing until a steady state is achieved. The oscillatory behavior (due to prolonged allocations and removals of the last couple of slots) observed in the first two sets of simulations, does not occur in the third set of simulations. Reaching the limits of the available spectrum resources is hence the most significant factor for increasing the convergence time. This is also confirmed in Table 4.4, where the third set of simulations displays much fewer removal procedures compared to the first two sets of simulations. The number of allocations for this set of simulations is higher, which can be attributed to more available frequency channels and accordingly more data slots to allocate.

4.4.2 Use case 2. Single-hop networks with external interference

With increasing wireless demands, also the probability of multiple wireless networks residing in the same geographical area is increasing, leading to mutual interference and reduction in achieved QoS. To satisfy reliability and throughput requirements, networks running DDMC-TDMA can detect time-frequency slots interfered by external networks and reallocate slots to unused slots that are not interfered.

In this use case, simulations are conducted with ad hoc networks using DDMC-TDMA and operating in the presence of other unknown external networks in their collision domain. These external networks may form any topology, use any medium access method, and do not employ coexistence and collision avoidance techniques. The ability of DDMC-TDMA based networks to mitigate external interference from these unknown external networks is examined. Sizes of DDMC-TDMA based networks are ranging from 10 to 50 with an incremental step of 10. For all executed simulations, it is assumed that external networks are inactive at the beginning of the simulation and start operating in the shared radio spectrum 250 seconds after the start. There are sufficient spectral resources, meaning sufficient time-frequency slots, to serve all traffic demands in the DDMC-TDMA based network and external networks (number of channels is 50% higher than the number of nodes in the DDMC-TDMA network). Spectrum usage of external networks is varied per simulation, so that induced interference to the DDMC-TDMA network is in the range of 1-5 Tx slots per transmitter node. For example, in the first set of simulations, one Tx slot of every transmitter node experiences interference, whereas, in the second set, 2 Tx slots are interfered per transmitter node, etc.

For the simplicity of executing simulations and comparing the performance of DDMC-TDMA in presence of external networks with different spectrum footprints, it is assumed that external networks follow the same MF-TDMA scheme as DDMC-TDMA based networks. In reality, external networks would have different channelization and spectrum utilization in the time domain. However, if partially overlapping (in the frequency domain, time domain or both) spectrum usage by external networks creates enough interference so that DDMC-TDMA based networks' receivers are unable to correctly decode packets, slots that are experiencing packet loss above the predefined PER threshold will be reallocated. Additionally, in real-world deployments, the influence of external networks that are fully or partially overlapping with slots allocated by DDMC-TDMA based networks may not be significant, due to distance between the networks, Tx power applied, modulation scheme, medium access scheme, etc., in which case there is no need for reallocation of the slots. So, regardless of the characteristics of the external networks, DDMC-TDMA only triggers reallocation of the slots that are being affected, by monitoring link statistics per slot.

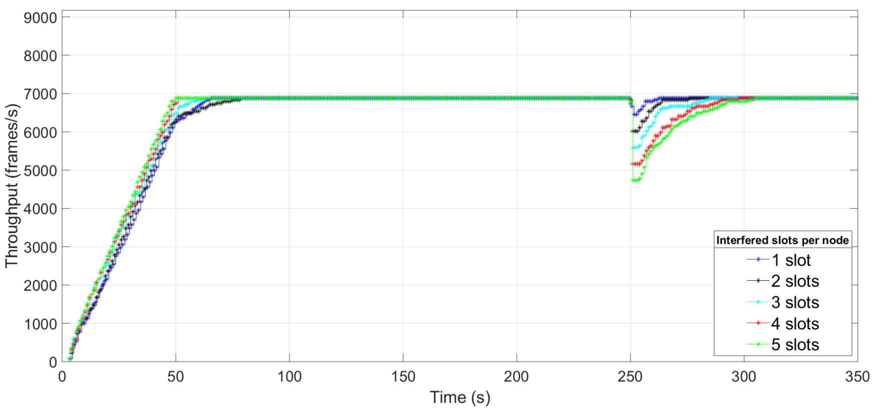


Figure 4.11: Throughput graph of 10 nodes network with different external interference.

The throughput graph of the ad hoc network consisting of 10 nodes and operating in presence of unknown external networks with different spectrum footprints is presented in Fig. 4.11. It can be seen that after an initial drop in the throughput, the DDMC-TDMA network will recover to its initial status by removing the interfered slots and allocating new slots that are free from interference. Regardless of the size of interference, the network employing DDMC-TDMA can mitigate the interference and recover the stable operation, meeting throughput demands. However, as is expected, throughput drop and recovery time are higher in the case of more interfered slots.

Figure 4.12: Superframe state of 10 nodes network before and after mitigation of external interference.

An example of accomplished external interference mitigation and reallocation of interfered slots to interference-free slots of the radio spectrum can be seen in Fig. 4.12. In this figure, the size of the DDMC-TDMA network is 10 nodes and the interference from external networks is 5 Tx slots per node, resulting in 50 interfered slots in total. The upper part of Fig. 4.12 shows the steady superframe state of the ad hoc network, which is achieved before the impact of external interference, whereas the lower part of Fig. 4.12 shows how the interfered slots have been reallocated to slots in the non-interfered section of the radio spectrum. In this way, the DDMC-TDMA network maintains its initial performance, while the external networks keep operating in the spectrum they occupied at the 250th second.

Interference recovery times for DDMC-TDMA networks with varying sizes

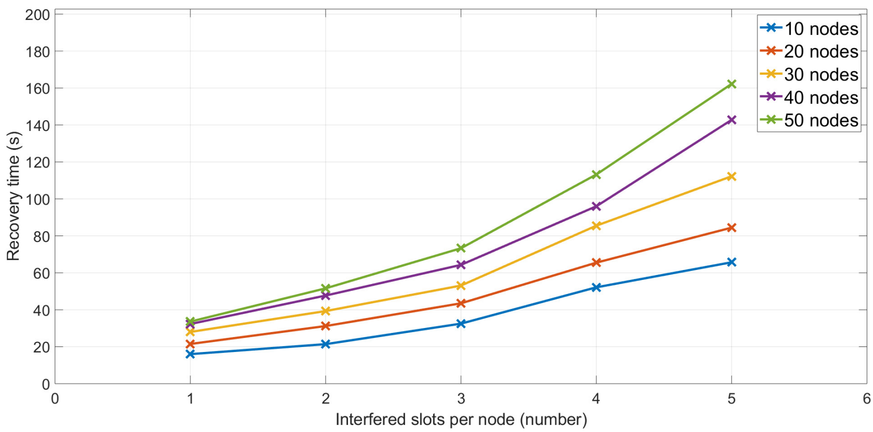


Figure 4.13: Recovery times of the networks with different sizes, operating in a presence of various external interference.

and with a different number of externally interfered Tx slots per node are presented in Fig. 4.13. Recovery times increase with the number of interfered Tx slots per node for all network sizes, and also with increasing network size. This is in line with the expectations, as with more nodes in the network, more slots are interfered, and more slots need to be reallocated.

An additional increase of recovery times may be attributed to the reallocation procedure, where due to the synchronous nature of the protocol, multiple links may allocate overlapping slots, as previously explained in the case of single-hop networks. In the conducted simulations, external networks occupy fixed parts of the radio spectrum. Nevertheless, DDMC-TDMA is able to mitigate dynamic interference induced by external networks, if their spectrum footprint is not changing rapidly over time.

4.4.3 Use case 3. Multi-hop networks spatial reuse and scaling examination

In an ad hoc network with multi-hop topology, the same time-frequency slots can be reused, due to the spatial separation of nodes in non-overlapping collision domains, where communication links do not interfere with each other. To optimize spatial reuse of spectrum, multi-hop networks must avoid the hidden and exposed node problems that may occur.

To demonstrate the reliable and scalable performance of DDMC-TDMA and its ability to achieve optimal spatial spectrum reuse in the case of multi-hop ad hoc networks, a set of simulations is executed. The number of nodes in the network is varied from 100 to 500 with a step of 100. For every size of the ad hoc network, sets of simulations with different network densities are executed. Random topology graph models of ad hoc networks are generated where every node of the network has 10 neighbor nodes in the first simulation set, increasing with a step of 5 in subsequent simulation sets, up to 30 neighbor nodes in the last simulation set. Every node in the network communicates with two neighbor nodes: transmitting to one neighbor node and receiving from another neighbor node. The superframe has the same configuration as in the single-hop use case: the number of available channels in the radio spectrum is fixed to 16 and the maximum number of available data slots is 256. However, please note that in a multi-hop network the same slot can be allocated multiple times in different collision domains. A low data rate is adopted and in order to satisfy its application throughput demands, every node is required to allocate 10 slots. This means, that, for instance, in a multi-hop ad hoc network consisting of 100 nodes, 1000 Tx slots must be allocated to accommodate all traffic.

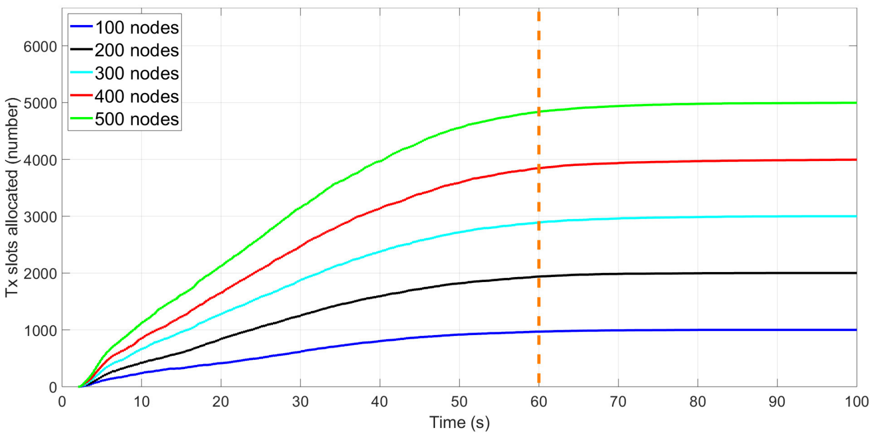


Figure 4.14: Tx slot allocation graph for the multi-hop network topology with the number of neighbor nodes fixed to 20.

Fig. 4.14 represents an example of allocated Tx slots of the whole multi-hop network over time. The number of neighbor nodes is fixed to 20 for this graph. We observe consistent convergence to steady network state and scalable DDMC-TDMA operation without any degradation of network performance when network size increases, as shown by the linear increase of allocated Tx slots with network size. Fig. 4.14 further shows that regardless of the network size, 95% of slot allocations is achieved within 60 seconds, after which it converges towards optimal resource utilization. Convergence towards optimal resource allocation is also observed for simulations with a different number of neighbor nodes (see Fig. 4.15). The fact that many more than 256 Tx slots can be allocated and that steady convergence is achieved proves the scalable and reliable operation of DDMC-TDMA and

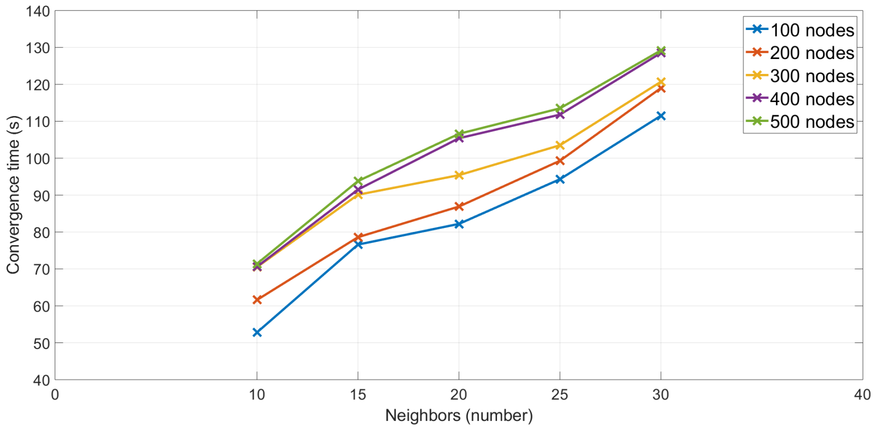
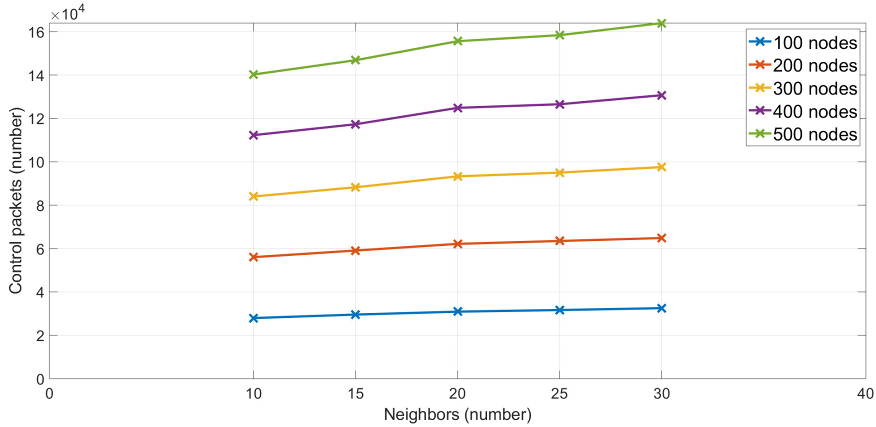


Figure 4.15: Convergence times versus the number of neighbor nodes for various sizes of multi-hop networks.

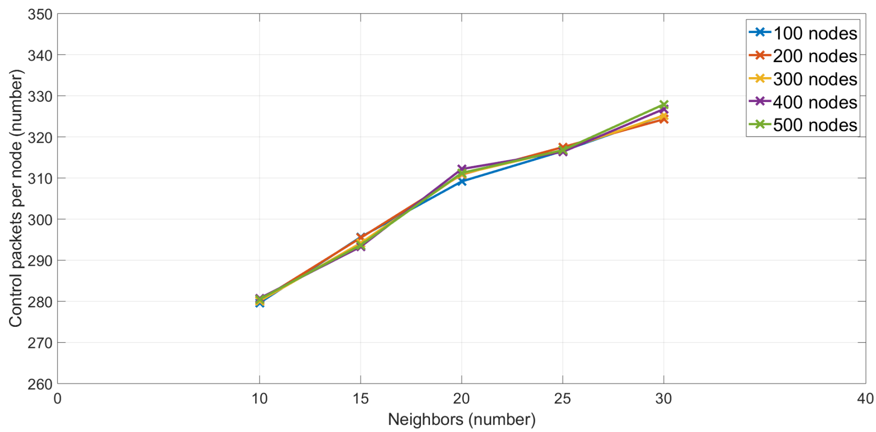
the capability to reuse slots spatially. At the same time, hidden and exposed node problems are also avoided, as this is an inherent feature of the protocol design. Occasionally, the hidden node problem might still occur due to parallel execution of the protocol or missed slot usage reports, but this will be resolved by the initiation of a slot removal procedure. Protection against the exposed node problem is described in more detail in the following use case.

Fig. 4.15 shows the scalability of the protocol in terms of convergence time with increasing network density. Similar to the single-hop use case, the convergence time increases linearly with network density, indicating that the scalability does not depend on the number of hops in a network. The limiting factor for scalability is the higher control slot utilization and a higher probability of overlapping slot allocations when the network density increases.

In Fig. 4.16 control overhead of multi-hop networks with different sizes and densities is presented; the overall control overhead of networks in Fig. 4.16a and per node control overhead in Fig. 4.16b. With more nodes in the multi-hop networks, more slots need to be allocated, therefore the control overhead is linearly increasing with a linear increase in network size, as shown in Fig. 4.16a. In addition, control overhead slightly increases with denser wireless deployments, which is more obvious from Fig. 4.16b. Per node control overhead increases from 280 packets up to approximately 315 packets for network densities of 10 and 30, respectively. Fig. 4.16b further proves the scalability of the algorithm, as per node control overhead only increases with an increase in network density and is not affected by an increase in network size. It is worth noting that control overhead depends on application demands. If there is no traffic for one or more nodes, there is no exchange of control messages, leading to zero control overhead for nodes in



(a)



(b)

Figure 4.16: Control overhead of multi-hop networks. (a) Overall. (b) Per node.

consideration. In executed simulations, traffic demands are constantly present for every node, so this feature of the algorithm is not shown.

4.4.4 Use case 4. Prevention of exposed node problem in multi-hop networks

In this section, the performance and behavior of DDMC-TDMA are analyzed in the case of network topologies and communication patterns that are vulnerable to the exposed node problem. The main focus of the conducted simulations is to validate how DDMC-TDMA is capable to avoid the exposed node problem in order to maximize spatial reuse of the spectrum. The network topology adopted for this use case is presented in Fig. 4.17.

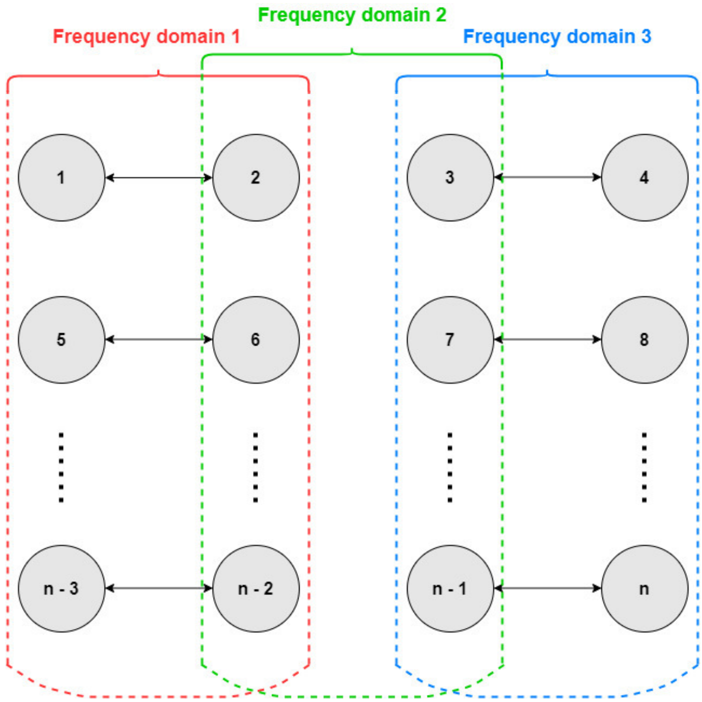


Figure 4.17: Network topology to trigger the exposed node problem.

This network topology is easy to scale up the network in size and allows to compare the performance for different network sizes. The network nodes are divided into 4 groups:

– group 1: nodes 1, 5, ..., *n*-3

– group 2: nodes 2, 6, ..., *n*-2

– group 3: nodes 3, 7, ..., *n*-1

– group 4: nodes 4, 8, ..., *n*,

where *n* (divisible by 4) is the number of nodes in an ad hoc network. Nodes from groups 1 and 2 are in the same collision domain and unidirectional data traffic is assumed from one node of group 1 to one node of group 2 or the other way around. The same applies to nodes from groups 3 and 4. Nodes from groups 2 and 3 do not communicate but are in the same collision domain. Such a network topology and point-to-point communication links lead to the occurrence of the exposed node problem.

The exposed node avoidance capability was analyzed for different numbers of frequency channels and different network sizes. The optimal number of channels (*n_channels*) per simulation is calculated as follows: for a network size of 4 nodes (see Fig. 4.17), the number of established communication links is 2. For a high data rate, in the optimal case, the two Tx nodes should allocate 16 data slots each. If the proposed protocol offers protection against the exposed node problem, one channel should suffice. For *n* nodes, this generalizes as follows:

$$n\_channels = \frac{n}{4} \tag{4.3}$$

Optimal spectrum utilization is achieved if every frequency-time slot in the superframe is allocated for any two non-interfering communication links, i.e. one link established between groups 1 and 2 and the other between groups 3 and 4. In this case, the number of allocated Tx slots is:

$$n\_allocated\_tx = n\_channels \cdot 16 \cdot 2 \tag{4.4}$$

where 16 represents the number of available data slots per channel. If we substitute Eq. 4.3 in Eq. 4.4, then the number of expected allocated Tx slots is:

$$n\_allocated\_tx = n \cdot 8 \tag{4.5}$$

Unlike most existing distributed scheduling protocols that consider Tx or Rx slots allocated in two-hop neighborhood unavailable for further utilization, the DDMC-TDMA approach detects the presence of exposed nodes and hence avoids spectrum underutilization. As is explained in Section 4.3.2.1, DDMC-TDMA avoids exposed node problem by keeping track of three different types of slot information, i.e. USED, USED Tx, and USED Rx. If no protection is offered against exposed nodes, the expected number of allocated Tx slots is half of the number calculated in Eq. 4.5.

Variable *n_allocated_tx* represents the optimal number of allocated slots in case that the exposed node problem is avoided. To prove the capability of DDMC-TDMA to reach this optimal number of allocated slots, simulations are performed with network size increasing from 4 to 116 nodes in steps of 8. For every simulation, a high application data rate is used. The achieved results meet the expectations: the number of allocated Tx slots is always equal to *n_allocated_tx* for DDMC-TDMA networks (4 nodes - 32 slots, 12 nodes - 96 slots, ..., 116 nodes - 928 slots). As the results are uniform for all simulated network sizes, it is shown that the exposed node avoidance feature of DDMC-TDMA is also scalable with increasing network size.

As previously mentioned in Section 4.3.2.1, the proposed protocol targets to reuse slots that are already used by the network nodes. In this way, more spectrum is left for external networks operating in the collision domain of the wireless

Table 4.5: Comparison of the optimal and achieved number of reused slots.

Nodes	Non-reused slots	Reused slots (achieved)	Reused slots (optimal)	Reused slots achieved percentage (achieved·100/optimal) (%)
4	0.2	9.9	10	99.00
12	1.4	29.3	30	97.67
20	1.8	49.1	50	98.20
28	2.8	68.6	70	98.00
36	3.2	88.4	90	98.22
44	2.7	108.7	110	98.82
52	3.4	128.3	130	98.69
60	5.8	147.2	150	98.13
68	5.1	167.6	170	98.59
76	6.9	186.9	190	98.37
84	9.2	205.7	210	97.95
92	8.9	225.7	230	98.13
100	11.0	244.8	250	97.92
108	12.7	264.4	270	97.93
116	14.5	283.5	290	97.76

network in consideration. To analyze the practical feasibility in terms of spectrum reuse, simulations are performed using the network topology from Fig. 4.17 with a varying number of network nodes. In addition to spectrum reuse, these simulations further demonstrate the capability of DDMC-TDMA to avoid exposed node problem. The application data rate adopted for these simulations is a low data rate. For example, in the case of 1 frequency channel available and a network consisting of 4 nodes, it is expected that 10 data slots are allocated and reused by 2 established and non-interfering communication links, while the remaining 6 data slots stay free. Generalizing this example, the optimal number of reused Tx slots can be calculated as:

$$n\_reused\_tx = n\_channels \cdot 10 \tag{4.6}$$

If we substitute Eq. 4.3 in Eq. 4.6, the optimal number of reused Tx slots is:

$$n\_reused\_tx = \frac{n \cdot 5}{2} \tag{4.7}$$

Optimal numbers of reused slots and achieved numbers of reused slots in conducted simulations are presented in Table 4.5. As they represent an average value within a set of 20 runs, presented numbers are not decimal numbers. For every network size, the optimal number of reused slots calculated based on Eq. 4.7 is presented under column 'Reused slots (optimal)'. However, due to the synchronous execution of the protocol and the limited number of slots that can be embedded in the proposed slots message, we can see from column 'Reused slots (achieved)' that achieved reusage of the slots is not reaching the optimal number. Instead of the preferred reuse of already allocated slots, it can be seen from the 'Non-reused slots' column that there are slots with only one communication link established. With an increase in network size, there is also an increase in non-reused slots. However, with the increased number of network nodes, there are more available channels in the radio spectrum and more slots to be allocated, thus an increased probability of occurrence of non-reused slots is expected. The best indication of spectrum reusage efficiency can be obtained if we calculate what percentage of optimal reusages is achieved. This percentage is presented in the last column of Table 4.5. It shows that the percentage-wise difference between the optimal and achieved number of reused slots is kept around the same value, thus no degradation of performance is induced with an increase in network size and an increase in the number of frequency-time slots to be allocated.

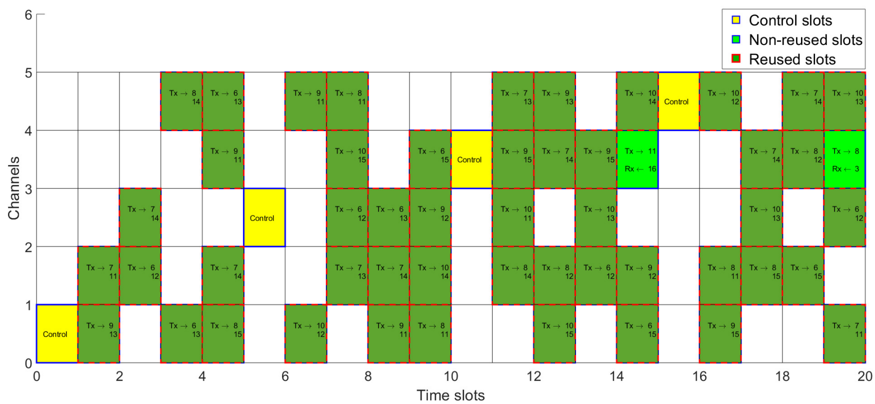


Figure 4.18: Superframe state of 20 nodes network running DDMC-TDMA with achieved near-optimal slot reusage.

An example of spectrum utilization for a 20 nodes network is presented in Fig. 4.18. Only in two time-frequency slots, single communication links are established. In an ideal case, these two links would occupy the same slot, leaving one more slot for external networks. All other slots are successfully reused by two communication links.

4.5 Conclusion

In this chapter, we have proposed a novel medium access protocol, called Dynamic Distributed Multi-Channel TDMA (DDMC-TDMA), that is capable to address the main shortcomings in wireless ad hoc networks, such as scalability, spectral efficiency, QoS, and coexistence of multiple co-located ad hoc networks. DDMC-TDMA was initially used and experimentally validated in the DARPA SC2 competition, where its flexibility and reliability in various challenging wireless scenarios for dynamic spectrum sharing with other unknown wireless technologies has been successfully proven.

In this chapter, the detailed architectural design of the DDMC-TDMA protocol and its modules has been presented. The central module is the scheduling table with different types of internally and externally used slots that are maintained per node. The architecture further includes procedures and associated control messages for slot allocation and removal operations. The exchange of control messages happens via dedicated control slots in a medium access scheme. The DDMC-TDMA protocol has been validated through ns-3 simulations for various single-hop and multi-hop wireless scenarios and it is shown that the protocol is scalable in all scenarios, without performance degradation. For single-hop networks, convergence times are proportional to the number of nodes (varied between 20 and 50) and the number of slots available for resource allocation. The time needed for achieving above 95% spectrum utilization is in the range of 55-100 seconds for basic DDMC-TDMA and in the range of 50-70 seconds for DDMC-TDMA with the optimization of randomized initiation of slot removal operations. After the convergence time, single-hop networks are able to converge to a steady state and fully utilize the available radio spectrum. The slot allocation mechanism has proven to be reliable and scalable for different network densities within the single-hop collision domain. It has also been shown that with small extensions on top of basic DDMC-TDMA, network performance can be further improved.

The performance of DDMC-TDMA has also been evaluated in presence of multiple co-located and unknown external networks. The size of the DDMC-TDMA based network is varied from 10 to 50 nodes, with each node utilizing 16 Tx slots and with external networks causing interference in 1 up to 5 Tx slots per node. Depending on the number of concurrently interfered slots, the recovery time, being the time needed until all affected slots are reallocated and a steady state is achieved again, is ranging between 18 and 160 seconds. It has been shown that with DDMC-TDMA, networks of different sizes and densities fully recover in the presence of different interference levels. This proves that DDMC-TDMA by design can coexist with unknown external networks operating in the same collision domain.

In the multi-hop network, regardless of the network size, 95% spectrum utilization is achieved within 60 seconds. All links are able to allocate sufficient slots to satisfy application demands by maximally exploiting spatial reuse of time-frequency slots. As a result, the spectrum footprint of ad hoc networks can be minimized, meaning that spectrum efficiency is optimized, hence preserving more spectrum for other networks operating in the same collision domain. Networks running DDMC-TDMA are able to avoid both hidden and exposed node problems leading to near-optimal spectrum utilization and improved QoS guarantees. For network topologies suffering from the exposed node problem, this problem is fully mitigated, regardless of network size. At the same time, spatial reuse is achieved for around 98% of slots.

The DDMC-TDMA protocol can be applied to any wireless network that requires the high-density deployment and needs to adapt to a dynamic environment while maintaining high reliability, high spectral efficiency, and seamless coexistence with external networks. In summary, the proposed DDMC-TDMA protocol offers an all-in-one solution tackling by design many well-known problems only partially addressed by other existing distributed TDMA scheduling protocols.

References

[1] Y. Song and J. Xie. *BRACER: A Distributed Broadcast Protocol in Multi-Hop Cognitive Radio Ad Hoc Networks with Collision Avoidance*. IEEE Trans. Mobile Comput., 14(3):509–524, Mar. 2015. doi:10.1109/TMC.2014.2328998.

[2] M. Hadded, P. Muhlethaler, A. Laouiti, R. Zagrouba, and L. A. Saidane. *TDMA-Based MAC Protocols for Vehicular Ad Hoc Networks: A Survey, Qualitative Analysis, and Open Research Issues*. IEEE Commun. Surveys Tuts., 17(4):2461–2492, Oct-Dec. 2015. doi:10.1109/COMST.2015.2440374.

[3] W. Zafar and B. Muhammad Khan. *Flying Ad-Hoc Networks: Technological and Social Implications*. IEEE Technol. Soc. Mag., 35(2):67–74, June 2016. doi:10.1109/MTS.2016.2554418.

[4] S. Sharmila and T. Shanthi. *A survey on wireless ad hoc network: Issues and implementation*. In Int. Conf. Emerg. Trends Eng., Technol. Sci. (ICETETS), pages 1–6, 2016.

[5] W. Ye, J. Heidemann, and D. Estrin. *An energy-efficient MAC protocol for wireless sensor networks*. In Proc. 21st Annu. Joint Conf. IEEE Comput. Commun. Soc., volume 3, pages 1567–1576, June 2002. doi:10.1109/INFCOM.2002.1019408.

[6] I. Rhee, A. Warrier, M. Aia, J. Min, and M. L. Sichitiu. *Z-MAC: A Hybrid MAC for Wireless Sensor Networks*. IEEE/ACM Trans. Netw., 16(3):511–524, June 2008. doi:10.1109/TNET.2007.900704.

[7] D. J. Vergados, D. D. Vergados, and M. Koutsogiannaki. *Capacity Optimization in TDMA Ad-Hoc Networks*. Wirel. Pers. Commun., 62(3):687–713, 2012. doi:10.1007/s11277-010-0089-2.

[8] M. Yu, C. Zhou, and M. Zhao. *De-jitter Dynamic Slot Allocation Algorithm Based on Cluster Analysis*. Comput. Sci., 44(7):94–97, July 2017.

[9] Y. Li, X. Zhang, J. Zeng, Y. Wan, and F. Ma. *A Distributed TDMA Scheduling Algorithm Based on Energy-Topology Factor in Internet of Things*. IEEE Access, 5:10757–10768, 2017. doi:10.1109/ACCESS.2017.2710304.

[10] M. A. Sheikh, N. B. Z. Ali, and A. Awang. *Distributive and self-sustainable scheduling algorithm for wireless sensor networks*. In 22nd Wireless Opt. Commun. Conf., pages 415–420, May 2013. doi:10.1109/WOCC.2013.6676403.

[11] D. Helen and D. D. Arivazhagan. *Applications, Advantages and Challenges of Ad Hoc Networks*. J. Acad. Ind. Res., 2(8):453–457, Jan. 2014.

[12] A. Jayasuriya, S. Perreau, A. Dadej, and S. Gordon. *Hidden vs. exposed terminal problem in ad hoc networks*. In Proc. Aust. Telecommun. Netw. Appl. Conf., pages 52–59, Jan. 2004.

[13] Defense Advanced Research Project Agency (DARPA). *Spectrum Collaboration Challenge (SC2)*. Available from: https://archive.darpa.mil/sc2/.

[14] NS-3 Consortium. *ns-3 Network Simulator*. Available from: https://www.nsnam.org/.

[15] C. Zhu and M. S. Corson. *A five-phase reservation protocol (FPRP) for mobile ad hoc networks*. In Proc. IEEE INFOCOM '98, Conf. Comput. Commun., 17th Annu. Joint Conf. IEEE Comput. Commun. Soc. (Cat. No.98), volume 1, pages 322–331, Mar. 1998. doi:10.1109/INFCOM.1998.659669.

[16] Y. Li, X. Zhang, T. Qiu, J. Zeng, and P. Hu. *A Distributed TDMA Scheduling Algorithm Based on Exponential Backoff Rule and Energy-Topology Factor in Internet of Things*. IEEE Access, 5:20866–20879, 2017. doi:10.1109/ACCESS.2017.2758340.

[17] I. Rhee, A. Warrier, J. Min, and L. Xu. *DRAND: Distributed Randomized TDMA Scheduling for Wireless Ad Hoc Networks*. IEEE Trans. Mobile Comput., 8(10):1384–1396, Oct. 2009. doi:10.1109/TMC.2009.59.

[18] S. Ramanathan. *A unified framework and algorithm for (T/F/C)DMA channel assignment in wireless networks*. In Proc. INFOCOM '97, volume 2, pages 900–907, Apr. 1997. doi:10.1109/INFCOM.1997.644573.

[19] K. Sato and S. Sakata. *A distance-measurement-oriented distributed TDMA scheduling algorithm for sensor networks*. In Int. Conf. Distrib. Comput. Sens. Syst. Workshops (DCOSS), pages 1–3, June 2011. doi:10.1109/DCOSS.2011.5982187.

[20] K. L. Bryan, T. Ren, T. Henry, and V. Fay-Wolfe. *Towards Optimal TDMA Frame Size in Wireless Sensor Networks*. Technical Report TR-xxx, Dept. of Computer Science and Statistics, University of Rhode Island, 2007.

[21] B. Dezfouli, M. Radi, K. Whitehouse, S. A. Razak, and T. Hwee-Pink. *DICSA: Distributed and concurrent link scheduling algorithm for data gathering in wireless sensor networks*. Ad Hoc Netw., 25:54 – 71, 2015. doi:https://doi.org/10.1016/j.adhoc.2014.09.011.

[22] D. J. Vergados, N. Amelina, Y. Jiang, K. Kralevska, and O. Granichin. *Toward Optimal Distributed Node Scheduling in a Multihop Wireless Network Through Local Voting*. IEEE Trans. Wireless Commun., 17(1):400–414, 2018. doi:10.1109/TWC.2017.2767045.

[23] *IEEE Standards 802.15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specification for Low-Rate Wireless Personal Area Networks (LR-WPANs)*, Sept. 2006.

[24] *IEEE Standard for Local and metropolitan area networks–Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs) Amendment 1: MAC sublayer*. IEEE Std 802.15.4e-2012 (Amendment to IEEE Std 802.15.4-2011), pages 1–225, Apr. 2012. doi:10.1109/IEEESTD.2012.6185525.

[25] P. Thubert. *An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4*. Internet-Draft draft-ietf-6tisch-architecture-30, Internet Engineering Task Force, November 2020. Work in Progress. Available from: https://datatracker.ietf.org/doc/html/draft-ietf-6tisch-architecture-30.

[26] Q. Wang and X. Vilajosana. *6TiSCH Operation Sublayer (6top)*. Internet-Draft draft-wang-6tisch-6top-sublayer-04, Internet Engineering Task Force, November 2015. Work in Progress. Available from: https://datatracker.ietf.org/doc/html/draft-wang-6tisch-6top-sublayer-04.

[27] Q. Wang, X. Vilajosana, and T. Watteyne. *6TiSCH Operation Sublayer (6top) Protocol (6P)*. RFC 8480, November 2018. Available from: https://rfc-editor.org/rfc/rfc8480.txt, doi:10.17487/RFC8480.

[28] D. Dujovne, L. A. Grieco, M. R. Palattella, and N. Accettura. *6TiSCH 6top Scheduling Function Zero (SF0)*. Internet-Draft draft-ietf-6tisch-6top-sf0-05, Internet Engineering Task Force, July 2017. Work in Progress. Available from: https://datatracker.ietf.org/doc/html/draft-ietf-6tisch-6top-sf0-05.

[29] D. Dujovne, L. A. Grieco, M. R. Palattella, and N. Accettura. *6TiSCH Experimental Scheduling Function (SFX)*. Internet-Draft draft-ietf-6tisch-6top-sfx-01, Internet Engineering Task Force, March 2018. Work in Progress. Available from: https://datatracker.ietf.org/doc/html/draft-ietf-6tisch-6top-sfx-01.

[30] T. Chang, M. Vučinić, X. Vilajosana, S. Duquennoy, and D. Dujovne. *6TiSCH Minimal Scheduling Function (MSF)*. Internet-Draft draft-ietf-6tisch-msf-18, Internet Engineering Task Force, September 2020. Work in Progress. Available from: https://datatracker.ietf.org/doc/html/draft-ietf-6tisch-msf-18.

[31] S. Kharb and A. Singhrova. *A survey on network formation and scheduling algorithms for time slotted channel hopping in industrial networks*. J. Netw. Comput. Appl., 126:59–87, 2019. doi:https://doi.org/10.1016/j.jnca.2018.11.004.

[32] A. Aijaz, A. Stanoev, and U. Raza. *GALLOP: Toward High-Performance Connectivity for Closing Control Loops over Multi-Hop Wireless Networks*. In Proc. 27th Int. Conf. Real-Time Netw. Syst., RTNS '19, page 176–186, 2019. doi:10.1145/3356401.3356413.

[33] S. M. Kamruzzaman and M. S. Alam. *Dynamic TDMA slot reservation protocol for cognitive radio ad hoc networks*. In 13th Int. Conf. Comput. Inf. Technol. (ICCIT), pages 142–147, Dec. 2010. doi:10.1109/ICCITECHN.2010.5723844.

[34] S. Agarwal, R. K. Shakya, Y. N. Singh, and A. Roy. *DSAT-MAC: Dynamic slot allocation based TDMA MAC protocol for Cognitive Radio Networks*. In 9th Int. Conf. Wireless Opt. Commun. Netw. (WOCN), pages 1–6, Sept. 2012. doi:10.1109/WOCN.2012.6331889.

[35] D. Rodenas-Herraiz, F. Garcia-Sanchez, A.-J. Garcia-Sanchez, and J. Garcia-Haro. *On the influence of the hidden and exposed terminal problems on asynchronous IEEE 802.15.5 networks*. Comput. Stand. Interfaces, 42:53 – 70, 2015. doi:https://doi.org/10.1016/j.csi.2015.04.006.

[36] K. Sohrabi, J. Gao, V. Ailawadhi, and G. J. Pottie. *Protocols for self-organization of a wireless sensor network*. IEEE Pers. Commun., 7(5):16–27, Oct. 2000. doi:10.1109/98.878532.

[37] R. Nelson and L. Kleinrock. *Spatial TDMA - A collision-free multihop channel access protocol*. IEEE Trans. Commun., 33:934–944, Sept. 1985.

[38] R. T. B. Ma, V. Misra, and D. Rubenstein. *An Analysis of Generalized Slotted-Aloha Protocols*. IEEE/ACM Trans. Netw., 17(3):936–949, 2009.

[39] I. Jabandžić. *A Dynamic Distributed Multi-Channel TDMA (DMC-TDMA) scheduling algorithm implementation for ns-3*, Jan. 2020. Available from: https://github.com/IrfanJabandzic/MfTDMA-ns-3.

[40] F. A. P. d. Figueiredo, D. Stojadinovic, P. Maddala, R. Mennes, I. Jabandžić, X. Jiao, and I. Moerman. *Scatter Phy: A Physical Layer for the DARPA Spectrum Collaboration Challenge*. In IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN), pages 1–6, 2019.

5

The CODYSUN Approach: A Novel Distributed Paradigm for Dynamic Spectrum Sharing in Satellite Communications

Similar to terrestrial networks, satellite systems apply a fixed spectrum allocation scheme. With a still relatively low number of satellites deployed, this scheme is performing well. However, with planned deployments of thousands of satellites in the near future, there is a need for improved allocation schemes. This chapter proposes two dynamic spectrum sharing techniques that can be used as a basis for future spectrum management in satellite communications. It shows that if compared to fixed spectrum assignment, there are significant gains with proposed techniques applied in downlink and uplink Ka-band communication.

⋆ ⋆ ⋆

This chapter is adapted from:

I. Jabandžić, F. Firyaguna, S. Giannoulis, A. Shahid, A. Mukhopadhyay, M. Ruffini, and I. Moerman, *The CODYSUN Approach: A Novel Distributed Paradigm for Dynamic Spectrum Sharing in Satellite Communications*

Submitted to MDPI Sensors.

Abstract With a constant increase in the number of deployed satellites, it is expected that current fixed spectrum allocation in satellite communications (SATCOM) will migrate towards more dynamic and flexible spectrum sharing rules. This migration is accelerated due to the introduction of new terrestrial services in bands previously exclusively used by satellite services. There are already several dynamic spectrum sharing (DSS) solutions for SATCOM, however, they are mainly centralized solutions and might lead to scalability issues with increasing satellite density. This chapter describes the main outcomes of the COllaborative and DYnamic approaches to increasing Spectrum UtilizatioN (CODYSUN) study that has investigated the benefits of new, dynamic, and distributed spectrum access techniques, over currently existing methods to share spectrum among satellite services. During the CODYSUN study two relevant SATCOM use cases have been selected for dynamic spectrum sharing: opportunistic sharing of dual satellite and terrestrial systems in (i) downlink Ka-band and (ii) uplink Ka-band. For the two selected use cases, two distributed DSS techniques have been designed and analyzed for increasing spectrum utilization in satellite-terrestrial systems and for minimizing the impact of interference between satellite and terrestrial segments compared to static spectrum allocation. Notable performance gains have been obtained.

5.1 Introduction

In the present day, the spectrum usage of SATCOM is tightly regulated (mainly based on planned spatial and frequency separation), leading to static and fixed spectrum allocation. Such an approach works well, as long as the number of satellite operators and satellites is limited, and the services do not have too high bandwidth requirements. However, the number of satellites that will be launched in the next few years is expected to grow rapidly, from projects such as Starlink, oneWeb, etc. For instance, SpaceX as of 24 October 2020 has launched 895 Starlink satellites. Furthermore, they plan to launch nearly 12,000 satellites with a possible extension to 42,000 satellites [1]. Considering the increasing satellite density, the static and fixed spectrum allocation model is not sustainable and will lead to a shortage of spectrum. Fixed spectrum assignment will eventually yield poor performance either due to interference or poor spectrum utilization.

Equivalent Power Flux Density (EPFD) rules for interference management in SATCOM are still working fine for low-density scenarios. However, these rules are only designed for protecting static geostationary (GSO) satellites from interference generated by non-geostationary (NGSO) satellites or other GSO satellites. In addition, these rules are very conservative as they apply a big protection margin that seriously limits the accommodation of new NGSO satellites. To overcome this, it becomes imperative to reuse and share the spectrum by moving towards spectrum sensing and DSS for improving spectrum utilization and mitigating the impact of interference in high-density SATCOM scenarios.

Traditionally, the coexistence of satellite systems with other satellite systems operating in the same frequency bands was possible by employing simple resource allocation algorithms such as round-robin scheduling, proportional fair scheduling, etc. [2]. However, Federal Communications Commission (FCC) reported that these traditional resource allocation algorithms result in the underutilization of the frequency bands [3]. To keep up with the growing density of satellites, it is essential to find optimal resource allocation solutions that not only ensure the coexistence among multiple satellite systems but also increase spectrum utilization. However, interference management is a challenging task in dual satellite systems in which two kinds of satellites i.e., GSO and NGSO operate in overlapping coverage areas in the same frequency bands [4]. This situation becomes much more complicated due to many systems with thousands of such satellites all sharing the same frequency bands.

In SATCOM, traditionally, Fixed Satellite Services (FSS) use C and K bands whereas mobile satellites use L and S bands. Due to the limited availability of L and C bands and the continuously increasing demands of broadband services via mobile satellites, Ku and Ka bands have also been assigned to mobile satellite services. There has been tremendous pressure on these bands due to the launch of new terrestrial services including mobile telephony, LTE, WiMAX, etc. [5]. Uplink Ka-band is one of the main target bands for the mmWave (FR2) operation of the 5^{th} Generation (5G) New Radio (NR) specification. Both Korea in June 2018 and United States in January 2019 concluded spectrum auctions on uplink Ka-band for terrestrial 5G services [6]. This introduction of terrestrial services in bands previously exclusively used by satellite services further emphasizes the need for migration from fixed to dynamic spectrum assignment and demands coexistence not only between multiple satellites but also with terrestrial services, when operating in the same bands.

There are several SATCOM scenarios that may benefit from DSS. Recently, Ka-band is gaining interest because of the more mature and cost-effective technology and because this part of the spectrum is still less congested than the more traditional C and Ku bands [6]. Ka-band is mainly used by satellite operators for end-to-end broadband services for consumer Internet access. It is anticipated that limited resources, expanding Internet use, and increasing density of satellite deployments will complicate future Ka-band deployments [7]. The European Conference of Postal and Telecommunications Administrations (CEPT) has adopted a decision for downlink communication in Ka-band (17.7-19.7 GHz), ERC/DEC/(00)07 [8], which gives guidance on the use of this band by FSS and terrestrial networks.

ECC/DEC(05)01 was amended in March 2016. According to this decision, FSS stations can be deployed anywhere but without the right of protection from interference generated by Fixed Service (FS) terrestrial links. DSS techniques could significantly increase the FSS spectrum usage by utilizing spatial, time, and/or frequency separation with dynamic (re)allocation of spectrum resources to protect FSS from interference by other FSS and terrestrial links. By CEPT decision ECC/DEC/(05)01 [9], which is amended in March 2013 and in March 2019, a segmentation is provided between FSS and terrestrial stations in uplink Ka-band (27.5-29.5 GHz). As the terrestrial segment of this band is underutilized throughout Europe, FSS stations could maximize their spectrum utilization by dynamically exploiting the spectrum allocated to FSS and terrestrial segments of the uplink Ka-band. At the same time, the DSS techniques need to provide interference protection of terrestrial incumbent users, as well as mutual interference protection of FSS links.

In order to increase spectrum utilization while still guaranteeing interference-free operation, we investigated the effects of novel DSS techniques in the Ka-band SATCOM in terms of performance gains and mitigation of interference components. The CODYSUN [10] study was carried out to investigate the current status quo of SATCOM and propose ways to enable advanced DSS techniques for solving the spectrum crunch problems in SATCOM. Therefore, in CODYSUN, we identified shortcomings of existing SATCOM solutions and proposed novel DSS techniques to avoid interference and reach the highest possible level of spectrum efficiency and reuse by exploiting all possible degrees of freedom (frequency, time, and space). This chapter discloses for the first time the proposed techniques and achieved results from the CODYSUN study.

The DSS techniques investigated in the CODYSUN study have been inspired by the Defense Advanced Research Projects Agency (DARPA) Spectrum Collaboration Challenge (SC2) [11], where, as part of team SCATTER (see Chapter 3), we designed and implemented DSS techniques for terrestrial networks. The goal of SC2 was to ensure that the exponentially growing number of military and civilian wireless devices and services would have full access to the increasingly crowded electromagnetic spectrum. By breaking the walls between isolated silos of the exclusively assigned spectrum and offering a wider shared spectrum band to multiple independent and collaborative networks, current problems of underutilization of spectrum in exclusive spectrum bands, leading to a huge waste of spectrum, can be mitigated. By exploiting the spatial separation, the same spectrum can be simultaneously used by multiple communication entities, heavily boosting spectrum utilization efficiency. Starting from the results obtained in DARPA SC2 for terrestrial networks, the CODYSUN study has revealed that similar DSS techniques can be applied for spectrum sharing across both satellite and terrestrial deployments.

The main contributions of this chapter can be summarized as follows:

1. After careful analysis and consultation with the European Space Agency (ESA), the two most prominent use cases were selected: 1) opportunistic sharing of dual satellite and terrestrial systems in downlink Ka-band; and 2) opportunistic sharing of dual satellite and terrestrial systems in uplink Ka-band. The major criterion for selecting relevant dynamic use cases was the presence of three systems (GSO, NGSO, and terrestrial networks) operating in the same frequency band. In comparison, current research on applying DSS techniques in SATCOM is focused on use cases where only two systems share the spectrum.

2. To solve the interference and spectrum underutilization problems in selected use cases, two DSS techniques were presented: 1) DSS1 - collaboration protocol; and 2) DSS2 - decentralized spectrum sensing.

3. A baseline scheme with fixed spectrum assignment was defined for comparison of the DSS techniques in terms of performance according to the following Key Performance Indicators (KPIs): spectrum utilization, aggregated system throughput, Packet Error Rate (PER), latency, system uptime, control overhead, etc.

4. The first demonstration of distributed DSS techniques for efficient spectrum sharing across multiple satellite and terrestrial systems, while protecting GSO and NGSO downlink and uplink transmission as well as terrestrial links.

The remainder of this chapter is structured as follows: Section 5.2 presents an overview of state-of-the-art related to dynamic spectrum sharing and interference mitigation in SATCOM scenarios, Section 5.3 includes detailed information on the two selected use cases and problem statement, in Section 5.4 the proposed approach and requirements are defined, Section 5.5 comprises brief information on the proposed DSS techniques, Section 5.6 presents a simulation scenario and performance analysis of the DSS techniques against the baseline, and finally, Section 5.7 concludes the chapter.

5.2 Related work

In order to address the spectrum scarcity problem in satellite-terrestrial systems, research on Cognitive Radio (CR) and cognitive techniques has gained significant interest. Different cognitive techniques are considered for spectrum sharing and interference mitigation in SATCOM scenarios such as spectrum sensing, databases, cognitive zones, beamforming, beam hopping, power control, etc.

The recent work on cognitive satellite scenarios can be categorized into two main groups: satellite-terrestrial coexistence and dual satellite coexistence. In satellite-terrestrial coexistence, the spectrum is shared between a satellite network and a terrestrial network, whereas in dual satellite coexistence two satellite networks operate simultaneously in the same spectrum band. In a satellite-terrestrial network, the satellite network can be considered as the primary user, whereas the terrestrial system is the secondary user or vice versa. Literature analysis shows that major research efforts on cognitive SATCOM sharing and interference mitigation are based on the construction and maintenance of centralized databases, where activity rules and main characteristics of primary users are stored. The key idea of a centralized database is that before assigning a channel to a secondary user, the vacancy of the channel must be verified from the database for the duration of the targeted transmission period [12].

In the remainder of this section, we first present studies that investigated coexistence scenarios in satellite-terrestrial systems. Next, contributions in dual satellite coexistence scenarios are addressed. Finally, we present a couple of coexistence studies in both satellite-terrestrial and dual satellite systems that are based on database technique.

In [13], the authors proposed a joint power and subchannel allocation algorithm for the uplink scenario of the cognitive satellite-terrestrial network where the cognitive GSO reuses the frequency band of incumbent terrestrial cellular networks in the S-band on 2.5-2.6 GHz. In [14] cooperative techniques and CR are applied to the satellite-terrestrial network operating in downlink and uplink in various frequency bands. The satellite segment consists of three GSOs, while the terrestrial segment is a 3rd Generation (3G)/4th Generation (4G) heterogeneous network. Coexistence is achieved by cooperative spectrum sensing, where each Earth Station (ES) detects primary users' transmitted signal and reports it to the satellite. Then, the satellite processes the received information from all the cooperative ESs and makes the final decision on the state of the spectrum, which is further broadcasted back to all the ESs. The focus of research in [15] is uplink communication in Ka-band where FSS satellite terminals try to reuse frequency bands of FS terrestrial Microwave (MW) links which represent the incumbent users. Three different power allocation algorithms are used to control the transmit power of the FSS satellite terminals, thus keeping the aggregated interference caused at the FS system below some acceptable threshold. We can see from the presented papers on satellite-terrestrial coexistence that various cognitive techniques are applied in different frequency bands. However, all techniques focus on the static use case where the satellite segment consists only of GSO satellites and FSS to static ESs, and they cannot be applied to dynamic systems that, in addition to GSO satellites and terrestrial networks, may consist of NGSO satellites and mobile Earth Stations in Motion (ESIMs).

Reference [16] proposes an adaptive power control technique for the coexistence of an NGSO satellite link with another NGSO/GSO satellite link for both downlink and uplink scenarios. The proposed technique is used to provide the required Signal-to-Noise Ratio (SINR) at the receiver and mitigate the in-line interference between the GSO and NGSO systems. In [17], the cognitive power control from [16] was enhanced based on the distance between the NGSO satellite and the NGSO earth terminal. The focus of this study is on the mitigation of downlink in-line interference from the NGSO satellite to the static GSO earth terminal in Ka-band and improving link quality for the NGSO system. In [18], the operation of the cognitive network with GSO and Low-Earth Orbit (LEO) broadband systems is studied in the downlink Ka-band, where LEO satellites are incumbent users, whereas GSO satellites are secondary users. To enhance spectral efficiency and protect the incumbent system, an optimization algorithm based on beam hopping and adaptive power control techniques is proposed. While the proposed solutions offer coexistence for dual satellite systems, none of them can be applied to simultaneously improve spectrum utilization, mitigate interference, and enable coexistence in the most complex scenario where GSO and NGSO satellite systems (with static or mobile ESs) and terrestrial networks operate in the same frequency bands.

Reference [19] analyzes the case where interference is caused to GSO systems by terrestrial cellular systems and NGSO systems respectively, both in the downlink and uplink Ka-band. The protection area is calculated, where no cognitive users (terrestrial system or NGSO system) are allowed to transmit, whereas, outside the protection area, the cognitive users are allowed to transmit concurrently with GSO systems. However, it is worth noting that this paper analyzes two separate scenarios. In the first one, GSO is protected against NGSO interference, while in the other, against terrestrial interference. Therefore, coexistence is again not analyzed for the complex scenario where GSO, NGSO, and terrestrial systems operate together in the same frequency band.

Authors in [20] address the cognitive GSO uplink in the Ka-band where satellite terminals reuse frequency bands of FS terrestrial MW links which are the incumbent users. In order for the interference impact of GSO satellite on FS links to stay within the regulatory interference limitations, a joint power and carrier allocation strategy was proposed, which requires the existence of a complete and reliable FS database. Further extension of study in [20] can be found in [21] where joint power and carrier allocation technique was proposed for uplink coexistence of GSO satellite with terrestrial FS in Ka-band. This technique was enhanced with a bandwidth allocation scheme that allocates bandwidth according to the user rate demands. This paper also analyzed the downlink coexistence of GSO and FS links in the Ka-band, where a joint beamforming and carrier allocation was introduced. Proposed strategies try to ensure the protection of the terrestrial FS system while maximizing the throughput of satellites. Availability of the FS database is assumed, which consists of FS antenna location and pointing directions.

In the CoRaSat project [22], Ka-band spectrum sharing between either single GSO or single NGSO and terrestrial users is mainly based on database techniques together with supplementary cognitive techniques as spectrum sensing. A centralized database and interference modeling is used to obtain the cognitive zones. A cognitive zone is defined as the area in which, for a given interference threshold, inside this zone, opportunistic operations are only allowed with the use of a cognitive technique to reduce the interference. Additional cognitive techniques only need to be applied within obtained cognitive zones, while outside these zones entities in the network can transmit without restrictions. Cognitive zones are also applied in [23] for dual satellite FSS system in 17.3–17.7 GHz band. The cognitive zone around incumbent broadcasting satellite service feeder links is determined by employing the characteristics of the links, which are stored and obtained from databases. Similar to the CoRaSat project, the cognitive FSS terminals can freely utilize the same frequency band outside cognitive zones. In the FREESTONE project [24], the possibilities to use frequency sharing techniques in SATCOM are investigated and several application scenarios and use cases are defined. One of the use cases is the coexistence of terrestrial FS and satellite FSS systems in 17.7–19.7 GHz Ka-band. A database design is proposed, taking into account characteristics of FS stations such as transmission powers, antenna parameters, and heights. The database requires information from the national FS registry and information about the secondary FSS terminals. Based on the database, it is calculated where and how the FSS users are allowed to operate in the band. Authors in [25] extend on scenarios evaluated in the FREESTONE project and analyze sharing between GSO and NGSO satellite in Ka-band. They describe how a spectrum database could be utilized in controlling and assisting coexistence in such a scenario. It is required that systems' main operational characteristics are stored in a spectrum database, such as frequency allocations, orbital positions, antenna patterns, etc.

It is worth noting that the introduction of dynamic entities like NGSO satellites or mobile ESs makes the use of a database approach very challenging [26]. Therefore, while solutions based on centralized databases may be appropriate for scenarios with a limited number of NGSO satellites (as assumed in all previously reported work), they might not be scalable to support upcoming large NGSO constellations. The downsides of approaches based on the centralized database are further investigated in Section 5.4. Similar to other studies on satellite-terrestrial and dual satellite coexistence, studies on coexistence techniques relying on databases have not yet addressed the problem of GSO, NGSO, and terrestrial networks simultaneously operating in the same frequency band.

5.3 Use case description and problem statement

The focus of this chapter is on the opportunistic transmission of data using satellite links operating at the Ka-band. The terrestrial users coexisting with the satellite systems are considered as the incumbent users, whereas GSO and NGSO satellites are secondary users with NGSO satellites having a lower priority than GSO satellites. In this work, we do not consider changing the spectrum management policies for the incumbent, but rather developing policies for GSO and NGSO entities so that the interference towards the incumbent is minimized. Therefore, only the satellite systems have DSS capabilities, whereas the terrestrial 5G Fixed Wireless Access (FWA) and MW links employ fixed spectrum allocation. The two most relevant and challenging use cases were selected: 1) opportunistic sharing of dual satellite and terrestrial systems in downlink Ka-band; and 2) opportunistic sharing of dual satellite and terrestrial systems in uplink Ka-band. These two use cases were selected based on the following criteria: 1) possible gains based on an initial qualitative evaluation; 2) importance and expected impact of the use case for the SATCOM industry; 3) implementation complexity; 4) need for new hardware support on satellites; and 5) need for new infrastructure on Ground Stations (GSs) and ESIMs.

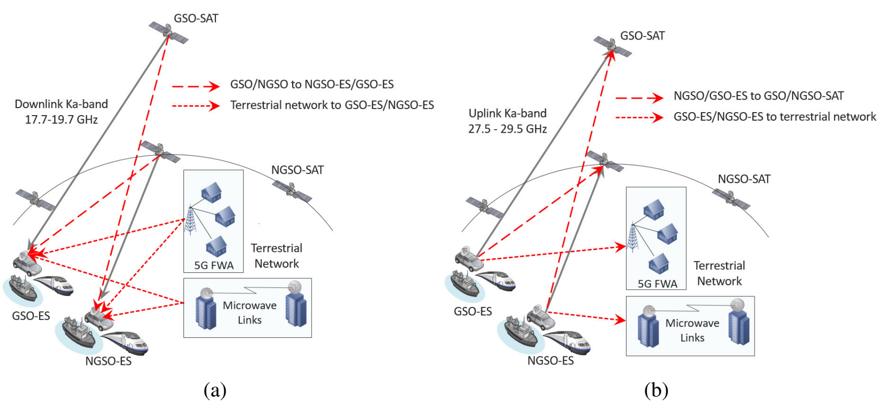


Figure 5.1: Selected use cases. (a) Downlink Ka-band. (b) Uplink Ka-band.

5.3.1 Opportunistic sharing of dual satellite and terrestrial systems in downlink Ka-band

This use case concerns GSO and NGSO satellite systems and terrestrial systems in the downlink Ka-band (17.7–19.7 GHz) and is shown in Fig. 5.1a. The terrestrial systems are mainly 5G FWA and MW links and it is assumed that they use Frequency Division Multiple Access (FDMA) for their transmission. The geostationary Earth Stations (GSO-ES) and non-geostationary Earth Stations (NGSO-ES) are mobile ESIMs and they represent trains, ships, cars, etc. Further, we consider that a control channel (CC) exists between different entities (satellites, ESIMs, GSs, and terrestrial networks). Since all the geostationary satellites (GSO-SAT) and non-geostationary satellites (NGSO-SAT) and the terrestrial network operate in the same band, their transmissions can have an impact on the ESIMs. For the downlink operation of GSO and NGSO satellites in the presence of the terrestrial network, the following interference components are considered:

- GSO-SAT towards NGSO-ES;
- NGSO-SAT towards GSO-ES;
- Terrestrial network towards GSO-ES;
- Terrestrial network towards NGSO-ES;
- GSO-SAT towards terrestrial network;
- NGSO-SAT towards terrestrial network.

The goal here is to protect the communication of incumbent terrestrial networks, manage the interference towards GSO and NGSO ESIMs coming from GSO-SAT, NGSO-SAT, and the terrestrial network and improve the performance (spectrum utilization, throughput, PER, latency, etc.) by applying DSS techniques.

5.3.2 Opportunistic sharing of dual satellite and terrestrial systems in uplink Ka-band

This use case deals with GSO and NGSO satellite systems and terrestrial systems operating in the uplink Ka-band (27.5–29.5 GHz) as shown in Fig. 5.1b. Similar to the downlink case, the terrestrial systems represent 5G FWA and MW links and it is assumed that they use FDMA for their transmission. Also, the assumption of a CC between different entities (satellites, ESIMs, GSs, and terrestrial networks) is applied here. Because of the spectrum sharing and possible spatial alignment of the ESIMs, it becomes possible that GSO, NGSO, and the terrestrial network are impacted by the transmission of the ESIMs. For the uplink case, we consider the following interference components:

- NGSO-ES towards GSO-SAT;
- GSO-ES towards NGSO-SAT;
- GSO-ES towards terrestrial network;
- NGSO-ES towards the terrestrial network.

The goal here is to protect incumbent terrestrial networks, manage the interference towards GSO and NGSO satellites and the terrestrial network coming from the ESIMs and improve the performance (spectrum utilization, throughput, PER, latency, etc.) by applying DSS techniques.

5.4 Proposed approach and requirements

In this section, we first motivate why we designed a distributed approach for applying DSS techniques and how this compares with more adopted centralized approaches. Further, we define what are the main requirements and constraints for applying distributed DSS techniques on existing satellite/terrestrial systems.

5.4.1 Distributed versus centralized approach

In our work, we have adopted a distributed approach for dynamic spectrum access, as opposed to current more widely accepted centralized database approaches (such as, for instance, Citizen Broadband Radio Service (CBRS) deployment in the United States [27]). The motivation to adopt a distributed approach rather than a centralized approach in this research is driven by the following observations:

- *Driver*: the main driver for a centralized database approach is to protect incumbents (primary users) while allowing opportunistic use of the incumbent's spectrum by secondary users as long as the incumbent does not experience any harmful interference. The main drivers for a distributed receiver-based approach are scalability, by only reacting upon interference events detected at the receiver, and spectrum efficiency, by exploiting spatial reuse of spectrum.

- *Mechanism*: a centralized approach grants access to opportunistic users in a centralized way by consulting a global database that relies on transmission models of transmitters and the definition of exclusion zones (geographical areas within which active opportunistic radio transmitters are not allowed), restriction zones (geographical areas within which active opportunistic radio transmitters are allowed under certain restrictive conditions, limited by Equivalent Isotropic Radiation Power (EIRP) or EPFD), and protection zones (geographical areas within which incumbent receivers will not be subject to harmful interference). The distributed approaches that we adopted only take an action when interference is detected at the receiver. As interference only happens at the receiver of a communication link, the receiver is the best place to detect any possible interference, in other words, a receiver can measure the 'ground truth'. A global centralized database, which is built based on transmitter models, cannot give more accurate information on the impact of interference than the local view as measured at the receiver. A model is by definition an approximation of the real world and always falls short of the complexities of reality. It is, for instance, very hard to accurately model multipath signal propagation, out-of-band emission, antenna properties (gain, side-lobe, back-lobe emission), Doppler shift, obstacles, weather conditions, etc.

- *Spectrum monitoring*: both approaches can be augmented with spectrum monitoring information. In the case of centralized approaches, geographically distributed spectrum monitoring devices are employed. A big issue here is accurate Radio Frequency (RF) calibration and abstracting hardware-specific features and artifacts of the spectrum monitoring device. Ideally, these devices should be capable of measuring the spectrum occupation for different directions around the spectrum monitoring device, increasing the complexity of the device. For comparison, in distributed approaches, when the spectrum is monitored at the receiver of a communication link using the same antenna configuration of the receiver, it will only capture relevant interference signals. There is further no need for accurate RF calibration, as threshold levels for harmful interference can be perfectly calibrated through correlation with link statistics.

- *Rules for spectrum sharing*: to compensate for inaccuracies and deficiencies of the models applied in a centralized database, large safety margins are applied, leading to very conservative rules for spectrum sharing and hence largely limiting the benefits thereof. In distributed approaches, resources are dynamically allocated, triggered by interference events, which are detected by monitoring link properties at the receiver. The receiver is not only the best place to detect any possible interference, but also the best place to make decisions on the spectrum reallocation of the interfered communication link.

- *Reaction time*: consulting a centralized database for taking decisions on spectrum resource allocation puts serious limitations on the reaction time, compared to local decisions using a fast dedicated CC.

- *Scalability*: centralized database approaches are not scalable when the number of involved entities is scaling up, due to the huge information exchange overhead to keep the database up to date. The number of exchanged messages that are needed to keep centralized databases up to date will further explode when dynamic entities are involved, which is the case with NGSO-SATs, ESIMs, and mobile users communicating with terrestrial systems. An increase in the dynamicity of the system may lead to outdated information stored in centralized databases. On the other hand, in distributed approaches, information exchange overhead between entities of satellite/terrestrial networks does not increase with the increase of network density. In this approach, information sharing between entities is event-based, information is exchanged only when required, which reduces control overhead and does not lead to performance degradation when network density increases.

5.4.2 Requirements

Since this research is aiming to redefine the framework of future SATCOM in order to achieve the highest possible spectrum efficiency, we are encouraged to identify and propose which aspects of SATCOM could be enhanced to be able to support dynamic spectrum slicing and allocation. Based on these considerations, the following requirements are identified throughout this work:

- A basic Multi-Frequency Time Division Multiple Access (MF-TDMA) Medium Access Control (MAC) protocol is considered to be available and running in all satellite-related entities. This is the most basic enabler of spectrum slicing in its two dimensions (frequency, time) that is adopted by the broader wireless community (also in 5G) for future dynamic and spectrum efficient resource allocation. Although current satellite systems use Frequency Division Duplexing (FDD) and further Time Division Multiplexing (TDM) in downlink and MF-TDMA in uplink as in 2nd Generation Interactive Digital Video Broadcasting Satellite System (DVB-RCS2), such a continuous TDM stream will be a major source of interference and this approach is not sustainable with the anticipated growth of satellite density. Therefore, we consider MF-TDMA for both downlink and uplink and elaborate its benefits in terms of spectrum utilization and interference avoidance capabilities. The baseline scheme that we considered in this work is equivalent to the current satellite standards, more precisely to DVB-RCS2 (TDM in downlink and MF-TDMA in uplink) [28].

- For the sake of simplicity, FDMA is considered for both 5G FWA and MW links. Note that the intention here is not to simulate the full stack of the terrestrial part, but rather to consider it as the incumbent.

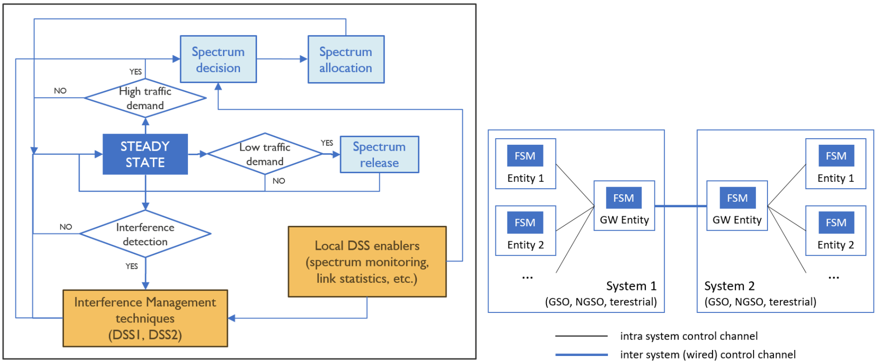
- There is time synchronization between all entities taking part in the MF-TDMA based DSS framework with an accuracy at least in the order of 1 up to 2 ms, requiring that the distance between two extremes of the satellite's beam, projected on the ground, is within bounds that allow the level of synchronization assumed. If terrestrial networks want to actively share spectrum in the time domain through the MF-TDMA MAC, then they also need to adhere to this time synchronization requirement.

- Each transmitter is capable to actively shift its transmission bursts in frequency and in time in order to make sure that the receiver receives the transmission within the intended frequency range and time slot.

5.5 DSS techniques and generalized Finite State Machine

In this section, we discuss the DSS techniques that have been implemented and evaluated in selected use cases. We first introduce a generalized Finite State Machine (FSM) that works as the core of the DSS techniques followed by a brief description of the two DSS techniques: 1) DSS1 – collaboration protocol (CP); and 2) DSS2 – decentralized spectrum sensing.

The DSS techniques we designed are generic and can be applied to other scenarios and in different satellite bands (such as L, C, X, Ku, etc.) than the two selected use cases operating in the Ka-band. More specifically, they can be applied in situations where: 1) interference management is an issue due to the coexistence of other satellites/operators in the same spectral band; 2) improving spectrum utilization is required; and 3) distributed interference management is required. As the proposed DSS techniques are generic and agnostic to Physical layer (PHY) technology (spectrum band, modulation, and coding techniques) and topology/constellations, they can be employed for solving the foreseen interference management and spectrum utilization problems in various use cases.



(a) (b)

Figure 5.2: (a) DSS Finite State Machine (FSM) running on satellite/terrestrial entities. (b) CCs between system entities (GSO systems, NGSO systems, terrestrial networks) with FSM.

Fig. 5.2 shows the generalized FSM and the generic entities, including gateway (GW) for the different systems considered in this study: GSO, NGSO, and terrestrial networks. Each entity runs an FSM that can collect local information such as spectrum occupation data (collected via spectrum monitoring), node position, antenna characteristics, etc., intra-system information from other entities operating within the same system as well as inter-system from entities operating in other systems. This information will be shared through established intra-system or inter-system CCs, depending on whether the entities communicate within or across systems. For instance, if the FSM operates on a GSO-ES, then it requires information from other GSO (satellite, ES), NGSO (satellite, ES), and terrestrial systems through their respective GSs/GWs and CCs.

The FSM includes the following building blocks: steady state, spectrum decision module, spectrum allocation module, spectrum release module, interference management techniques, and local DSS enablers. The starting point of the FSM is the steady state. At the bootstrap, we assume that there is less dynamicity and the system starts in the steady state. This steady state operation could be achieved through an initial fixed spectrum assignment. While the system is in the steady state, as the dynamicity increases the following events could happen: 1) low traffic demand; 2) high traffic demand; and 3) interference detection. In case of low traffic demand from the application side, the spectrum release module will be executed which will release some spectrum resources so that those resources become available for other entities, and the system will resume its operation in the steady state. In case of high traffic demand from the application side, the spectrum decision and spectrum allocation modules will be executed which will allocate new spectrum resources for meeting the traffic demand. These low and high traffic demand events will be identified by inspecting the data queues for incoming packets from the application layer. The job of the spectrum decision module is to find out the most suitable slots/channels to be used by GSO/NGSO satellites and ESIMs based on different channel characteristics. It is the receiver of a communication link that proposes suitable slots/channels to the sender by utilizing: 1) local DSS techniques such as spectrum monitoring; and 2) transmission characteristics received from other entities that might interfere. The job of the spectrum allocation module is to implement different spectrum allocation and access mechanisms for: 1) protecting incumbent terrestrial systems; 2) satisfying the priority of GSO over NGSO systems; and 3) preventing interference among different GSO and NGSO transmissions. Lastly, the most important event from the steady state is interference detection where interference can be detected in a variety of ways such as SINR degradation, interference sources detection, etc. All active links are monitored when the system is operating in the steady state, and in case of detecting an interference event, the interference management techniques, local DSS enablers, and spectrum decision module will be executed in order to move the system back to the steady state. Two DSS techniques have been designed for managing the interference and support flexible spectrum (re)allocation and release. These two interference management techniques are based on CP (called DSS1) and spectrum sensing (called DSS2).

5.5.1 DSS1 - collaboration protocol

DSS1 is based on a decentralized CP that enables information dissemination between different node entities (within the same network or across different networks) that are taking part in the collaboration. The CP is used for advertising information such as node positions, transmission characteristics of active transmitters, etc. (see Section 5.5.1.2). This approach aims to provide each node with the necessary and relevant information about current spectrum usage and active transmissions from other nodes within its interference range. Having this information, a node can make informed decisions on its spectrum usage and avoid interfering with other nodes or getting interfered by others.

We further enhanced the CP protocol in DSS1 to support an optional alarm message. This alarm message is initiated by terrestrial nodes and advertises a problem when suffering from interference at the receiver. This allows satellite transceivers with link adaptive capabilities to protect terrestrial links that cannot adapt. Once the FWA link detects interference, it notifies all the ESIM and satellite transmitters participating in CP to reallocate slots that might be causing this interference. Upon reception of the alarm message, each entity in the network should filter it based on the active transmissions and react if necessary. The interfering satellite link will reallocate slots that directly overlap with slots allocated by the FWA link. In addition, also slots that might cause Adjacent Channel Interference (ACI) are reallocated. DSS1 technique based on CP enhanced with alarm messages is referred to as DSS1+ in the rest of the chapter.

In DSS1 and DSS1+, the FSM shown in Fig. 5.2a is executed at GS, GW, and/or satellite or ESIM. All the states explained in the generalized FSM remain the same and the only difference is that DSS1 relies on disseminating required information via the CP. The main advantage of DSS1 is that it does not require any changes to the existing operation of satellite systems, and only requires the introduction of the CP that can run in the GSs controlling the satellites and ESIMs or the GWs controlling terrestrial links.

5.5.1.1 Network architecture of collaboration protocol

The proposed CP network architecture with multiple systems consisting of different entities and established user and control links is depicted in Fig. 5.3. As the focus of this work is not on feeder links, they are not presented here, however, all the approaches presented here can be extended to feeder links as well. The CP disseminates its CP messages between involved entities over a wired terrestrial backbone network via their GSs and GWs, supported by existing wireless CCs.

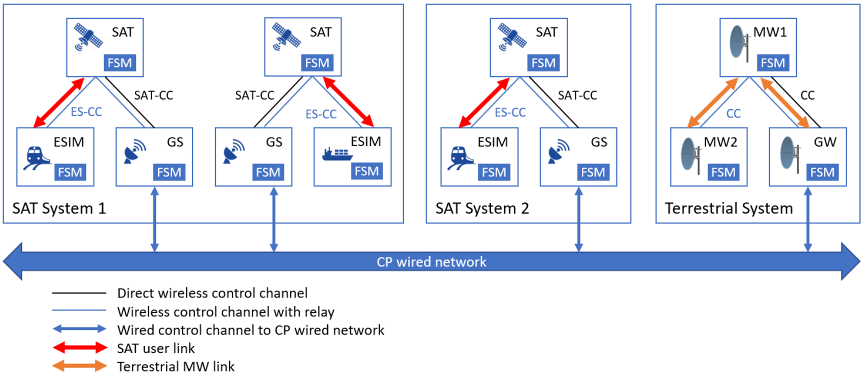


Figure 5.3: Network architecture of CP consisting of different entities with their respective user and control links.

The CP can be terminated in the GS and pass any decision to the satellite or ES via standard control messages over standard CCs: in the case of satellites, there is direct satellite control channel (SAT-CC) between GS and satellite, whereas, in the case of ES, the Earth Station control channel (ES-CC) is relayed between GS and ES via the satellite. In this case, the GS will act as a CP proxy for the satellite and ES. Alternatively, the CP can also run on the satellite and ES to take local decisions by extending the standard CCs (SAT-CC and ES-CC) to pass special CP messages. The CP network is also connected to the terrestrial network, where GWs have a direct connection to the CP wired network, while other MW terminals will communicate via a terrestrial wireless CC (possible via relaying over MW mesh network).

Minimizing the overhead of the CP protocol in the wired terrestrial Internet infrastructure is not within the scope of this study. There are multiple well-known techniques for minimizing dissemination overhead like the publish-subscribe paradigm or filtered controlled flooding that can be selected to accomplish this target. We however make sure that the overhead over the wireless control link is minimized, and this is part of the evaluation metrics.

5.5.1.2 Collaboration protocol structure

CP is used to disseminate at least the following type of messages:

- Position of a node (mandatory message, generated by all transmitter nodes or their respective GSs/GWs);
- Active transmission characteristics (mandatory message, generated by all transmitter nodes or their respective GSs/GWs);

- FWA alarm message in case of detected interference (optional message, generated only by receivers in terrestrial networks or their respective GWs).

The position of a node can be defined in multiple ways, depending on the coordinate system employed. In our study, we used Earth Centered Earth Fixed (ECEF) coordinate system [29]. For ECEF, three parameters x, y, z are needed to give the position of an object. Each node advertises its location periodically. To reduce the overhead of the GS-satellite control link we assume that satellites rely on their GSs to advertise their position in the CP network.

Regarding the active transmission characteristics, it will describe any active transmission from any interface of a transmitting node. For the transmission to be clearly described and receivers to be able to calculate the possible interference to them, the following info needs to be included in such a message for each interface:

- Transmission power setting of the interface;
- Destination focal center point coordinates;
- Gain as a function of transmission cone azimuth and elevation;
- Time information about the transmission (continuous, periodical, start and end of transmission, allocated MF-TDMA slot[s], etc.).

The CP architecture is designed so that each active satellite and ESIM should have all relevant information only about transmissions that could affect them, their power levels, their timing information and therefore could build a synthetic interference map.

The FWA alarm message should contain the location of the receiver node and time-frequency resource usage information (i.e. utilized MF-TDMA slot[s]).

5.5.2 DSS2 - spectrum sensing

The second DSS technique (DSS2) is based on decentralized spectrum sensing and it requires spectrum monitoring capabilities at each receiver taking part in the satellite network. There exist technical solutions to achieve continuous runtime wide spectrum sensing without the need for a time-sharing approach between normal operation and spectrum sensing. Hence, each node can know the spectrum usage situation present in the receiver's immediate vicinity. By processing the raw I/Q samples, local energy maps can be created that can be employed for decision-making on spectrum allocation. In DSS2, the FSM from Fig. 5.2a that includes spectrum sensing will be executed at the receiving entities.

DSS2 does not require a CP or any exchange of state/status data between different communicating entities, it only requires a dedicated CC between the communicating entities for slot (re)allocation or release. Upon detection of an interference event, each receiving entity decides locally which free slots are available for mitigating interference based on locally acquired spectrum monitoring information (local energy map). Then, the receiver will alert the transmitter and initiate a slot reallocation procedure over the dedicated CC to replace the interfered slot with the new slots. Whereas DSS1 can only detect known interference from CP advertisements, DSS2 can detect any interference source at the receiver. However, DSS2 requires changes to the radio modem to integrate spectrum monitoring at the receiver.

5.5.3 Link setup procedure in DSS techniques

For a link to be established between two entities in our selected use cases, we assume that a CC is always available between two entities trying to set up a link between them, meaning for all possible cases between ESIM, GS, and GSO/NGSO satellites. By extending dedicated control links that already exist today (SAT-CC and ES-CC), through the wired infrastructure, a slot allocation for either uplink or downlink can happen through the control link.

Here we shall describe how a satellite-GS link or satellite-ES link will be established using the information provided from the CP and the synthetic interference map in DSS1 or local energy maps in DSS2. As an example, let us assume that the link will be an uplink connection from GS or ES to the satellite. At the time that the link needs to be established, the GS/ES transmitter will alert through SAT-CC or ES-CC that it wants to set up a link with X slots (X will be determined from the incoming traffic volume and can be an estimation that can be adapted later on the lifecycle of the link) towards the destination satellite. Then the satellite will examine its synthetic interference map or local energy map and based on the interference levels depicted, will choose X interference-free slots and send them back to the GS/ES through the SAT-CC or ES-CC.

For DSS1, in case the CP information is not forwarded to the satellite or ES but is kept on the GS (serving the satellite or ES), the GS will decide on behalf of the satellite or ES based on the available information in the synthetic interference map and select the interference-free slots, informing the satellite or ES through the SAT-CC or ES-CC to switch its receiver into the selected slots. In any case, for DSS1 the synthetic interference map will be the main tool for picking up new MF-TDMA slots that are interference-free and can be allocated to the new link. In DSS2, by having a clear view of the channel occupation on the receiver side, all present sources of known or unknown interference can be avoided and therefore links can be set up with optimal link characteristics.

By defining a way where transmitter and receiver of a new or pre-existing link can negotiate through a control link and reach a final agreement on slot allocation (employing their synthetic and local energy maps), the proposed DSS techniques can offer a completely decentralized solution of dynamic spectrum allocation and minimize interference.

5.5.4 Interference detection and avoidance procedure in DSS techniques

The procedure of interference detection and initiation of an interference avoidance procedure will be presented in this section. Following a similar approach in all proposed DSS techniques for interference detection, we used PHY metrics like SINR and MAC metrics like Packet Success Rate (PSR) falling below a predefined threshold as an interference event trigger. Similar to the link setup procedure, the interference avoidance procedure for either downlink or uplink is executed based on the existing control link infrastructure, with minimal extensions applied.

Let us analyze an example where a satellite-GS link or satellite-ES link is already established and starts to experience interference. At the time an interference event is triggered, the nodes participating in the link must take some actions to correct the link quality. Since it is the receiver that will identify the interference event, it is up to the receiver to consult its synthetic interference map in DSS1 or local energy map in DSS2 and pick new free and non-interfered slots to replace the ones that are getting interfered. Then the receiver will alert the transmitter node to alter the schedule of the transmission by removing the interfered slots and enabling the new MF-TDMA slots. In case the GS is responsible for making the decision on spectrum allocation, the satellite will just inform the GS about the interference event detection and then the GS will decide to move the interfered slots to other time-frequency slots. Next, the GS will inform the satellite about the newly chosen slots. In both cases, the link will recover by abandoning the interfered slots and replacing them with new interference-free slots.

Let us also note that in case there is no MF-TDMA in place but just an FDMA approach utilized on the link, the same link setup and interference avoidance procedures can be applied for channel selection instead of slot selection.

5.6 Simulation setup and results

First, the simulation scenario, baseline reference scenario, adopted assumptions, and selected KPIs are presented. Afterward, the DSS techniques are compared to a baseline reference scenario with fixed spectrum allocation by running an extended set of simulations. Simulations were performed in a satellite environment based on the components and functionalities being reused and extended from Satellite Network Simulator 3 (SNS3) [30]. SNS3 is a satellite network extension based on the well-known ns-3 simulator environment [31].

5.6.1 Simulation scenario

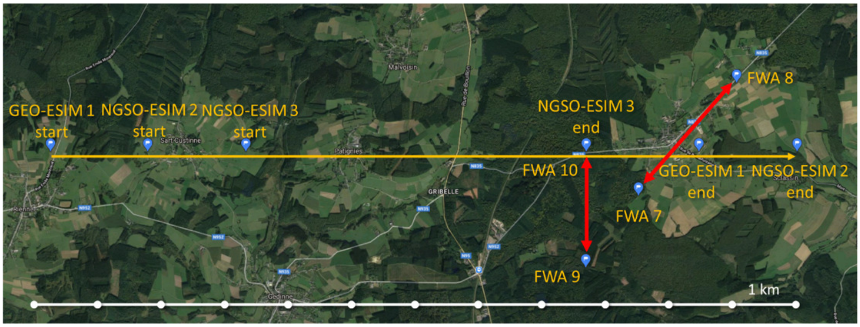
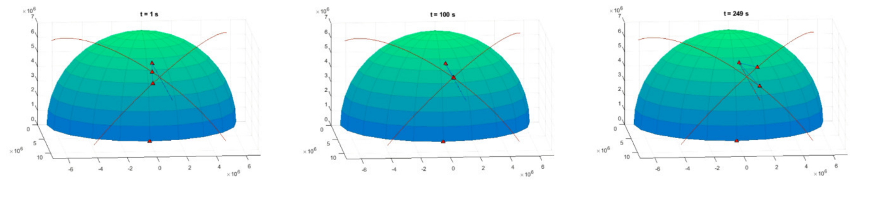


Figure 5.4: Simulation scenario on Earth's surface.

Fig. 5.4 shows the simulation scenario on Earth's surface. It includes three ESIMs (one GSO-ESIM and two NGSO-ESIMs) communicating to three satellites. ESIMs are moving in a straight line with variable speeds of 25–55 m/s (i.e., between 90 and 200 km/h, representing a range of typical train speed) between the 'START' and the 'END' points. The reason for having three ESIMs in the collision domain (overlapping of satellite footprints) of the GSO-SAT and the two NGSO-SATs, and later on in the collision domain of the FWA links when they come closer to the FWA nodes, is to demonstrate at minimum the performance of the DSS techniques and how they manage to solve the interference when spatial separation is not possible. Since the proposed DSS techniques operate in a distributed manner, the schemes can scale well to a large number of ESIMs by exploiting time, frequency, and spatial dimensions to avoid interference. Further, there are 2 NGSO orbits at an altitude of 2000 km, an inclination of 50° from the Equator, and ascending nodes at 27.38 W and 142.61 W. In each orbit, only 1 NGSO-SAT is considered, as the scenario lasts 250 seconds and there is no need for link roaming between satellites. Both the NGSO-SATs point towards their respective focal points around 50 N, 5 E, covering an area of 70-75 km around their focal points on the Earth's surface. Fig. 5.5 shows the trajectory of the 2 NGSO-SATs and their positions at times 1, 100, and 249 seconds. There is one GSO-SAT at an altitude of 36,000 km and it associates with one ESIM on the Earth's surface. There are two point-to-point MW FWA links: 1) FWA 7-8; and 2) FWA 9-10. The first FWA link is parallel to the surface of the Earth, whereas for the second link the elevation angle from FWA 9 to FWA 10 is 32° aiming to provide an alignment with the GSO-SAT.

There are a couple of main interference events we aimed to create in this scenario:

- Alignment between ESIMs, NGSO-SATs, and GSO-SAT occurs around the



(a) (b) (c)

Figure 5.5: Position of the 2 NGSO satellites. (a) At 1st second. (b) At 100th second. (c) At 249th second.

100th second of the simulation.

- ESIMs getting very close to FWA 10 and also aligning with the FWA 9-10 link.
- ESIMs passing through the FWA 7-8 link while also being very close to both transmitter FWA 7 and receiver FWA 8.

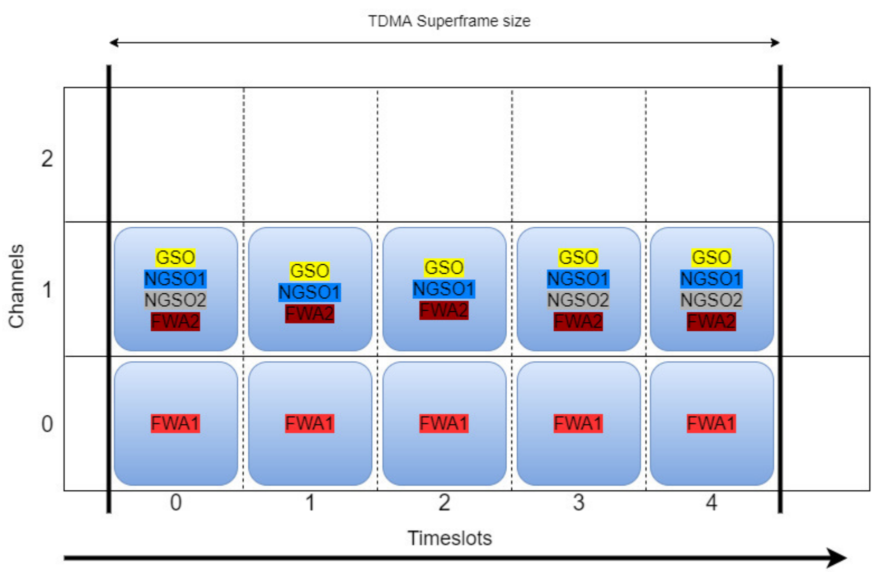


Figure 5.6: Initial spectrum allocation status.

All satellite links start the scenario by using channel 1. FWA 7-8 and 9-10 links use channels 0 and 1, respectively. By assigning the same channel to multiple links, the advantage of spatial separation of all nodes at the start is exploited while enabling the detection of interference during the scenario in order to study the impact of the proposed DSS techniques. The initial spectrum allocation status is presented in Fig. 5.6.

5.6.2 Baseline

With DSS techniques applied in satellite-terrestrial deployments, entities are able to move out of initially allocated slots. However, for benchmarking purposes, a baseline scenario without DSS is defined. For this reference scenario, all DSS mechanisms are disabled for the GSO and NGSO links, and static spectrum usage is assigned. Therefore, in the baseline, satellite-terrestrial systems keep initial spectrum allocation status throughout the simulation, in which all the involved entities including GSO-SAT, NGSO-SAT, ESIMs, and terrestrial networks operate in the same frequency bands. As motivated in Section 5.4, a basic MF-TDMA MAC protocol is considered for both downlink and uplink, running in all satellite-related entities. However, for the baseline scheme, restrictions are imposed on the MF-TDMA MAC scheme to comply with the current DVB-RCS2 satellite standard (TDM in downlink and MF-TDMA in uplink).

The main goal of the proposed DSS solutions is to increase the spectrum utilization performance if compared to the baseline without compromising on the protection of the incumbent (terrestrial network) and the priority of GSO-SATs over NGSO-SATs.

5.6.3 Assumptions

For the GSO and NGSO links, we consider up to four application flows with different throughput demands, which are enabled or disabled during the scenario. Therefore, the simulation scenario involves dynamically changing application loads for all satellite links. For the two FWA links, we consider a continuous application flow. The application layer packet size was selected to match the size of the PHY Protocol Data Unit (PDU) (64800 bits) to avoid fragmentation and aggregation requirements as this was not the focus of this study. Specifically, the UDP/IP/MAC headers have a size of maximum 54 bytes, so the application layer packet was set to 64368 bits. It must also be noted that each MF-TDMA slot supports sending a maximum of 3 PPDUs and that the selected Modulation and Coding Scheme (MCS) for all satellite communications was 16APSK, 2/3 coding rate, and 0.1 roll-off factor. The selected MCS requires an SINR of 9 dB in order to offer successful decoding of a PHY frame. Therefore, all satellite-terrestrial links are designed to have a reception threshold of 9 dB SINR, below which they lose connection. The nominal level of SINR is set to around 10 dB for satellite links and around 20 dB for FWA links, offering a safety margin of 1 dB and 11 dB for satellite and FWA links, respectively.

Several other assumptions have been made to simplify the simulation scenario without heavily impacting the results:

- The simulations consider antenna properties (main lobe, sidelobes, backlobes, beam tracking, co-channel interference, and ACI) that are suggested in International Telecommunication Union (ITU) documents ITU-R S.672-4 [32], ITU-R S.1528 [33], and ITU-R S.580-6 [34]. ACI is modeled at -27 dB attenuated signal power compared to main channel power.

- The simulator includes models for Received Signal Strength Indicator (RSSI) measurements and spectrum monitoring.

- For the reaction time of collaboration and resource allocation protocols, propagation delays are considered.

- Perfect modem synchronization, meaning zero time to reconnect after interference is solved.

- All satellite systems are considered single-beam systems.

- Feeder links are not simulated.

- A single carrier is adopted per PHY channel.

- ESIM antennas track the connected satellite.

- GSO/NGSO focal points are fixed.

- NGSO satellites are assumed to be LEO satellites.

- During the baseline scenario finetuning, a satisfactory PER threshold value for initiation of interference avoidance procedure is determined to be above 0.01%.

- The default MF-TDMA superframe structure of 5 time slots and 3-5 channels is applied for all simulations.

- The channel bandwidth of each channel is 12.5 MHz, consisting of actual 10 MHz signal usage and 2.5 MHz guard space. Guard spaces provide Doppler Effect protection, supporting speeds of ESIMs up to 2500 km/h.

5.6.4 Key Performance Indicators

In addition to SINR graphs, other KPIs measured through an extensive simulation study and used to evaluate the results are:

- Throughput is the aggregated application layer throughput achieved in each scenario run across all active links.
- PER (%) is the average application layer PER observed in each scenario run across all active links.
- Uptime (%) is the average percentage of simulation time that links were operational with 99.9% PSR and above across all active links.
- Spectrum utilization represents the percentage of allocated spectrum used for actual transmissions. Spectrum utilization of 100% means that all transceivers use their full allocated capacity (all allocated slots are used for transmission of 3 PPDUs).
- Latency is the average application layer latency observed in each scenario run across all active links.
- Wireless and wired protocol overhead of the DSS techniques in terms of information sharing via a control channel between GSO, NGSO, ESIM, and the terrestrial network.
  - CP overhead is the overhead per second induced from running the CP protocol across all wired and wireless links.
  - The slot selection protocol overhead is the overhead per second of the slot allocation protocol across all wireless links.

5.6.5 Results

All the KPIs are presented and analyzed for satellite and terrestrial links separately, in order to pinpoint the effects of DSS techniques on different networks in the system. In the end, 3 simulation sets were executed, two for uplink and one for a downlink scenario, each evaluating the two proposed DSS techniques and comparing them to baseline scenario. For a better understanding of the results, a detailed explanation of the circumstances leading to the results in the baseline of downlink and uplink is presented, whereas for the rest of the simulation runs, only comments are made where it is important to pinpoint specific aspects of the results.

5.6.6 Downlink use case

For the downlink use case, we analyzed the impact of NGSO/GSO downlink transmission and terrestrial links on the GSO and NGSO ESIMs. We also analyzed if FWA links need protection during downlink transmission. Further, we examined the ability of the proposed DSS techniques to provide interference protection to the satellite links and, if needed, to terrestrial links. In this simulation set, the available spectrum consists of 3 channels.

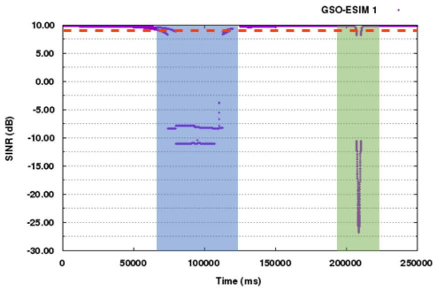
5.6.6.1 Baseline

Fig. 5.7 shows the SINR graphs of the FWA receivers, NGSO-ESIMs, and the GSO-ESIM. The FWA links are experiencing an SINR drop of 2.75 dB in the case of FWA 9 and no drop for FWA 8. Both FWA links remain operational though as the safety margin for correct reception SINR for them is around 9 dB. On the other hand, the GSO-ESIM receiver suffers on multiple occasions. In presented SINR graphs where measurements are dropping below the reception SINR threshold of 9 dB, this threshold is represented with a red dashed line. The first big dip below the reception SINR threshold of 9 dB (see dashed red line in Fig. 5.7a) is because of the alignment of all NGSO and GSO communication paths around the 100th second of the simulation run (see blue area). The second dip (green area) happens around the 210th second of the simulation time because of reaching the receiver FWA node 9. At that time the GSO-ESIM receiver is perfectly aligned to the signal of FWA node 9, impacting its SINR and dropping the connection.

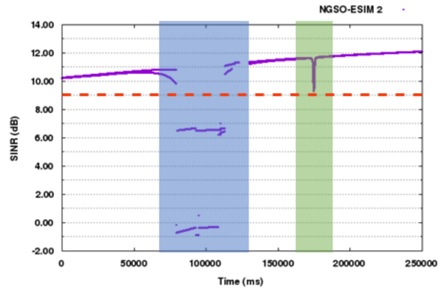
In SINR graphs of NGSO links, we can distinguish again the 2 major interference events, the NGSO-GSO geometry alignment at around 80th second (blue area), and then the FWA 9-10 link interference effect at around 170th second, as the ESIMs pass behind FWA receiver node 9. There is another slight impact for both of them around 170-200 seconds that corresponds to both of ESIMs crossing the second FWA link at different times. The impact of interference is quite higher in NGSO-ESIM 3 (yellow area) than in NGSO-ESIM 2 (green area) because NGSO-ESIM 3 actually stops at the FWA receiver and thus suffers interference for a longer time than NGSO-ESIM 2.

The fact we can see two or more lines per SINR measurement is because of the condensed measurement of SINRs at the ms scale. We can see multiple levels of SINR measurements that are grouped per slot corresponding to multiple interference events impacting the signal. For instance, some slots might get interfered differently than others based on their ACI and co-channel interference. Hence, we often observe two or more lines in most SINR graphs.

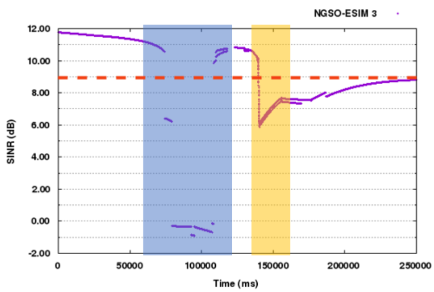
In Fig. 5.8a, the overall system PER graph is presented. It is clear that PER increases in the specific areas of the SINR graphs, where SINR drops below 9 dB for the GSO and NGSO links. NGSO 3 cannot recover after 120 seconds, as in



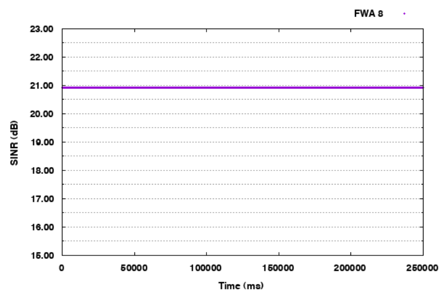
(a)



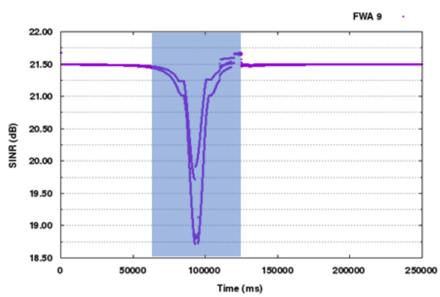
(b)



(c)



(d)

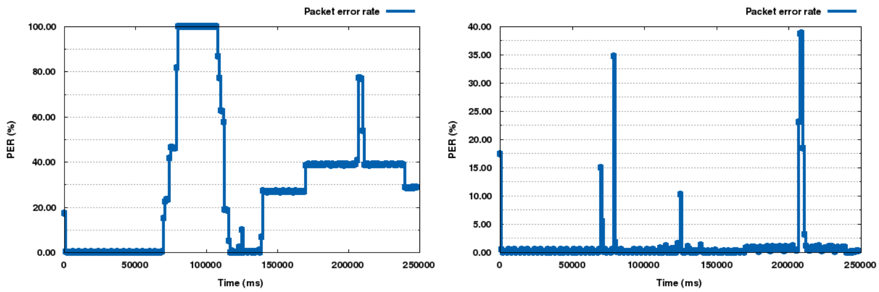


(e)

Figure 5.7: SINR graphs in the downlink use case (baseline). (a) GSO-SAT 1. (b) NGSO-SAT 2. (c) NGSO-SAT 3. (d) FWA 8. (e) FWA 9.

the scenario it is designed that it stops close to FWA node 9 at that time and never moves again, therefore it is clearly interfered for the remaining of the scenario from the FWA link.

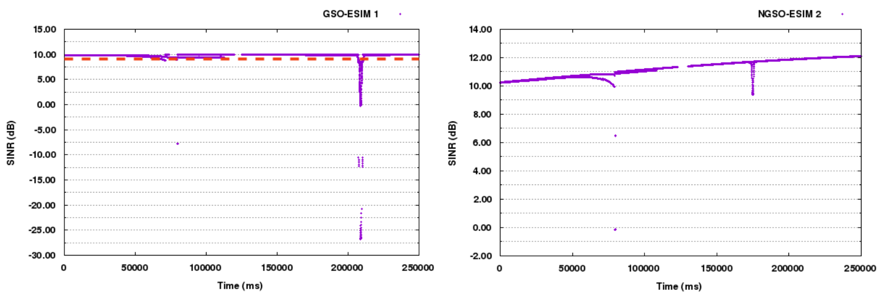
We continue now to present results from the proposed DSS techniques and compare them versus this baseline.



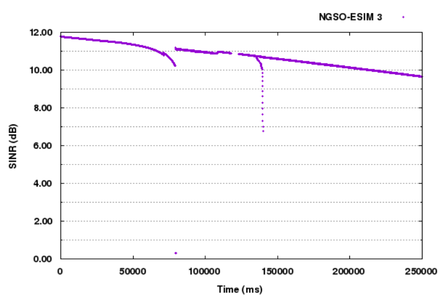
(a) (b)

Figure 5.8: PER graph for satellite links in the downlink use case. (a) Baseline. (b) DSS1.

5.6.6.2 DSS1



(a) (b)

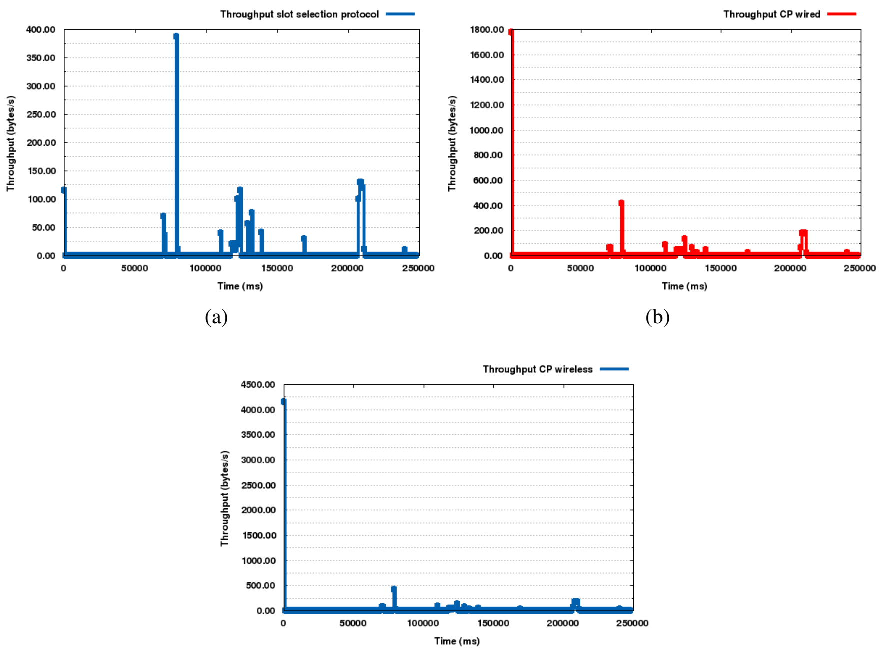


(c)

Figure 5.9: SINR graphs in the downlink use case (DSS1). (a) GSO-SAT 1. (b) NGSO-SAT 2. (c) NGSO-SAT 3.

Observing the SINR graphs of DSS1 in Fig. 5.9, we can conclude that DSS1 is capable of resolving almost all interference events in the satellite links since SINR drops below 9 dB for short periods before reacting and moving to free spectrum slots. There are, as it seems, enough slots for the three satellite links to not only avoid direct but also possible ACI from neighboring nodes. It is also clear that the nodes react only when the SINR drops below the error threshold. FWA node 9 in DSS1 is only losing 2.2 dB SINR, dropping from 22 to 19.8 dB, whereas FWA node 8 is only showing a loss of 0.1 dB in the worst case. We observed approximately 18.7 dB SINR in the baseline, marginally gaining 1.1 dB in DSS1 compared to baseline. Since FWAs were not impacted negatively and no big differences were observed in the results achieved with DSS, we did not present SINR graphs for FWAs. It is worth mentioning that in scenarios where FWAs links are heavily interfered by satellite communication, their protection would be guaranteed by applying the DSS1+ technique.

The PER graph explains even better how successfully DSS1 was able to solve the interference problems between the satellite links as shown in Fig. 5.8b. The PER only spiked at the start of each interference event, being resolved less than a few seconds later.



(a) (b) (c)

Figure 5.10: Control overhead in the downlink use case (DSS1). (a) Slot selection overhead (wireless). (b) Collaboration protocol overhead (wired). (c) Collaboration protocol overhead (wireless).

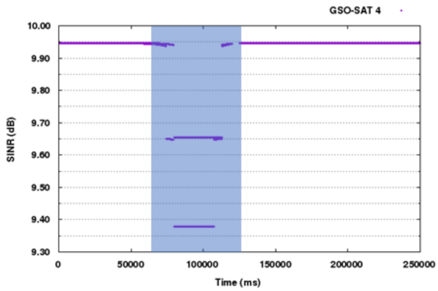
Finally, to show if DSS1 is costly in terms of control overhead, the overhead of the CP and slot allocation protocol is presented in Fig. 5.10. As it can be seen, both CP and slot allocation protocol overhead is minimal, only increasing when actions need to be taken to resolve interference events. Except for bootstrapping of the system where all nodes advertise their positions and their active links, after that time the maximum CP overhead, both wired and wireless, is 500 bytes/s in total. That proves that no periodic exchange of information happens except in the case of mobile ESIMs where they need to update their position information. The wireless slot allocation overhead imposed on the control link of the satellites is no more than 400 bytes/s in the worst case. The related CP advertisement messages and the slot allocation messages are all designed to be event-driven and therefore scale nicely with the network density.

5.6.6.3 DSS2

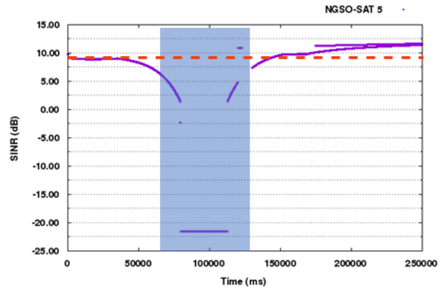
Due to similarity with SINR graphs presented for DSS1 and to avoid duplication, we did not present SINR graphs for DSS2. Similar to DSS1, DSS2 manages to resolve all interference events on GSO and NGSO links, with SINR measurements keeping above the 9 dB threshold for the most part. All links are stable and are reacting when SINR drops below 9 dB within a few seconds. FWA nodes 8 and 9 do receive some interference from the downlink signals of the satellite links, but they keep within their safe margin, never dropping below 15 dB SINR. The FWA links had a 0% packet loss, like DSS1 and the baseline, as downlink communications do not seem to have a major impact on the earth FWA links. If this impact is not negligible, DSS2 would not be able to protect FWA links, unless FWA nodes are supporting spectrum sensing and operating under DSS2 framework, which would make them adaptive and protected against external interference. It is found out that DSS2 provides marginally lower packet loss and marginally better throughput in the dynamic part of our scenario compared to DSS1. A closer look at our detailed logs reveals that DSS2 lost 1900 packets in total, while DSS1 lost 2000 packets, providing a small advantage, for a total loss in both of about 0.005% packets.

5.6.7 Uplink use case

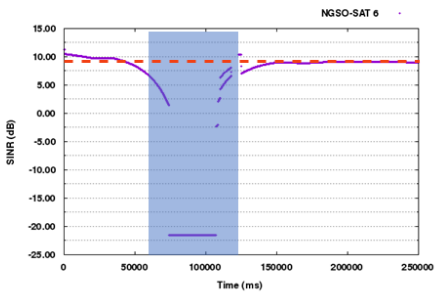
In this section, we discuss the uplink use case and examine the impact of NGSO and GSO uplink transmission on the FWA links and the GSO and NGSO satellites. We also examine the ability of the proposed DSS techniques to provide interference protection to both the FWA and the satellite links during uplink transmission. For the uplink case, the FWA link direction between FWA 9 and 10 was reversed so that the FWA link now radiates towards the sky and not towards the earth. This allows us to analyze if the FWA links can be impacted by the GSO and NGSO ESIMs uplink transmission towards space.



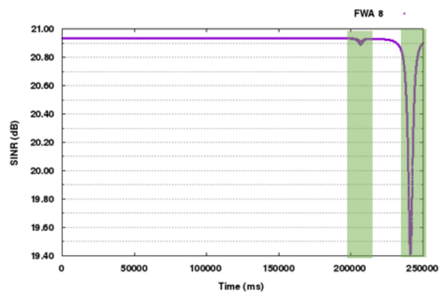
(a)



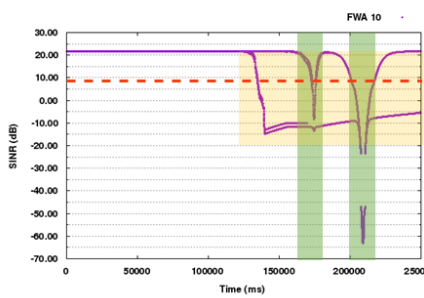
(b)



(c)



(d)



(e)

Figure 5.11: SINR graphs in the uplink use case (baseline). (a) GSO-SAT 4. (b) NGSO-SAT 5. (c) NGSO-SAT 6. (d) FWA 8. (e) FWA 10.

In this simulation set, we allowed the system to allocate the spectrum into 5 available channels. An initial set of simulations was performed with 3 available channels, but it was noticed that DSS techniques cannot resolve interference because the two NGSO links are suffering from adjacent channel emissions from the GSO ESIM. It was observed that, although the NGSO links moved out of the channel/slots used by the GSO ESIM to avoid co-channel interference, NGSO links still suffer a lot of interference in adjacent slots of GSO ESIM transmissions during alignment with the GSO. Such ACI can only be mitigated by moving away further from the GSO ESIM slots. For this reason, 5 channels are the minimum number of channels that allow NGSO links to avoid both ACI and co-channel interference from the GSO link.

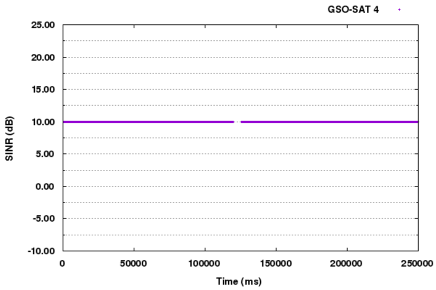
5.6.7.1 Baseline

The SINR results for the baseline scenario simulation run are presented in Fig. 5.11. As expected, the satellite connections are not stable in the baseline as all links reuse the same channel for operation. During the convergence phase (see blue regions), both NGSO links suffer from interference but then recover to their normal operation. The GSO link is only slightly impacted by the interference from the NGSO links: we can observe that the SINR of the GSO receiver maximally decreases by 0.5 dB, which is within the 1 dB safety margin. The FWA links are impacted as the ESIMs are moving closer to them, aligning with the FWA beam. The impact on FWA 8 is much smaller though, not more than 1.5 dB, whereas FWA 10 is seriously impacted by ESIMs passing close to its position. The interference impact of NGSO-ESIM 3 stopping behind the FWA 8 node is visible from 120 to 250 seconds of simulation time (yellow regions), whereas the passing of the other two ESIMs behind the FWA receiver is also quite visible in the SINR graph as deep spikes of SINR drops (green areas). Once interference occurs in the second part of the simulation run, the FWA links are heavily impacted and unable to operate.

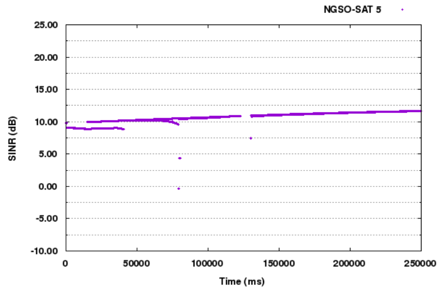
In the next sections, we analyze if the proposed DSS techniques are capable to resolve the satellite link interference events and also to protect the FWA links.

5.6.7.2 DSS1

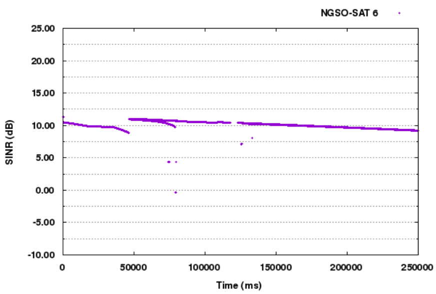
DSS1 resolves a good portion of the interference events, especially in the satellite links. Satellite nodes (GSO-SAT 4, NGSO-SAT 5, and NGSO-SAT 6) are all healthy now, keeping their SINR values over 9 dB in almost all cases throughout the simulation run (see Fig. 5.12). FWA links continue to experience high interference though, making it obvious that we were unable to protect them. Since the FWA nodes cannot select dynamically their channel usage, they cannot recover when interfered. The main problem comes from the GSO link that is very stable and does not suffer from any interference and hence does not take any action for slot reallocation. It is therefore impossible to protect terrestrial FWA links which seriously suffer from high packet losses. To remedy this situation, we have envisioned an upgrade on the CP for DSS1 allowing FWA nodes to inform other systems participating in the collaboration about their interference status. With this upgrade, FWA links can alert other DSS capable nodes and trigger slot reallocation.



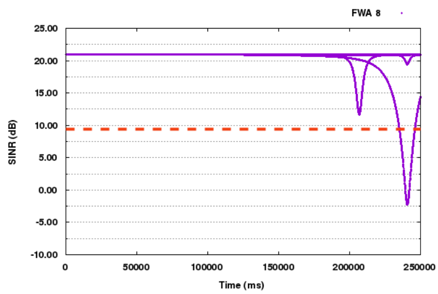
(a)



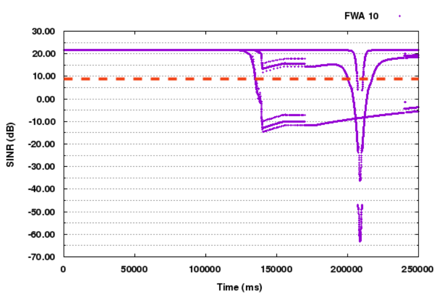
(b)



(c)



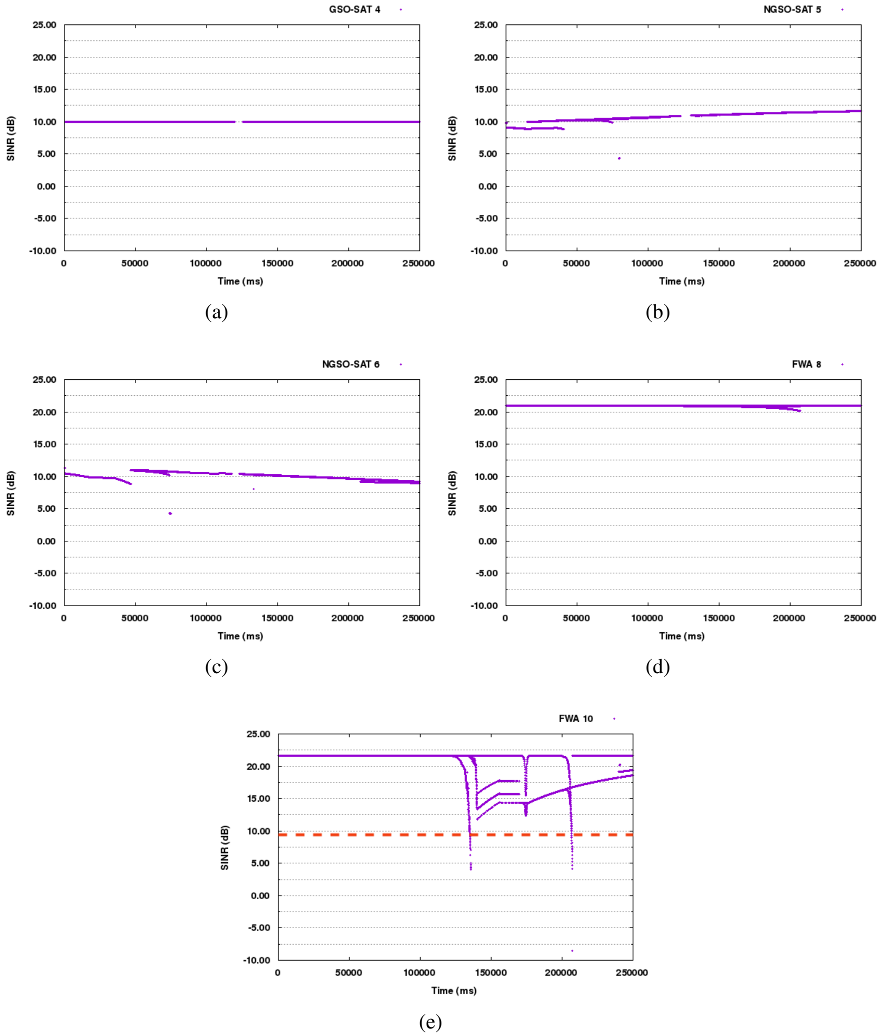
(d)



(e)

Figure 5.12: SINR graphs in the uplink use case (DSS1). (a) GSO-SAT 4. (b) NGSO-SAT 5. (c) NGSO-SAT 6. (d) FWA 8. (e) FWA 10.

Overall, DSS1 solves most of the interference problems of the satellite links, whereas terrestrial links still suffer from interference. With DSS1 applied, the FWA links are more resistant to interference than in the baseline case. To better protect terrestrial links, the CP needs to be extended to alert the satellite links about the status of the FWA links (DSS1+).



(a) (b) (c) (d) (e)

Figure 5.13: SINR graphs in the uplink use case (DSS1+). (a) GSO-SAT 4. (b) NGSO-SAT 5. (c) NGSO-SAT 6. (d) FWA 8. (e) FWA 10.

5.6.7.3 DSS1+

With the introduction of FWA alarm messages in CP of DSS1 technique, we can see from the SINR graphs in Fig. 5.13 that FWA links are much better protected now. Moreover, there is no side effect of the extra FWA protection mechanism on satellite links. The protection mechanism is offering a significant improvement, returning the FWA links almost back to optimal operation, with FWA links operating close to 99% uptime now.

The introduction of additional CP messages is also not causing a significant increase in control channel overhead. FWA alert messages are used sparsely, only in case interference is detected, in line with the event-based design approach of the CP. Hence as expected, the induced overhead is minimal (see Table 5.3).

5.6.7.4 DSS2

As SINR measurement graphs are similar to DSS1, they are not presented here. Detailed analysis shows that with DSS2 the offered service was marginally better (see Table 5.3). Satellite links behave almost optimally, however, just like DSS1, DSS2 fails to protect FWA links. When further examining the reason for this, we concluded that since DSS2 is based on spectrum sensing at the receiver, and the receiver of the GSO link is 36,000 km away, it cannot sense the presence of FWA links at the surface of the Earth. Furthermore, as the GSO link has zero PER, it does not need to initiate any slot reallocation procedure. On the other hand, the powerful GSO-ESIM transmitter creates a huge amount of interference on the Earth's surface, impacting all links that are not capable to adapt their spectrum allocation.

To summarize, DSS2 results are quite similar to DSS1, solving the problems of interference for the satellite links but not resolving the FWA issues, as FWA nodes do not support spectrum sensing and adaptive spectrum allocation. The only way to resolve the remaining issues using DSS2 is to also integrate the DSS module that supports spectrum sensing and dynamic spectrum access in FWA nodes, making them also adaptive and interference-aware. In this way, they would be able to simply move out of the spectrum that is affected by the GSO link.

5.6.8 Summary of results and main findings

A summary of the measured KPIs (as defined at the start of Section 5.6) is presented in Tables 5.1, 5.2, and 5.3 for baseline and DSS techniques. For completeness, we also presented results for the 3 channels uplink use case. Red numbers in the tables indicate unsatisfactory results. For the satellite links, the main focus was to benefit from opportunistic spectrum sharing with primary terrestrial links while not introducing harmful interference to the primary links, maximize uptime, and increase spectrum utilization by using collaborative and dynamic spectrum sharing techniques. As can be seen from the averaged results, avoiding interference may come at the cost of a short downtime (or reduced uptime) and a small packet loss during downtime.

The key results and findings for the two use cases are summarized below:

- Opportunistic sharing of satellite and terrestrial systems in downlink Ka-band:

Table 5.1: KPIs for the downlink use case.

	KPI	DOWNLINK - 3 channels		
		BASELINE	DSS1	DSS2
Satellite links	Throughput [KB]	4313.06	6244.79	6237.26
	PER [%]	30.40%	1.13%	1.25%
	Uptime (%)	70.40%	99.18%	97.75%
	Spectrum utilization [%]	88.10%	96.72%	96.84%
	Latency [msec]	81.06	76.17	75.7
	CP overhead [B/sec]	0	24.04	0
	Slot selection overhead [B/sec]	0	6.65	7.71
FWA links	Uptime	100.00%	100.00%	100.00%
All links	Uptime	82.24%	99.51%	98.65%

Table 5.2: KPIs for the uplink use case, 3 channels.

	KPI	UPLINK - 3 channels		
		BASELINE	DSS1	DSS2
Satellite links	Throughput [KB]	4566	5734	5737
	PER [%]	27.54%	8.76%	8.70%
	Uptime (%)	69.28%	90.08%	90.05%
	Spectrum utilization [%]	88.12%	96.72%	96.91%
	Latency [msec]	101.75	87.49	86.39
	CP overhead [B/sec]	0	2627.02	0
	Slot selection overhead [B/sec]	0	128.1	125.99
FWA links	Uptime	84.64%	97.26%	95.71%
All links	Uptime	75.42%	92.91%	92.01%

– All DSS techniques offer significant advantages over the baseline in terms of link stability and uptime. All interference events can be detected by monitoring link statistics at the receiver. The overall uptime is increased from around 82.2% for the baseline scenario to 98.6-99.5% when DSS is applied.

– All DSS techniques are capable to quickly recover from interference generated by known satellite and terrestrial systems. When an interference event occurs, only a few packets are lost, and any event is resolved quickly by restoring a stable connection within a few seconds in the worst case.

Table 5.3: KPIs for the uplink use case, 5 channels.

	KPI	UPLINK - 5 channels			
		BASELINE	DSS1	DSS1+	DSS2
Satellite links	Throughput [KB]	4566	6273	6275	6284
	PER [%]	27.54%	0.80%	0.75%	0.61%
	Uptime (%)	69.28%	97.79%	98.01%	97.93%
	Spectrum utilization [%]	88.11%	96.79%	96.78%	96.78%
	Latency [msec]	101.75	78.09	78.31	77.33
	CP overhead [B/sec]	0	464.74	462.72	0
	Slot selection overhead [B/sec]	0	7.92	8.35	4.46
FWA links	Throughput [KB]	4435.83	4671.66	5231.79	5086.71
	PER [%]	15.34%	10.84%	0.18%	10.87%
	Uptime	84.67%	96.54%	99.92%	96.77%
All links	Uptime	75.43%	97.29%	98.01%	97.46%

  - Terrestrial MW FWA links do not suffer from interference by satellite links.

- Opportunistic sharing of satellite and terrestrial systems in uplink Ka-band:
  - All DSS techniques offer significant advantages over the baseline in terms of link stability and uptime. The overall uptime is increased from around 75.4% for the baseline to 98% when DSS is applied. 98% uptime can only be achieved when:
    * ACI is taken into account, in particular when the high transmission power is involved like an uplink transmission by the GSO-ESIM. Also, side-lobe emissions can create interference events and should not be ignored.
    * Terrestrial MW FWA receivers send alert messages via the CP to warn satellite systems causing interference and to request them to move out of time-frequency slots also used by terrestrial FWA links.

Some additional key findings, independent of the use cases, are:

- Besides the capability of mitigating interference events, an additional advantage of using an MF-TDMA MAC scheme combined with proposed DSS techniques is the increased spectrum efficiency as the DSS techniques can instantaneously adjust the allocated spectrum to match the incoming application layer load, ensuring that no more than needed spectrum is used at any given time. It can be seen that in both downlink and uplink use cases, applying DSS techniques leads to an increase of spectrum utilization from 88% to 96%. We anticipate that the proposed DSS techniques would lead to much higher spectrum utilization gains in more complex and larger-scale satellite-terrestrial deployments.

- The wireless control overhead of all DSS techniques is negligible. The control overhead is only a few kB of data being exchanged per interference event between the related communicating entities, and remains constant per interference event, not scaling up with a number of deployed satellites but with the limited number of satellites/ESIMs affecting or getting affected per interference event.

- In terms of implementation, DSS1 is the simplest technique, but it can only mitigate known interference sources (advertised by CP messages), whereas DSS2 with its spectrum monitoring capabilities can mitigate unknown as well as known interference.

- There is only one drawback with using event-driven DSS techniques: although the downtime is very limited, short periods with high packet loss of a few seconds are possible. For this simulation study, conservative link statistics measurements were used for detecting interference events. Interference was only detected after packets have been lost. This could be mitigated by applying techniques for the prediction of interference events (e.g., more advanced SINR analytics for early detection of SINR degradation or pattern recognition of recurring interference events using artificial intelligence and machine learning techniques) and proactive resource reallocation before packet loss occurs.

5.7 Conclusion

The current static and regulated spectrum allocation in SATCOM leads to underutilization of the spectrum. This regulated spectrum allocation cannot meet the emergence of new satellite services and satellite operators that are recently entering into SATCOM. Therefore, to support a dense deployment of new satellite services and accommodate the new operators, it is required to reuse and share the spectrum. The reuse and sharing of the spectrum can be enabled by employing DSS techniques. Two decentralized DSS techniques have been analyzed in terms of spectrum reuse for opportunistic sharing of satellite and terrestrial systems in downlink Ka-band and uplink Ka-band. The first technique relies on a decentralized CP that connects all entities across different systems and is used for advertising relevant information, while the second technique is based on decentralized spectrum sensing and requires spectrum monitoring capabilities.

Achieved results demonstrate that the proposed DSS techniques show remarkable interference mitigation capabilities as compared to the baseline, reducing the downtime of the established communication links by 60-99%. Interference mitigation between GSO and NGSO links is proven. Whereas FWA links are not affected in the downlink use case, protection of FWA links in the uplink use case was achieved with small upgrades on the DSS techniques. It has become evident that with a few more enhancements a seamless experience can be provided to the application layer users in all cases. The control overhead is negligible, measured in a few kB of data being exchanged per interference event between the related peers, and remains constant per interference event, not scaling up with the number of deployed satellites but rather with satellites/ESIMs affecting or getting affected per interference event. An additional advantage is the increased spectrum efficiency as the DSS techniques enable instantaneous adjustment of the allocated spectrum to match incoming application layer load, making sure no more than the needed spectrum is used at any given time. This effect in conjunction with the use of a smaller basic transmission unit in the form of a slot instead of a channel enables increased spectrum efficiency, optimally mapping incoming traffic to allocated spectrum. In both downlink and uplink scenarios, spectrum utilization was increased from 88% in baseline to 96% with applied DSS techniques. We anticipate seeing higher gains in spectrum utilization in the case of higher density satellite-terrestrial network deployments, where spatial reuse of spectrum plays a bigger role.

The work presented in this chapter is showing preliminary results of DSS techniques for enabling the coexistence of satellite and terrestrial networks. The evaluated scenarios are local and include few user terminals. In reality, the coverage area of satellite systems includes a much higher number of terrestrial deployments. Therefore, in future work, large-scale testing is needed to investigate the scalability of proposed solutions and to pinpoint the areas where the design needs to be fine-tuned. Besides, with large-scale testing, we will be able to determine the impact of proposed techniques in scenarios with large-scale satellite-terrestrial deployments. Another interesting point to discuss in the future is improved protection of terrestrial links. Improved protection of terrestrial networks may be achieved, for instance, by improving CP alert message for anticipating packet loss or by extending the terrestrial networks with the same DSS techniques as applied in satellite systems. Besides, the anticipation of packet loss by applying advanced techniques for the prediction of interference events could enable proactive resource reallocation before packet loss occurs and further reduce downtime in satellite links.

Acknowledgment

This work was conducted and funded through the ESA ARTES Future Preparation Programme.

References

[1] C. Henry. *SpaceX submits paperwork for 30,000 more Starlink satellites*. Available from: https://spacenews.com/spacex-submits-paperwork-for-30000-more-starlink-satellites/.

[2] S. K. Sharma, S. Chatzinotas, and B. Ottersten. *Satellite cognitive communications: Interference modeling and techniques selection*. In 2012 6th Advanced Satellite Multimedia Systems Conference (ASMS) and 12th Signal Processing for Space Communications Workshop (SPSC), pages 111–118, 2012. doi:10.1109/ASMS-SPSC.2012.6333061.

[3] P. Kolodzy. *Spectrum policy task force, Federal Communications Commission: Washington, DC, 2002. Rep. ET Docket, (02-135).*

[4] S. Chatzinotas, S. K. Sharma, and B. Ottersten. *Frequency Packing for Interference Alignment-Based Cognitive Dual Satellite Systems*. In 2013 IEEE 78th Vehicular Technology Conference (VTC Fall), pages 1–7, 2013. doi:10.1109/VTCFall.2013.6692209.

[5] P. Chini, G. Giambene, and S. Kota. *A survey on mobile satellite systems*. Int. J. Satell. Commun. Netw., 28:29–57, 2010.

[6] N. Cassiau, G. Noh, S. Jaeckel, L. Raschkowski, J. Houssin, L. Combelles, M. Thary, J. Kim, J. Dore, and M. Laugeois. *Satellite and Terrestrial Multi-Connectivity for 5G: Making Spectrum Sharing Possible*. In 2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), pages 1–6, 2020. doi:10.1109/WCNCW48565.2020.9124728.

[7] D. Tarchi, A. Guidotti, V. Icolari, A. Vanelli-Coralli, S. K. Sharma, S. Chatzinotas, S. Malekil, B. Evans, P. Thompson, W. Tang, and J. Grotz. *Technical challenges for cognitive radio application in satellite communications*. In 2014 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), pages 136–142, 2014. doi:10.4108/icst.crowncom.2014.255727.

[8] *ERC Decision of 19 October 2000 on the Shared Use of the Band 17.7-19.7 GHz by the Fixed Service and Earth Stations of the Fixed-Satellite Service (Space-to-Earth)*, 2000.

[9] *The Use of the Band 27.5-29.5 GHz by the Fixed Service and Uncoordinated Earth Stations of the Fixed-Satellite Service (Earth-to-Space)*, 2005.

[10] Imec, Antwerp Space N.V., Trinity College Dublin. *CODYSUN - Collaborative and Dynamic Approaches to Increasing Spectrum Utilisation*, 2021. Available from: https://artes.esa.int/projects/codysun.

[11] Defense Advanced Research Project Agency (DARPA). *Spectrum Collaboration Challenge (SC2)*. Available from: https://archive.darpa.mil/sc2/.

[12] M. Höyhtyä, J. Ylitalo, X. Chen, and A. Mämmelä. *Chapter 14 - Use of databases for dynamic spectrum management in cognitive satellite systems*. In S. Chatzinotas, B. Ottersten, and R. De Gaudenzi, editors, Cooperative and Cognitive Satellite Systems, pages 453–480. Academic Press, 2015. doi:https://doi.org/10.1016/B978-0-12-799948-7.00014-1.

[13] Z. Chen, D. Guo, K. An, B. Zhang, X. Zhang, and B. Zhao. *Efficient and Fair Resource Allocation Scheme for Cognitive Satellite-Terrestrial Networks*. IEEE Access, 7:145124–145133, 2019. doi:10.1109/ACCESS.2019.2943908.

[14] M. Jia, X. Gu, Q. Guo, W. Xiang, and N. Zhang. *Broadband Hybrid Satellite-Terrestrial Communication Systems Based on Cognitive Radio toward 5G*. IEEE Wireless Commun., 23(6):96–106, 2016. doi:10.1109/MWC.2016.1500108WC.

[15] E. Lagunas, S. K. Sharma, S. Maleki, S. Chatzinotas, and B. Ottersten. *Power Control for Satellite Uplink and Terrestrial Fixed-Service Co-Existence in Ka-Band*. In 2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall), pages 1–5, 2015. doi:10.1109/VTCFall.2015.7391092.

[16] S. K. Sharma, S. Chatzinotas, and B. Ottersten. *In-line interference mitigation techniques for spectral coexistence of GEO and NGEO satellites*. Int. J. Satell. Commun. Netw., 34(1):11–39, 2016. doi:https://doi.org/10.1002/sat.1090.

[17] A. Pourmoghadas, S. K. Sharma, S. Chatzinotas, and B. Ottersten. *Cognitive Interference Management Techniques for the Spectral Co-existence of GSO and NGSO Satellites*. In I. Otung, P. Pillai, G. Eleftherakis, and G. Giambene, editors, Wireless and Satellite Systems, pages 178–190. Springer International Publishing, 2017.

[18] C. Wang, D. Bian, S. Shi, J. Xu, and G. Zhang. *A Novel Cognitive Satellite Network With GEO and LEO Broadband Systems in the Downlink Case*. IEEE Access, 6:25987–26000, 2018. doi:10.1109/ACCESS.2018.2831218.

[19] C. Zhang, C. Jiang, L. Kuang, J. Jin, Y. He, and Z. Han. *Spatial Spectrum Sharing for Satellite and Terrestrial Communication Networks*. IEEE Trans. Aerosp. Electron. Syst., 55(3):1075–1089, 2019. doi:10.1109/TAES.2018.2889585.

[20] E. Lagunas, S. K. Sharma, S. Maleki, S. Chatzinotas, J. Grotz, J. Krause, and B. Ottersten. *Resource Allocation for Cognitive Satellite Uplink and Fixed-Service Terrestrial Coexistence in Ka-Band*. In M. Weichold, M. Hamdi, M. Z. Shakir, M. Abdallah, G. K. Karagiannidis, and M. Ismail, editors, Cognitive Radio Oriented Wireless Networks, pages 487–498. Springer International Publishing, 2015.

[21] E. Lagunas, S. K. Sharma, S. Maleki, S. Chatzinotas, and B. Ottersten. *Resource Allocation for Cognitive Satellite Communications With Incumbent Terrestrial Networks*. IEEE Trans. Cogn. Commun. Netw., 1(3):305–317, 2015. doi:10.1109/TCCN.2015.2503286.

[22] A. Vanelli-Coralli, A. Guidotti, D. Tarchi, S. Chatzinotas, S. Maleki, S. Krishna Sharma, N. Chuberre, B. Evans, M. Lopez-Benitez, W. Tang, J. Grotz, and K. Liolis. *Chapter 10 - Cognitive radio scenarios for satellite communications: the CoRaSat project*. In S. Chatzinotas, B. Ottersten, and R. De Gaudenzi, editors, Cooperative and Cognitive Satellite Systems, pages 303–336. Academic Press, 2015. doi:https://doi.org/10.1016/B978-0-12-799948-7.00010-4.

[23] S. Maleki, S. Chatzinotas, J. Krause, K. Liolis, and B. Ottersten. *Cognitive Zone for Broadband Satellite Communications in 17.3–17.7 GHz Band*. IEEE Wireless Commun. Lett., 4(3):305–308, 2015. doi:10.1109/LWC.2015.2411597.

[24] VTT, Univ. Oulu, Bittium, Airbus Defence & Space. *ESA ARTES 1 Project FREESTONE: Frequency Sharing Techniques With Other Networks or Radio Services*, 2012. Available from: https://artes.esa.int/projects/freestone-frequency-sharingtechniques-other-networks-or-radio-services.

[25] M. Höyhtyä, A. Mämmelä, X. Chen, A. Hulkkonen, J. Janhunen, J. Dunat, and J. Gardey. *Database-Assisted Spectrum Sharing in Satellite Communications: A Survey*. IEEE Access, 5:25322–25341, 2017. doi:10.1109/ACCESS.2017.2771300.

[26] M. Höyhtyä, J. Kyröläinen, A. Hulkkonen, J. Ylitalo, and A. Roivainen. *Application of cognitive radio techniques to satellite communication*. In 2012 IEEE International Symposium on Dynamic Spectrum Access Networks, pages 540–551, 2012. doi:10.1109/DYSPAN.2012.6478178.

[27] CBRS Alliance. *Citizens Broadband Radio Service (CBRS)*. Available from: https://www.cbrsalliance.org/.

[28] ETSI. *Digital Video Broadcasting (DVB); Second Generation DVB Interactive Satellite System (DVB-RCS2); Part 1: Overview and System Level*

specification, Jul. 2020. Available from: https://www.etsi.org/deliver/etsi_ts/101500_101599/10154501/01.03.01_60/ts_10154501v010301p.pdf.

[29] J. Sanz Subirana and J.M. Juan Zornoza and M. Hernández-Pajares. *Reference Systems and Frames*, 2011. Available from: https://gssc.esa.int/navipedia/index.php/Reference_Systems_and_Frames.

[30] Magister Solutions Ltd. *Satellite Network Simulator 3 (SNS3)*. Available from: https://www.sns3.org/content/home.php.

[31] NS-3 Consortium. *ns-3 Network Simulator*. Available from: https://www.nsnam.org/.

[32] ITU. *Satellite antenna radiation pattern for use as a design objective in the fixed-satellite service employing geostationary satellites*, 1997. Available from: https://www.itu.int/dms_pubrec/itu-r/rec/s/R-REC-S.672-4-199709-I!!PDF-E.pdf.

[33] ITU. *Satellite antenna radiation patterns for non-geostationary orbit satellite antennas operating in the fixed-satellite service below 30 GHz*, 2001. Available from: https://www.itu.int/dms_pubrec/itu-r/rec/s/R-REC-S.1528-0-200106-I!!PDF-E.pdf.

[34] ITU. *Radiation diagrams for use as design objectives for antennas of earth stations operating with geostationary satellites*, 2003. Available from: https://www.itu.int/dms_pubrec/itu-r/rec/s/R-REC-S.580-6-200401-I!!PDF-E.pdf.

6
Conclusion and Future Work

“Let us put our minds together and see what life we can make for our children.”

–Sitting Bull (1831 - 1890)

The electromagnetic spectrum is a precious and limited resource, continuously gaining in value as demand is constantly rising. However, we are misusing it; license-free frequency bands are becoming overcrowded while licensed frequency bands are often underutilized. The world is turning into one big interconnected global communication network, incorporating not only humans but also animals, machines, infrastructure, etc. The arising concern is if inefficient utilization of the spectrum is going to be the bottleneck that slows down or completely halts this on-going expansion? In this thesis, we argue that end-user experience, technological advancements, development of new wireless technologies and their introduction to the market, etc., should not be restricted in any way by the model of spectrum management we are using today. We support the claim that the current model of spectrum management is not a maintainable concept in the long-term future and migration to dynamic sharing of the spectrum is a necessary step. Further, we consider that all spectrum bands (licensed, unlicensed) should migrate to a dynamic spectrum sharing paradigm that unifies the way spectrum is accessed. Therefore, this thesis introduces novel dynamic spectrum sharing approaches for dealing with various challenges that arise in the field of spectrum management. Among others, this includes solutions for underutilization of the radio spectrum, the coexistence of heterogeneous wireless networks, and scheduling in networks with the Time Division Multiple Access (TDMA) Medium Access Control (MAC) layer.

The remainder of this chapter summarizes the main contributions of this PhD research and identifies the potential directions for future work.

6.1 Conclusion

In Chapter 2 we have proposed a dynamic licensing scheme that can be a solution for solving the problem of underutilization in licensed frequency bands. Further, for the problem of coexistence between heterogeneous technologies operating in overcrowded license-free bands, we have proposed another model for sharing the spectrum. Both proposals are based on fragmenting the spectrum into time-frequency slices and assigning a minimum number of slices to each network, based on their demands. For the execution of proposed models, it is assumed that the entry point (EP) (access point (AP), base station (BS), or gateway (GW)) of each network can communicate with Central entities (CEs) of Regulatory Committees (licensed bands) or with EPs of other networks (unlicensed bands). Once the networks acquire the requested resources, each network can independently decide how the devices in the network are going to share and access the acquired spectrum. With the proposed dynamic licensing scheme, wireless operators that want to operate in licensed bands are only required to pay for the number of slices they have acquired. This would result in more innovation freedom in the field of telecommunications, more new technologies and services entering the market, all leading to reduced consumer prices and better user services. Complete migration from the current model of using the spectrum to fully dynamic sharing might take years, even decades. During this period, there will be many technological advancements that can make this shift easier. Proposed models will benefit from any technological advancement in the area of wireless communications and can be adapted accordingly, however, their basic concepts will stay the same.

Chapter 3 presents how the MAC layer with built-in dynamic spectrum sharing capability fits into an end-to-end wireless system where Cognitive Radio (CR) and Machine Learning (ML) techniques are used to make the decisions on allocation and removal of the slots. In this chapter, also a brief description of other system layers is given, and it is described how they interact with each other to achieve coexistence with unknown wireless networks and provide collaborative spectrum usage. Chapter 4 demonstrates how distributed and dynamic spectrum sharing can be achieved on the node level. The proposed Dynamic Distributed Multi-Channel TDMA (DDMC-TDMA) scheduling protocol enables every node in the network to acquire the required number of slots in the Multi-Frequency Time Division Multiple Access (MF-TDMA) scheme. Acquired slots are guaranteed to be free from internal and/or external interference. As such, DDMC-TDMA is a promising solution for dealing with problems of underutilization and coexistence both in licensed and license-free bands. DDMC-TDMA is also proven to be scalable in single-hop and multi-hop networks. Moreover, it is shown that networks with DDMC-TDMA are able to avoid hidden and exposed node problems, with the avoidance of the exposed node problem providing near-optimal spatial reuse of the spectrum.

Chapter 5 covers main research challenges in satellite communications, namely underutilization due to fixed spectrum assignment and coexistence of satellite and terrestrial systems operating in the same frequency bands. The downsides of fixed spectrum assignment are pinpointed and two dynamic spectrum sharing techniques are proposed as a suitable replacement. The performance of proposed techniques is evaluated in most complex scenarios where geostationary (GSO), non-geostationary (NGSO), and terrestrial systems all operate in the same frequency band and overlapping geographical areas. Dynamic spectrum sharing techniques are only applied to satellite systems, which are required to protect incumbent terrestrial systems. The comparison against the performance of the fixed spectrum model, applied to the same scenarios, shows remarkable improvements in interference mitigation and increased spectrum utilization. The achieved results support the claim that dynamic spectrum sharing is the way forward for spectrum management in satellite communications and that fixed spectrum allocation will become unsustainable with further increases in the size and density of satellite networks.

6.2 Future work

In this dissertation, some major steps were taken to address the challenges that are present in wireless communications today, more specifically, challenges in the areas of radio spectrum management and dynamic sharing of the spectrum. This section discusses future work that could complement the contributions of this dissertation.

For the spectrum sharing models that are proposed in Chapter 2 to be applied in real-world, there is a need for a set of regulatory rules and standards, which will limit the misuse of the sharing frameworks and protect their main concepts; minimizing the spectrum footprint of each network, while enabling them to deliver their desired Quality of Service (QoS) at any moment and in any location. This further requires the implementation of monitoring of the spectrum usage and penalization of the operators that do not adhere to the regulations. Operators need to be prevented from overdimensioning their spectrum usage, which would bring back the same problems of inefficient spectrum utilization we are seeing today. The proper selection of spectrum slices that are going to be assigned to different operators is another thing that needs to be addressed. Inadequate assignment of the slices may lead to the appearance of unused spectrum left between allocated spectrum slices (fragmentation of spectrum). Such spectrum fragments may be too small for utilization by any technology and remain unused for a long time. Thus, the design of a proper assignment algorithm is needed. This algorithm should provide an optimal assignment policy and make sure to avoid the occurrence of unusable spectrum fragments. Both spectrum sharing models presented in Chapter 2 should be evaluated either through simulations or experimentally, demonstrating that they are able to offer interference-free operation while solving the problem of spectrum underutilization.

DDMC-TDMA scheduling protocol has been tested in emulation and simulation environments and showed promising results for dynamic spectrum sharing and creating collision-free TDMA schedules. Conducting experiments in real-world wireless networks would be the next step to determine if the achieved results are comparable or some adjustments to the protocol are needed. Although DDMC-TDMA proved to be a generic solution that works in various use cases, testing its performance in additional scenarios would give further insights into its applicability. For example, one interesting scenario is a scenario where wireless networks consist of mobile nodes. It may be beneficial to verify how successfully DDMC-TDMA adapts to the mobility of the nodes, more specifically its ability to utilize spectrum that becomes available due to topology changes and ability to avoid internal interference induced by movement of the nodes to the same collision domain. Another use case that may be considered is the use case where multiple DDMC-TDMA based networks operate in the same collision domain. In this use case, it is worth verifying how each network can adapt its schedule to coexist with other networks and how fast they converge to a steady state. DDMC-TDMA has been implemented to be a Physical layer (PHY) agnostic protocol and therefore easily integrated with any existing or future wireless PHY standards. However, it has only been tested on top of an LTE-based PHY, and its performance could be verified with other standardized PHYs.

To complement our research on dynamic spectrum sharing in satellite communications, additional research is needed to validate and compare the potential of proposed dynamic spectrum sharing techniques. More specifically, we need to experimentally evaluate in small-scale tests how effective the proposed techniques are with dynamic scenarios in real satellite networks, in particular for the detection of interference events between satellite and terrestrial links. In a real small-scale network, the behavior of the participating satellite and terrestrial nodes can be examined in detail and proposed techniques adapted if necessary. Besides experimental testing, simulated models should be extended to support simulations of large-scale networks, which would give us further insights into the scalability of the proposed techniques and may lead to better fine-tuning of their design. It would also be interesting to see how the dual satellite and terrestrial systems coexist if terrestrial systems are not considered incumbent, but they operate following the same dynamic spectrum sharing methods as satellite systems. This would actually lead to the unified dynamically shared spectrum paradigm across all wireless technologies and application domains, which we envisioned during this PhD research.

We are all part of the big global network, let us take the next step to improve it.