Grant Agreement No.: 671705
Research and Innovation action
Call Identifier: H2020-ICT-2014-2

5G PPP

SpeedFG

quality of Service Provision and capacity Expansion through Extended-DSA for 5G

D4.4: Final definition of the RM solution options and recommendations (public version)

Version: 1.1

Deliverable type	Report
Dissemination level	PU
Due date	31/10/2017
Submission date	15/11/2017
Lead editor	Keith Briggs (BT)
Authors	Keith Briggs, Michael Fitch (BT); Ioannis-Prodromos Belikaidis, Voula Vassaki, Andreas Georgakopoulos, Evangelos Kosmatos, Evangelia Tzifa, Kostas Tsagkaris, Aikaterini Demesticha, Vera Stavroulaki, Panagiotis Vlacheas, Dimitrios Kelaidonis, Yiouli Kritikou, Athena Ropodi, Panagiotis Demestichas, Marinos Galiatsatos, Aristotelis Margaris, Dimitrios Kardaris, Evangelos Argoudelis (WINGS); Antonio de Domenico (CEA); Haeyoung Lee, Dionysia Triantafyllopoulou, Faouzi Bouali (UNIS); Shahid Mumtaz, Jonathan Rodriguez, Kazi Saidul Huq, Imran Khan (IT); Dimitrios Kritharidis, Konstantinos Chartsias, Thanasis Oikonomou (ICOM).
Reviewers	Valerio Frascolla (Intel)
Work package, Task	WP4, T4.3
Keywords	5G, Radio Resource Manager, RRM, performance evaluations, regulatory and operator constraints

Abstract

The RRM (Radio Resource Manager) entity is a key component of the SPEED-5G architecture. It consists of two parts, a centralised part (cRRM) and a distributed part (dRRM). This deliverable is devoted to performance evaluations of the various algorithms which will run in the cRRM, but also includes an outline of regulatory and operator constraints, and of software architecture for the cRRM.

Document revision history

Version	Date	Description of change	List of contributor(s)
v0.1	2017-07-25	Draft table of contents	Haeyoung Lee, Keith Briggs
v0.4	2017-08-07	Added content to section 2	Michael Fitch
v0.8	2017-08-22	Added IT chapter, cleaned up formatting	Keith Briggs
v0.12	2017-09-11	Merged versions 0.10 and 0.11	Keith Briggs
v0.13	2017-09-11		IT
v0.14	2017-09-11	Chapter 9 updates	WINGS
v0.15	2017-09-13	Chapter 9 corrections, and other updates	Keith Briggs
v0.16	2017-09-14	Many corrections	Keith Briggs
v0.20	2017-09-15	More corrections	Keith Briggs
v0.27	2017-09-18	First version for review	Keith Briggs
v0.28	2017-09-19	ICOM chapter added	Keith Briggs
v0.29	2017-09-20	BT chapter updated	Keith Briggs
v0.30	2017-09-21	Response to Valerio's review	Keith Briggs
v0.31	2017-09-22	RRM chapter updated	Keith Briggs
v0.32	2017-09-25	Response to Seiamak's review	Keith Briggs
v1.0	2017-10-30	ICOM updates and final revisions	Keith Briggs
v1.1	2017-11-15	Re-enter omitted inputs	Haeyoung Lee, Faouzi Bouali

Disclaimer

Copyright notice

[1] http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en_US

Executive Summary

This deliverable gives the final definition of the RRM framework including the centralized RMM (cRRM) with its interfaces to other parts of the core network: the spectrum manager, Key Performance Indicators (KPI) collector, and Operational Support Services (OSS) and to dRRMs, and the dRRM with its interfaces to the cRRM and to the MAC layer. The proposed RRM design is capable of incorporating algorithms from multiple vendors, whether those algorithms are centralized or distributed. Moreover, the proposed RRM framework fully decouples the underlying algorithms by the introduction of an abstraction layer (AL) and supports multiple interfaces transparently from the algorithmic point of view. This means that communications with the entities outside the RRM are the responsibility of the AL.

This deliverable outlines the proposed RRM software architecture and describes a demulator (demonstrator+emulator) software design for evaluation of this architecture. The proposed architecture has been tested in software prior to implementation in a full hardware testbed. The demulator thus becomes in effect a reference model for the RRM framework, hosting different RRM-related algorithms, and enabling a key target of SPEED-5G, enhanced dynamic spectrum access (eDSA).

This deliverable also incorporates performance evaluations of several RRM algorithms from the SPEED-5G consortium, which are being integrated into the proposed RRM framework in order to be demonstrated in the six SPEED-5G PoCs. More specifically:

Algorithm 1: Efficient licensed-assisted access operation in same call cell based on reinforcement learning: This algorithm is designed for operation in dense heterogeneous cellular networks with Licensed-Assisted Access (LAA) small cells, capable of operating in both licensed and unlicensed spectrum.

Algorithm 2: This algorithm is related to the radio access technology (RAT)/spectrum/channel selection based on machine learning: This algorithm comprises RAT, spectrum and channel selection based on machine learning and takes into account the 3.5 GHz band for achieving better performance, especially in dense and congested 5G environments.

Algorithm 3: Radio resource allocation with aggregation for mixed traffic in a WiFi coexisted heterogeneous network: This algorithm is intended for resource allocation with aggregation to support different level of quality of service (QoS) of different traffic types in the cellular network where the WiFi network coexists.

Algorithm 4: Fuzzy Multiple Attribute Decision Making (MADM) strategy for spectrum management in multi-RAT environments: This algorithm performs a context-aware user-driven RAT selection in dense small cell environments to support a mixture of delay-sensitive and delay-tolerant applications.

Algorithm 5: Co-primary spectrum sharing in uplink SC-FDMA networks: This algorithm allows mobile network operators (MNOs) employing infrastructure sharing to efficiently share the available spectrum resources, taking advantage of information coming from the Physical and MAC layers, in order to avoid inter-operator interference and achieve improved Quality of Service (QoS) for real-time applications

Algorithm 6: Dynamic resource allocation algorithms for coexistence of LTE-U and WiFi: This algorithm provides resource allocation for co-existing LTE-U and WiFi networks to maximize throughput and hence minimize the interference.

Finally, the solutions for the evolution of the PtMP backhaul are presented in this deliverable.

Table of Contents

List of Figures

List of Tables

Abbreviations

AL	Abstraction Layer
CA	Carrier Aggregation
CM	Connection Manager
CQI	Channel Quality Index
cRRM	Centralized Radio Resource Manager
DCF	Distributed Coordination Function
dRRM	Distributed Radio Resource Manager
FBMC	Filter Bank Multicarrier
FLC	Fuzzy Logic Controller
GAA	General Authorized Access
GPS	Generalized Pattern Search
GW	GateWay
hMAC	Higher MAC
HO	Hand over
HW	Hardware
IPSec	Internet Protocol Security
ISM	Industrial, Scientific and Medical
KPI	Key Performance Indicator
LAA	Licensed-Assisted Access
LLS	Link Level Simulation
LSA	Licensed Spectrum Access
LTE-U	LTE in Unlicensed band
LWA	LTE-WiFi Aggregation
MAB	Multi-armed bandit
MAC	Medium Access Control
MADM	Multiple Attribute Decision Making
MADS	Mesh Adaptive Direct Search Algorithm
MNO	Mobile Network Operator
MP	Multiprocessing
OAM	Operation And Management
OMA-DM	Open Mobile Alliance-Device Management
OSS	Operations Support System
PAL	Priority Access Licenses
Phy	Physical interface

PSNR	Peak signal to noise ratio
QoS	Quality of Service
QoE	Quality of Experience
RAN	Radio Access Network
RAT	Radio Access Technology
RAT	Radio Access Terminal
RLC	Radio Link Controller
RPC	Remote Procedure Call
RRC	Radio Resource Controller
RRM	Radio Resource Management
SAS	Spectrum Access System
SCTP	Stream Control Transmission Protocol
SLS	System Level Simulator
SM	Spectrum Mobility
SOM	Self-Organizing Maps
SON	Self-Organized Network
SS	Spectrum Selection
SSIM	Structural similarity
SW	Software
UCB	Upper Confidence Bound
UE	User Equipment
USRP	Universal Software Radio Peripheral

1 Introduction

1.1 Objective

This document presents a performance evaluation of the most promising solutions for implementation in the radio resource manager (RRM) which have been studied in tasks T4.1, T4.2 and T4.3 and selected by SPEED-5G partners.

The SPEED-5G shortlist of solutions was given in chapter 2 of D4.2. It is important to note that the partners in the project have widely varying interests, meaning that different weights are given to different criteria. The criteria can be grouped into two general categories: (1) general criteria applying to all project partners, and (2) specific criteria applying more to commercial partners.

Different parts of the work, such as definition, modelling, and initial evaluation have been undertaken in T4.1, T4.2, and T4.3, respectively. The resulting algorithms and solutions target improvements regarding latency, throughput, signalling cost, resource balancing, as well as network availability. It is specifically the performance evaluation of these algorithms which is the main focus of this deliverable.

The design of the SPEED-5G RRM framework must be capable of supporting diverse algorithms from multiple vendors irrespective of whether the solution is centralized or distributed. In order to do this, the proposed framework and the algorithms are completely decoupled. The decoupling is achieved by automatic mechanisms for algorithm addition, removal or update operations supported by the framework.

The framework and algorithms decoupling allows the same framework to be used in a centralized or distributed environment. Depending on where it is deployed (remote or local), HW and SW restrictions may differ. The framework design allows asynchronous mechanisms so that algorithms may run in parallel. Depending on the algorithm requirements, parallelisation may be done during the whole decision process; but the framework is also capable of running the algorithms in sequential manner when the outputs of some are used as inputs to others. This particular framework structure has been design for support of the enhanced RRM algorithms proposed in SPEED-5G.

One of the main SPEED-5G concepts is that the network intelligence is allocated in the RRM entity. For that reason, RRM is connected with most of the network elements. In the same way as the LTE X2 interface is used between cells to exchange information about handovers and for self-organizing networks (SON), the SPEED-5G project has defined a new set of interfaces, in which the information is grouped in a logical way depending on the task. These interfaces have been fully defined in WP5.

The RRM (radio resource manager) is a crucial component of the SPEED-5G project, as it is responsible for carrying out the eDSA (enhanced Dynamic Spectrum Access) functions. As such, the design proposed in earlier SPEED-5G deliverables needs to be tested in software before it is implemented in a full hardware testbed. This is why the concept of a demulator is proposed. A *demulator* is a piece of software which represents the combined functionality of a *de*monstrator and an e*mulator*. Its intended function is to (1) demonstrate the working of the RRM algorithms in several scenarios closely connected to the SPEED-5G use-cases, and (2) incorporate emulation code for RRM functions which can eventually be used in the real system. The demulator is effectively a *reference model* for the RRM; in other words a functionaly correct implementation, even if not efficient and not suitable for large-system deployment.

Recall that the cRRM is a high-level component of SPEED-5G, which operates at slow time scales, typically minutes or longer. One of its principal functions is to choose an appropriate RAT (radio access technology) for a specified service, upon request from a cell which has UEs requiring to be serviced. Because of the slow time-scale of operation, the RRM exists as a self-contained entity quite high in the network hierarchy and communicates with cells over TCP.

This document has the following structure. Chapter 2 covers radio resource management from the

regulatory point of view, and from the viewpoint of network operators. This indicates the constraints under which the RRM must operate. Chapter 3 discusses the software architecture of the current implementation of the cRRM, which will be used for the SPEED-5G PoCs. Chapters 4 to 9 then give performances evaluations, mostly based on simulations, of the six chosen algorithms.

2 Radio resource management

2.1 Licensing regimes and spectrum policy

For the foreseeable future, allocation of spectrum in Europe will be the responsibility of the national regulator working within guidelines set by the European standards organisations CEPT and ETSI. The present discussion relates mainly to spectrum below 6 GHz. We will cover in this chapter:

- How licensing is done at the moment, and how it might evolve with 5G
- RRM algorithms and how they interact with the licensing
- Regulator response to 5G

2.1.1 How licensing is done at the moment

Spectrum is currently categorised as licensed, lightly-licensed, or unlicensed. Strictly though, when an operator purchases a licence, it gives them 'licence' to operate equipment within that part of the spectrum. Different equipment in the same system usually operates under different licence types. For example with a cellular system like LTE, the mobile operator holds a licence to operate base-stations in a certain part of the spectrum and under certain conditions like maximum transmit power, but the UEs in the system are operated on an unlicensed basis because they are owned by ordinary people who cannot be expected to hold a licence, and they do not stay in one location. The operator licence will specify the spectrum the UEs use and other conditions that the UEs must adhere to (like maximum transmit power), but they are still operated on an unlicensed basis.

2.1.1.1 Licensed spectrum

A spectrum licence may be one of two basic types, which are Spectrum Access (technology-neutral) and Technology-Specific. Almost all spectrum is SA, and all new bands being auctioned across Europe will be SA, but there are few bands that are TS for specific reasons such as when several operators have to share the same spectrum, or when the regulator wants to encourage fairness among operators by preventing some from using certain technologies in those bands. Examples of TS bands are 1800 MHz, which is specified for 3G and LTE only, and the low power GSM band at 1900 MHz that is constrained to 2G or LTE only. The trend is towards SA for all bands, and in say 5 years it is probable that all bands will be SA.

The licence may call up one or more interface specifications which impose additional constraints, or may contain ETSI or 3GPP specifications which the equipment must adhere to.

The licence specifies conditions that the equipment must adhere to when accessing the spectrum and such conditions would typically include whether the equipment is TDD or FDD, channel widths, maximum transmit power usually as a density (in units of dBm/MHz), and masks for unwanted emissions both in-band and out of band. The allowable transmit power is highest of all licensing options, and can be over 1 kW for a 20 MHz system. Licence auctions are held to gain sole right to use the spectrum, and the operator who wins will be issued a licence so they can operate their equipment in that spectrum. An example of use of this type of spectrum is LTE, and the sole use rights of the spectrum allows an operator to be in control of the interference and they can offer controlled QoS to their customers.

With TDD bands, the licence also specifies allowable ratios of the downlink to uplink, since signals from powerful base-stations into receivers of base-stations belonging to operators who are using the adjacent bands can cause de-sensing. So it is not possible to have a TDD system that is more than about 80% downlink or uplink. There are also guard bands put in place with low power restrictions as another mechanism to mitigate this problem. It is possible in the future that all operators will need to synchronise the frame start-times to a tight degree using a common clock.

2.1.1.2 Lightly licensed spectrum

With lightly licensed spectrum, or more accurately with equipment that is lightly licensed to operate within this spectrum, there is no concept of spectrum auctions or sole use, but an operator that uses this equipment does so in certain parts of the spectrum that are set aside for it. There are some conditions for using lightly licensed spectrum, typically the conditions are, for all items of equipment that can transmit, (a) informing the regulator of the locations of all equipment and (b) paying the regulator a small fee per year for every piece of equipment. Different operators can use the same band at the same time, but a co-ordination procedure is required to ensure the systems are geographically separated and use suitably low powers so as not to cause harmful interference to one another. These conditions make the use of lightly-licensed equipment not feasible for mobile applications, but it can be used for fixed point to multipoint. One example of lightly-licensed equipment is TV Whitespace equipment, where the location of every transmitter is fixed and known, and a database is run by the regulator to grant the equipment permission to operate and inform it of the channel and the power it can transmit without causing harmful interference to other users. Typically the transmit power is below 10 W. The QoS that is offered in lightly licensed systems is usually best effort only because of the non-exclusive use of the spectrum. Of all the types of spectrum, it is this one, lightly licensed, that is the least attractive commercially.

2.1.1.3 Unlicensed spectrum

Unlicensed equipment can be used in certain 'unlicensed bands', which are technology neutral. Typical technologies accessing unlicensed spectrum are Bluetooth and WiFi, and there is no limit to the number of different operators that can operate equipment in the bands. There is a limit on the maximum transmit power though, that depends on the spectrum band used but is typically 100 - 1000 mW, and this limit is present for safety reasons and to restrict the range of operation to a few 10s of metres. Equipment generally follows procedures for one-at-a-time access through mechanisms like listen before talk, and must adhere to certain IEEE and/or ETSI type approval specifications to avoid interference to other users of the band (such as radar systems). The QoS offered on services using this equipment is best effort.

2.1.2 RRM algorithms and how they interact with the licensing

With state-of-the-art technologies LAA and MuLTEfire, we have the situation where LTE is used in unlicensed bands and therefore they have the need to co-exist with WiFi in those bands and this is done through a process called listen-before-talk (LBT).

LAA works by having downlink carriers in both unlicensed and licensed bands, with traffic aggregated at the PDCP layer, to increase the downlink data-rate. The UE will establish uplink and downlink in a licenced band, and in addition will have a downlink carrier in an unlicensed band. So there is an 'anchor' connection in the licensed band, with the downlink enhanced by the unlicensed carrier. The anchor connection can be used for high QoS traffic, with best effort traffic sent over the unlicensed carrier.

MuLTEfire goes a step further than LAA by removing the need for a licensed anchor and the entire LTE communications is done over unlicensed bands complete with mobility support. LTE has advantages over WiFi in terms of (a) coverage through the use of HARQ mechanisms, (b) latency and (c) more efficient use of the spectrum through the use of scheduling, but it may not be possible to realise all of these advantages if operated in unlicensed bands. This is because (a) the interference is not under the control of the operator, (b) there are lower transmit power limits in these bands (Europe has lower limits than the US), and (c) there is the need in Europe for dynamic frequency selection (DFS) for radar avoidance.

Table 2 gives the impact of the SPEED-5G RRM algorithms against the licensing conditions.

RRM algorithm	Summary	Impact from/to regulation
LAA for small cells based on reinforcement learning (Chapter 7)	Chooses the best unlicensed channel depending on interference in that channel and QoS requirement of traffic	The algorithm also needs to take into account DFS requirements in certain 5 GHz bands. (It is taken into account for sensing delay and retries.)
RAT, Spectrum, and channel selection based on hierarchical machine learning (Chapter 9)	Using dRRM it chooses the best option for the downlink taking into account a pool of bands and various licensing schemes (licensed, unlicens-ed) and the need to fulfil certain traffic requirements.	Would give maximum advantage in technology neutral bands, which currently does not include 1800 MHz band, so needs opening up for filter-bank multi-carrier (FBMC). If using TDD, the restrictions on the ratio of uplink to downlink needs to be considered.
Resource Allocation and Aggregation for mixed traffic in HetNets with WLAN (Chapter 6)	Aggregation for mixed traffic in a WiFi coexisted heterogeneous network. Performs load balancing across WiFi and licensed spectrum. Using knowledge of the available capacity on the unlicensed spectrum it decides which UEs can use WiFi.	Differently from LAA/LTE-U, it is expected this is able to use Wi-Fi in both 2.4 GHz and 5GHz unlicensed bands for cellular mobile services under the current licensing regimes.
Context-aware user-driven decision making (Chapter 5)	A connection manager (CM) is introduced on the UE side to collect the various components of the context and acts according to a policy that is remotely adjusted by the network manager. Based on this, a fuzzy multiple attribute decision making (MADM) implementation of the CM is developed to select the best RAT for a set of heterogeneous applications.	The approach does not directly influence regulation, but could exploit any set of RATs using different bands and/or licensing regimes.
Co-primary spectrum sharing (Chapter 4)	Takes into consideration the users' buffer status and real-time delay constraints, as well as the operator priorities and the constraints of a realistic LTE system in order to perform uplink resource allocation in a QoS and	Provides a mechanism for two or more operators to share licensed spectrum. This requires a significant shift in regulator policy, and they would need to put safeguards in place against harmful interference.

	energy efficient manner.	
Dynamic resource allocation algorithms for coexistence of LTE-U and WiFi (Chapter 8)	Maximizes network throughput in the multi-operator scenario for 5G mobile systems by jointly considering a licensed & unlicensed band, user association and power allocation subject to minimum rate guarantee and co-channel interference threshold.	Same as above.

Table 1: SPEED-5G RRM algorithms against licensing conditions.

We can summarise as follows. Moving into 5G, we have the increased availability of software defined radio (SDR) with flexible tuning, bandwidth and modulation and coding schemes. Also there is the RRM research in this project and in others, which aims to improve spectrum utility, and assumes the use of SDR. Two changes to the regulatory regime are proposed as follows:

- Allowing more than one operator to share licensed spectrum provided mechanisms are in place to assure fairness and avoid harmful interference,
- Allow any air interface technology to operate in any part of the spectrum (i.e. bands should be technology neutral) and be flexible on occupied bandwidth.

Finally we comment on D2D communications which are growing in attraction in the scope of 5G. D2D devices would need licence-exemption. If they were WiFi-based they would already be exempt. If they were in licenced bands and the licence-exemption regulations for mobile devices in the band in question did not already cover licence-exempt use in a D2D scenario then the regulation would need amending to allow it.

2.2 Operational Support Systems (OSS)

OSS or Network Management (the terms are assumed interchangeable), is a non-trivial endeavour that must be faced by network operators, bringing about a host of new challenges beyond 3G and 4G. The number of nodes, the homogeneity of the access technologies, the conflicting management objectives, resource usage minimization, and the division between limited physical resources and elastic virtual resources is driving a complete change in the methodology for efficient network management. A 5GPPP white paper [46] covers the four management categories in 5G of provisioning, security, QoS support, and fault tolerance. These categories are not adequate, we believe, and representation is being made into 5GPPP to address this inadequacy. All these concepts will be defined, and elaborated on, in the following paragraphs.

Management categories have gone through two stages of evolution since they were standardised in the ISO model by the ITU in the 1990s. Table 2 shows the evolution of OSS through the standards, going left to right along the columns with ITU-T and then to the TMF eTOM model [49] and then to 5G.

FCAPS (ITU-T)	FAB (TMF eTOM model)	5G-PPP
Fault	Assurance	Fault tolerance
Configuration	Fulfilment	Provisioning
Accounting	Billing	Not covered - provisioning
Performance	Assurance	QoS support, fault tolerance
Security	Fulfilment	Security

Table 2: Evolution of OSS

The eTOM model consisted of only three categories which were fulfilment, assurance and billing, and they can be mapped onto the older ITU-T categories as shown. The Telemanagement Forum went on to develop interfaces for software modules to enable them to be developed by separate organisations, and also to develop business models for all the layers in the OSS stack. This is the principal forum for operators. The FAB categories are not adequate when moving into 5G. The 5G network management configurations are summarised from the 5GPPP white paper [46]- Provisioning will ensure that the network is adequately provisioned with resources that are sufficient to deal with current demand levels while maintaining QoS at an agreed level.

- Security management protects network data and its performance through accurate detection of intrusion, privacy and denial of service as well as autonomous anomaly detection.

- QoS support enables Network Slicing to support several defined QoS levels simultaneously and kept logically isolated by the same physical network.

- Fault tolerance enables the network to recognise emerging faults or error conditions and pre-emptively deal with them, or intercept unexpected faults or errors as quickly as possible to minimise any reduction in QoS.

Billing is a critical part of the network management system for operators, and indeed was one of the major categories in the FCAPS and FAB models but it does not seems to have been sustained into 5G, which is a gap that we will need to be addressed, and this will be pursued via the 5GPPP Network Management Working Group, which may eventually achieve consensus and a de-facto standard. For now, we will consider it part of the provisioning.

Another thread we need to consider here is the baseline requirements of the RRM, so that these can be evaluated in the context of OSS and also mapped to the RRM algorithms. From D4.1, the requirements are:

- Admission and prioritisation of traffic
- Load balancing
- Spectrum, RAT, and antenna selection
- Channel selection
- Inter-RAT co-ordination
- KPI monitoring and maintenance

The first five of these are facilitated by RRM algorithms that are described in detail later in this deliverable. The last is a function of the OSS.

The mapping of RRM algorithms to the current 5GPPP OSS categories is shown in Table 3.

#	chapter	RRM Algorithm	Relevant OSS category
1	7	LAA for small cells based on re-inforcement learning	Provisioning
2	9	RAT, spectrum, channel selection based on hierarchical machine learning	QoS support
3	6	Resource Allocation and Aggregation for mixed traffic in HetNets with WLAN	Provisioning; QoS support
4	5	Context-aware user-driven decision making	QoS support; provisioning; billing
5	4	Co-primary spectrum sharing	Provisioning; QoS support
6	8	Dynamic resource allocation algorithms for coexistence of LTE-U and WiFi	Provisioning; QoS support

Table 3: Mapping of RRM algorithms to context and framework category

From the table we can see that all of the RRM algorithms map to the Provisioning and QoS support OSS categories, as would be expected. The project does not address fault tolerance or security.

3 RRM overview

3.1 Novelty of RRM algorithms

Each of the RRM algorithms has made progress beyond the state of the art, and this progress is shown in Table 4 below.

#	chapter	Algorithm	Novelty (beyond SoA)
1	7	LAA for small cells based on reinforcement learning	Re-inforcement learning, combined with FBMC, leads to higher spectrum utilisation
2	9	RAT, spectrum, and channel selection based on hierarchical machine learning	Prioritization of channel assignment based on QoS requirements based on channel segregation and machine learning aspects. Multi-RAT support
3	6	Resource Allocation and Aggregation for mixed traffic in HetNets with WLAN	By calculating the access index integrating the estimated user mobility, user information and traffic types, it identifies appropriate users to connect to WLAN. By two-step allocation, while it supports robust connection (with primary carrier in licensed band), it can also support higher rate (with supplemental carrier either in licensed or unlicensed band).
4	5	Context-aware user-driven decision making	Combination of (1) Fuzzy logic to cope with the uncertainty level and lack of information typically associated with UEs and (2) MADM to efficiently combine the heterogeneous components of the context
5	4	Co-primary spectrum sharing	- Joint co-primary spectrum sharing and resource allocation - Consideration of MNO priority in spectrum sharing - Constraints of uplink resource allocation allocation of sets of contiguous resource blocks per user-specified LTE procedures of user requests for uplink transmission grants and buffer status reporting - Estimation of the packet delays based on the received scheduling requests by the users
6	8	Dynamic resource allocation algorithms for coexistence of LTE-U and WiFi	- Jointly considering licensed & unlicensed band, user association and power allocation subject to minimum rate guarantee and co-channel interference threshold. - Nonlinear Optimization by Mesh Adaptive Direct Search (NOMAD), for resource allocation in 5G Heterogeneous Networks, to reach sub-optimal solution.

Table 4: How the RRM algorithms push beyond the state of the art

3.2 Mapping of RRM algorithms to RRM framework

Figure 1 shows the RRM framework from D4.2; note at the bottom right the categories of SPEED-5G algorithms listed in the previous section.

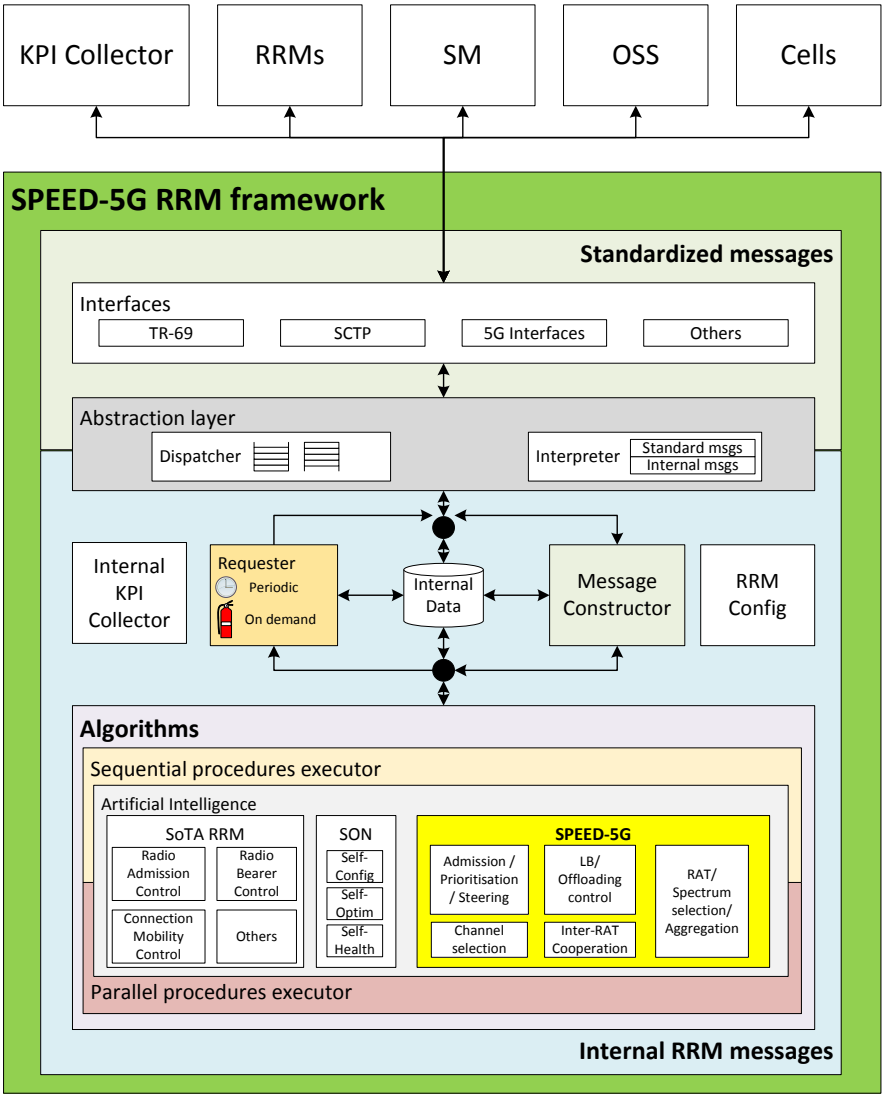


Figure 1: Showing where OSS fits into the RRM framework, and showing at the bottom right-hand side of the figure, the five Speed-5G RRM categories

The mapping of the algorithms to the framework categories is shown in Table 5 below.

	RRM Algorithm	Context	Map to Speed-5G framework category
1	LAA for small cells based on re-inforcement learning	DL, cRRM/dRRM	Load balancing and offloading; Channel selection
2	RAT, spectrum, channel selection based on hierarchical machine learning	DL, cRRM/dRRM	Channel selection; RAT, spectrum selection, aggregation
3	Resource Allocation and Aggregation for mixed traffic in HetNets with WLAN	DL, cRRM/dRRM	Admission, prioritisation, steering; Load balancing and offloading
4	Context-aware user-driven decision making	DL, UE	RAT, spectrum selection, aggregation
5	Co-primary spectrum sharing	UL, dRRM	Admission, prioritisation, steering; Inter-RAT co-operation
6	Dynamic resource allocation algorithms for coexistence of LTE-U and WiFi	DL, cRRM	Load balancing and offloading

Table 5: Mapping of RRM algorithms to framework categories

3.3 Software design

The software architecture being used to implement the cRRM was detailed in D4.2. Here only a minor update is required. This chapter therefore documents some design decisions made subsequent to D4.2.

The main function of the cRRM is to interface between hardware and RRM algorithms, and thus to act as an abstraction layer presenting a uniform API to devices. To achieve this, a choice to use XML-RPC as the only communication protocol was made (http://xmlrpc.scripting.com/). Though it is an old standard, XML-RPC has the advantage of wide support in most programming languages. The cRRM API is thus of the remote procedure call type. Typical functions implemented include:

1. Register cell – every cell must register before further requests will be handled.
2. Report signal-to-interference ratio (SIR) as measured by a UE.
3. Request new channel – this triggers an algorithm to compute a better channel
4. Goodbye – this de-registers a cell.

To ensure portability, and to enable high-level software design, the chosen implementation language is Python (version 3.5 or higher). The software makes use of the multiprocessing (mp) module of python, for maximum flexibility and portability over different thread implementations. The design is built upon several base classes, intended to be subclassed to cover various specific instances. The result is the following main classes:

1. DuplexQueue: uses two SimplexQueues of the mp module, to allow duplex message-passing between RRM components. Buffering is handled transparently. Each pair of components will install message handlers at each end of the DuplexQueue.
2. Logger: handles all system logging, writing tsv files for post-processing.
3. RRM_base: base class from which instances of RRM can be derived. Provides basic messaging, cell registration, and logging facilities, not explicit resource management functions.
4. MAC_base: base class from which instances of the MAC can be derived.
5. Cell_base: base class from which instances of cells can be derived. Stores information on RATs of which the cell is capable.
6. UE_base: base class from which instances of UEs can be derived.

The cRRM is designed to be able to run the SPEED-5G RM algorithms, even when source code is not available. It achieves this by running the algorithm from a compiled executable as a separate process, communicating via command-line arguments.

3.4 The demulator

In order to develop and debug the cRRM code, and separate software tool called the demulator (demonstrator+emulator) was developed (Figure 2). This emulates a set of cells, and communicates with the cRRM exactly as a set of real cells would do. The GUI displays on the left the complementary cumulative distribution of signal-to-interference ratio, and on the right, a heat map displays the SIR visually. Messages to and from the cRRM are scrolled in the box at the bottom. The demulator also forms SPEED-5G PoC4.

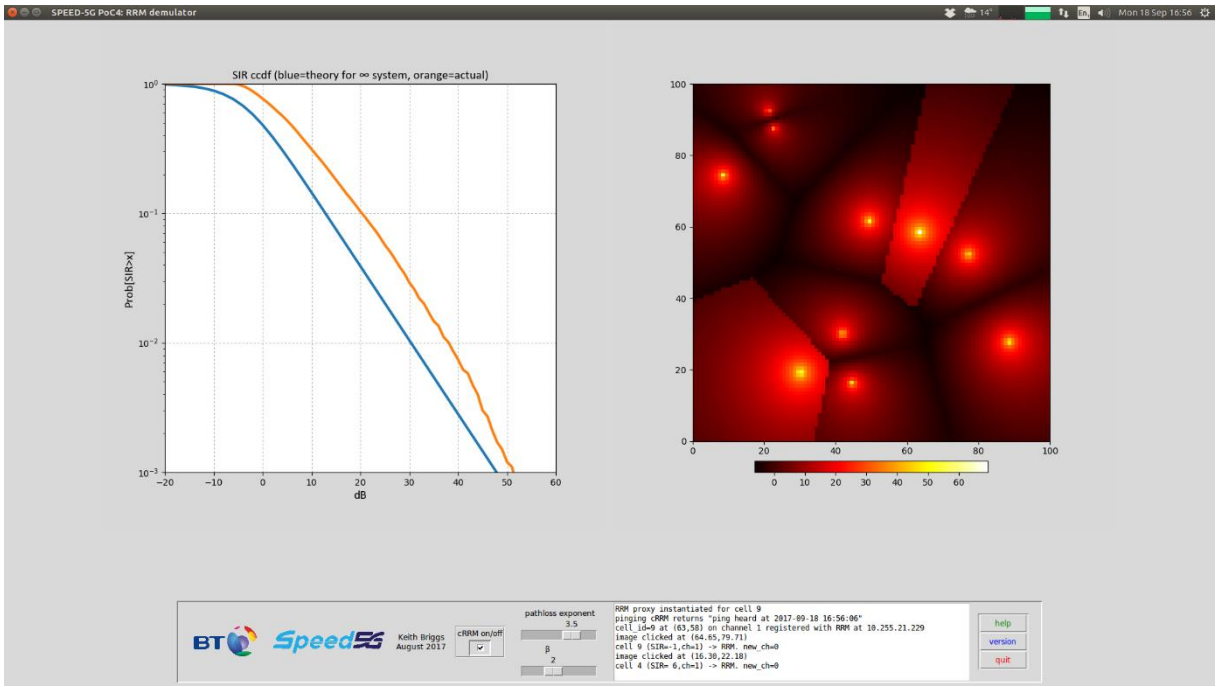


Figure 2: The demulator with its GUI

4 Co-primary spectrum sharing

In this section, we provide updated performance evaluation results of the co-primary uplink spectrum sharing algorithm, which was firstly introduced in D4.1 [1] and D4.2 [11]. We consider a common pool of shared spectrum, for the case of two Mobile Network Operators (MNOs) operating in the uplink direction of LTE and employing infrastructure sharing. The objective of the spectrum sharing process is to assure exclusive access to the shared spectrum in order to avoid inter-operator interference and enhance Quality of Service (QoS) provision. This scheme can be applied to both macro and small cells.

The proposed algorithm considers the main constraints in uplink Single Carrier – Frequency Division Multiple Access (SC-FDMA) resource allocation, i.e., the allocation of contiguous sets of resource blocks of the localized physical layer to each user, and the imperfect knowledge of the users' uplink buffer status and packet waiting time.

In D4.1, the optimal resource allocation was formulated as a discrete connected cake-cutting problem, where different agents are allocated consecutive subsequences of a sequence of indivisible items [1]. This problem is NP-hard, therefore a suboptimal algorithm was introduced in D4.2, which performs resource allocation using information on the estimated uplink packet delay, the average delay and data rate of past allocations, as well as the required uplink power per resource block and the MNO's priority over the spectrum [11].

4.1 State of the art

In this section we provide a brief overview of the literature in the area of co-primary spectrum sharing.

The authors in [2] consider a common pool of shared spectrum between two MNOs. A centralised scheduler is responsible for managing the sharing procedure, with the aim to assure exclusive access to the shared spectrum and avoid inter-operator interference. The centralised controlled allocates the shared bands considering the Channel Quality Index (CQI), in order to maximise the cell capacity. Therefore, more elaborate performance evaluation metrics, such as the mean delay, to reduce the packet timeout rate of real-time applications, as well as the rate and delay of past allocations, to improve the system fairness, are not considered. Moreover, the need for the centralised controller to coordinate the spectrum sharing process, although it can guarantee reliable interference protection between the sharing MNOs, has several disadvantages, including the increased complexity and the higher signalling overhead.

A cooperative co-primary spectrum sharing scheme between two MNOs that employ infrastructure sharing, is proposed in [3]. The spectrum pool considered includes both dedicated spectrum for each MNO, as well as the shared spectrum. The partitioning of the available spectrum to orthogonally and non-orthogonally shared frequency sub-bands is performed in an adaptive manner, taking into consideration the channel conditions of the users. The MNOs share their inter-RAN CSI via fibre optic-based backhaul as they simultaneously may need to use the shared band. Also, an inter-RAN precoder is applied to minimise inter-operator interference. Similarly to the previous case, the spectrum sharing is performed only considering the channel quality information, and not more QoS-related metrics.

In [4] and [5], the spectrum pool is available for simultaneous access of two MNOs in the same area, using beamforming as the coordination technique. The most serious challenge in the case of beamforming is the need to share, channel quality information among the base stations (BSs) /access points (APs) of different MNOs, as well as interfering channel quality information among BSs of one MNO and users of the other MNO. Such information exchange needs to be carried out in a timely manner and with sufficient accuracy to maximise the benefits provided by the application of beamforming.

The authors in [6] consider a common pool of radio resources, which is shared between MNOs of a multi-operator small cell network, with the aim to achieve long term fairness of spectrum sharing without the need for coordination between the small cell APs. A learning technique based on Gibbs sampling is used by a decentralised control mechanism in order to allocate the appropriate amount of spectrum to each AP. Five algorithms are compared addressing co-primary multi-operator resource sharing under heterogeneous traffic requirements and the performance is assessed through extensive system-level simulations. The main performance metrics are user throughput and fairness between operators. However, the traffic model considered for the performance evaluation of the different algorithms assumes continuous constant rate transmission. Therefore, the effect of the proposed spectrum sharing scheme on the mean delay, which is of ultimate importance in the case of real-time applications is not evaluated.

4.2 Simulation assumptions

In the context of T4.3, additional simulations were performed, in order to evaluate the performance of the co-primary uplink spectrum sharing algorithm, taking into consideration the simulation parameters defined in WP5.

The performance evaluation parameters are as follows (updated parameters are marked in *italics*):

Parameters	**Values**
Carrier Frequency	*2 GHz*
Physical layer parameters	*Channel Bandwidth: 20 MHz,* Subframe length (T_{sf}): 1 ms, *Number of resource blocks (N_{RB}^{UL}): 100*
Resource block format	Number of subcarriers per resource block (N_{SC}^{RB}): 12, Number of symbols per resource block (N_{symb}^{UL}): 7, Subcarrier spacing: 15 kHz
Reference Signal transmissions	2 Reference Signal transmissions per subframe
TDD configuration	Configuration 1, DL:UL 3:2
Modulation and Coding Schemes	QPSK ½, 16-QAM ½ and 64-QAM ¾
Path loss model	*ITU-R UMi based model [according to Table B.1.2.1-1 in TR36.814], UMa, NLOS*
Transmitter antenna gain (UE)	0 dBi
Receiver antenna gain + connector loss (BS)	*5 dBi*
Thermal Noise Spectral Density	*-174 dBm/Hz*
UE Noise Figure	*9 dBm*
Interference Margin [7]	1 dB
Control Channel Overhead [7]	0 dB
Shadowing	*Log normal, σ=6dB*
Fading	Rayleigh
Maximum UE transmission power	*20 dBm*
Target received power ($P_{0,PUSCH}$)	-57 dBm
Uplink path loss compensation factor (α)	0.7

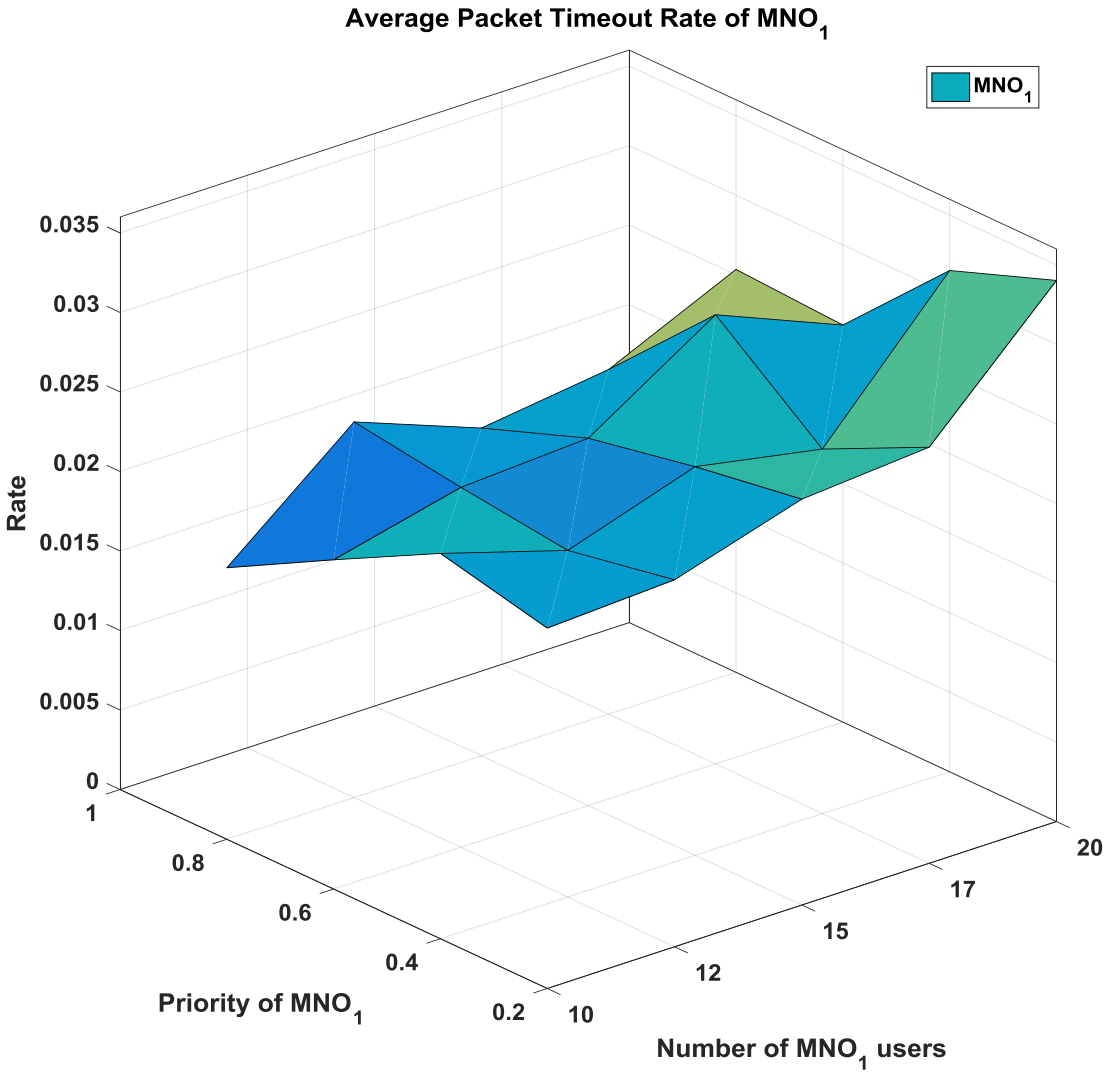
Maximum tolerable delay ($d_{th,i}$)	20ms
RLC mode	Unacknowledged mode (UM)
Traffic model [8]	H264 video traffic QCIF 176×144
Protocol header sizes	RTP/UDP/IP with ROCH Compression: 3 bytes, PDCP: 2 bytes, RLC: 3 bytes, MAC: 2 bytes, CRC: 3 bytes
Moving average calculation factor (β)	0.2
Inter-site distance	*100 m*
Antenna Height (UE)	*1.5 m*
Antenna Height (BS)	*20 m*
Average building height (h)	*20 m*
Street width (W)	*20 m*
Simulation time	67 s

Table 6: Performance evaluation parameters

In terms of signalling overhead, the proposed algorithm does not burden the system with additional overhead, as it solely relies on the standard signalling used for resource allocation; that is, channel quality information, buffer status requests and signalling requests.

4.3 Performance evaluation

Figure 3 depicts the average packet timeout rate with respect to the number of users per Mobile Network Operator (MNO) and the shared spectrum access priority of MNO_1, which is referred to as p_1. The priority of MNO_2, is defined as $p_2 = 1 - p_1$. The total number of users N is in the range of $[20, 40]$, while the number of users of each MNO is defined as: $N_1 = \left\lfloor \frac{N}{2} \right\rfloor$ and $N_2 = N - N_1$. The packet timeout rate is defined as the number of packets that expire in the unit of time, since in real-time applications excessive scheduling delay results in the need to discard the expired packets, as they are no longer useful at the receiver side. As shown in Figure 3, for both MNOs the packet timeout rate follows an increasing course with the number of users, since an increasing number of users results in higher congestion, which leads to longer queuing time and, eventually to more packet expirations. The effect of the shared spectrum priority, p_k, is also shown in this figure. As it can be seen, the mean packet timeout rate of MNO_1 follows a declining course with the increase of p_1, while the packet timeout rate of MNO_2 follows the opposite course, since $p_2 = 1 - p_1$. This is a result of the fact that the proposed algorithm takes into consideration the shared spectrum access priority of each MNO during the resource allocation.


Average Packet Timeout Rate of MNO1
MNO1
Rate
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
1
0.8
0.6
0.4
0.2
Priority of MNO1
10
12
15
17
20
Number of MNO1 users

(a)

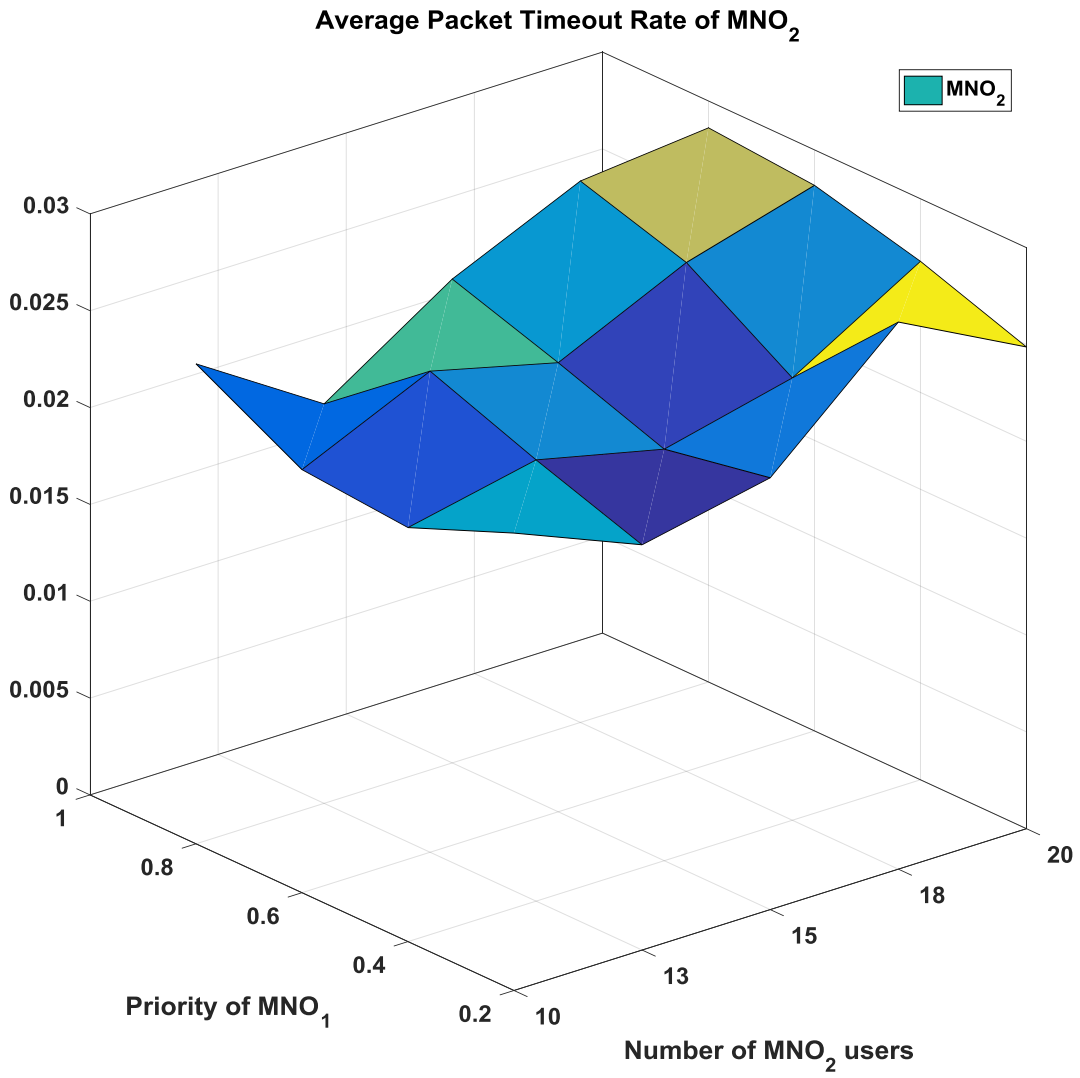
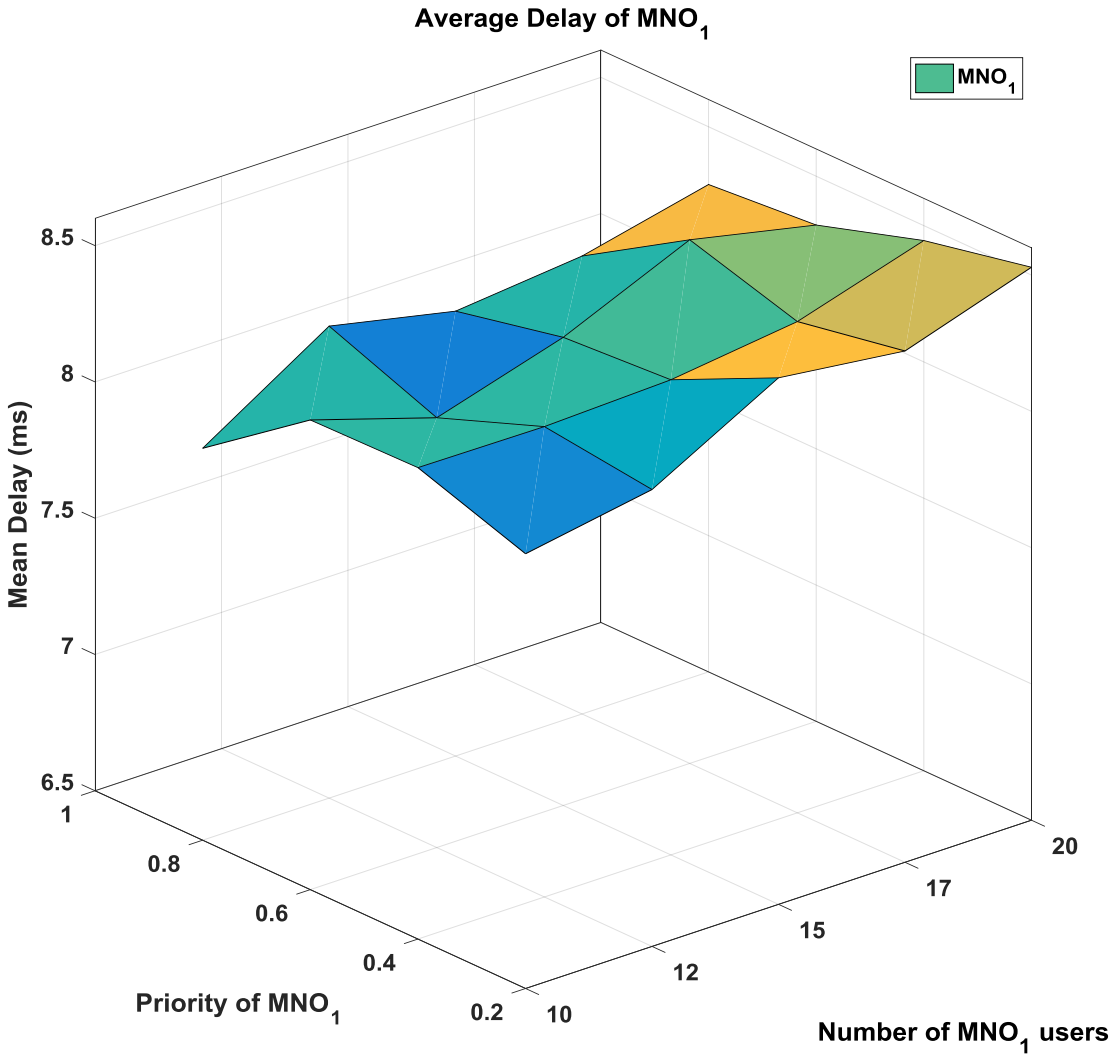


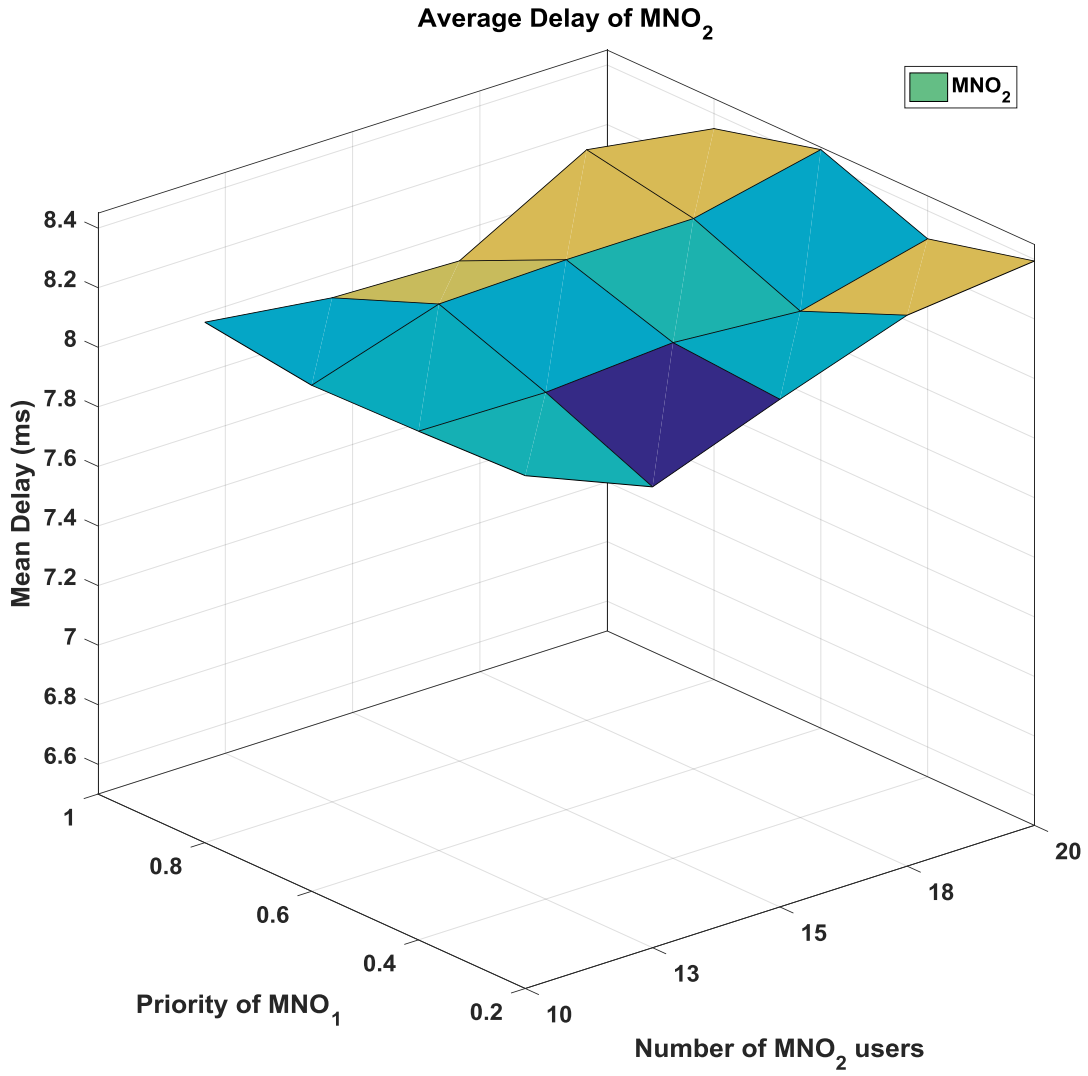
Figure 3: Average packet timeout rate of (a) MNO_1 and (b) MNO_2 versus the number of users and the shared spectrum access priority.

(b)

Figure 4 depicts the average packet delay versus the number of users and the shared spectrum access priority. As it can be seen, for both MNOs, the average packet delay follows a slightly increasing course with the number of users, as a result of the increased congestion and the need for longer queuing times. Moreover, it can be seen that the average delay follows a slightly declining course with the increase of the shared spectrum priority of the MNO, due to the fact that the proposed algorithm prioritizes users based on the priority of the MNO they are associated with.


Average Delay of MNO1
MNO1
Mean Delay (ms)
8.5
8
7.5
7
6.5
1
0.8
0.6
0.4
0.2
Priority of MNO1
10
12
15
17
20
Number of MNO1 users

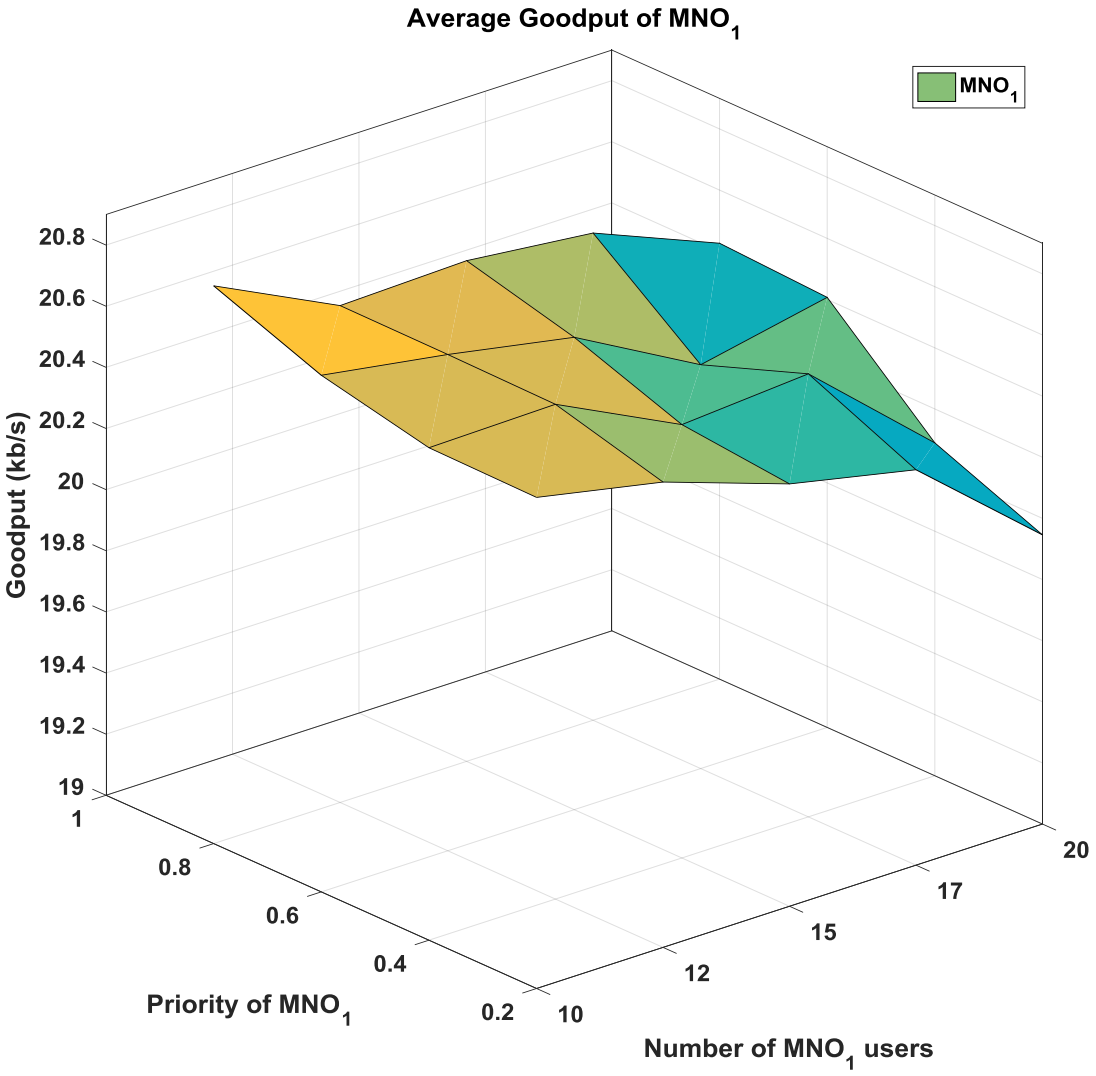
(a)



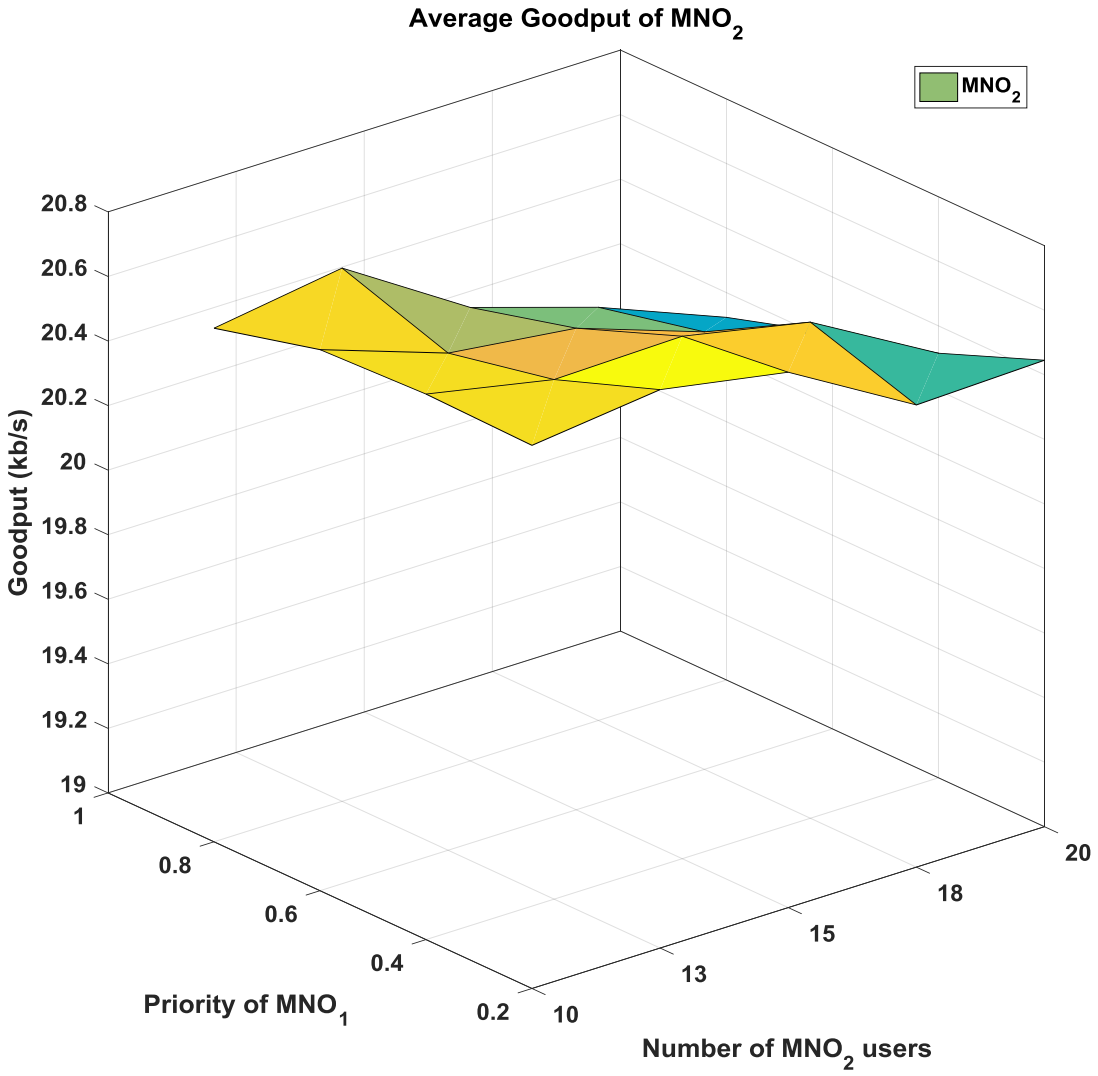
(b)

Figure 4: Average delay of (a) MNO_1 and (b) MNO_2 versus the number of users and the shared spectrum access priority.

Figure 5 depicts the average goodput of both MNOs. The goodput is defined as the throughput at the application layer, that is, the rate of useful bits that reach the application layer in the unit of time. As it can be seen, in both cases the goodput follows a declining course with the increase of the number of users, as a result of the increasing congestion, which leads to excessive packet delays and timeouts. Moreover, it is shown that the average goodput increases with the increase of the MNO spectrum access priority.


Average Goodput of MNO1
MNO1
Goodput (kb/s)
20.8
20.6
20.4
20.2
20
19.8
19.6
19.4
19.2
19
1
0.8
0.6
0.4
0.2
Priority of MNO1
10
12
15
17
20
Number of MNO1 users

(a)



(b)

Figure 5: Goodput of (a) MNO_1 and (b) MNO_2 versus the number of users and the shared spectrum access priority.

Figure 6 depicts the fairness of both operators with respect to an increasing number of users of both operators and the shared spectrum access priority. Fairness is evaluated using the Jain Index of Fairness, which is defined as $FI = \frac{(\sum_{i \in UE} Th_i(t))^2}{|UE| \cdot \sum_{i \in UE} Th_i^2(t)}$, where $Th_i(t)$ is the throughput of user i. As it can be seen, fairness of the system that employs the proposed algorithms is considerably high, as a result of the fact that the proposed algorithm takes into consideration the average packet delay $\overline{D}_{i,k}^{UL}(t)$ and the average data rate $\bar{R}_{i,k}^{UL}(t)$ in the user prioritization, therefore favouring users who have experienced high average delay and low average data rate in past allocations.

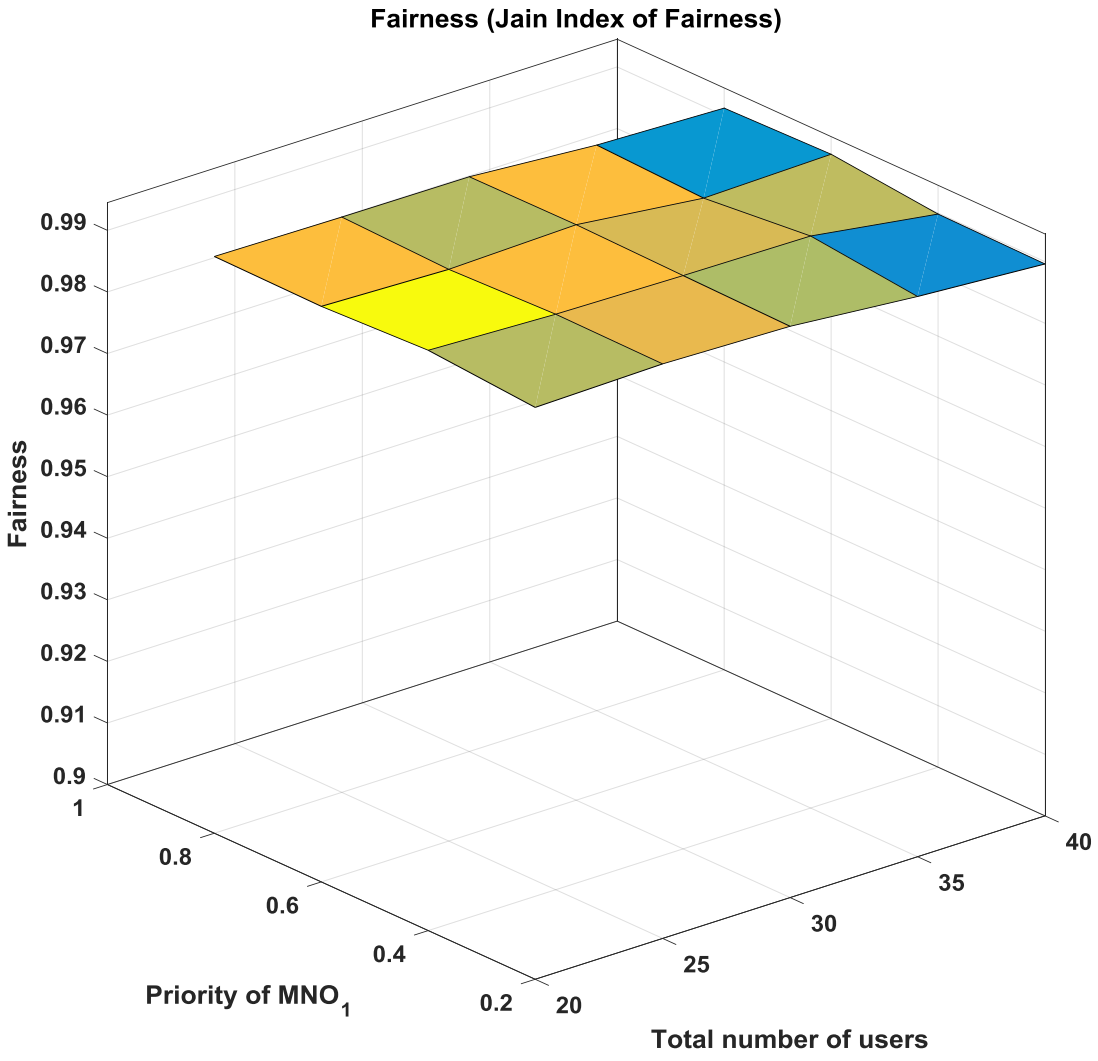


Figure 6: Fairness (Jain index) versus the total number of users and the MNO priority.

4.4 Summary

In this section, the performance of the proposed co-primary uplink spectrum sharing algorithm was evaluated, considering the simulation parameters defined in WP5. The performance evaluation results indicated the effect of the increasing number of users per MNO, as well as each MNO's spectrum access priority, to Key Performance Indicators including packet timeout rate, mean delay, goodput and fairness.

5 Context-aware user-driven RAT selection

This chapter exploits the generic framework built in [10] to enable a context-aware user-driven mode of operation. As an initial use case, the framework was applied in [11] to perform an intelligent offloading to WLAN for a mixture of best-effort and voice-over-IP (VoIP) applications. Motivated by the proven usefulness of the proposed framework in supporting spectrum management, this chapter extends it to support all emerging LTE-based candidates to access unlicensed bands (i.e., LTE-U, LAA and muLTEfire) together with their co-existence mechanisms (i.e., adaptive duty cycle and LBT). Based on this extension, the proposed strategy is applied to efficiently exploit licensed and unlicensed bands to support interactive (that is, delay-sensitive) video sessions in ultra-dense environments [12].

5.1 State-of-the-art on RAT selection

The RAT selection problem has been extensively investigated in the context of 3GPP (e.g., WCDMA and LTE) and non-3GPP (e.g., WLANs) access networks [13]-[15]. Most of these works considered a common radio resource management based on a tight coupling architecture of the considered RATs. However, such coordinated approach may not be valid to complement new 5G RATs. As a matter of fact, recent 5G architectures do not integrate most of legacy RATs (e.g., GSM, WCDMA and WLANs) or prefer to interconnect them only at the core network level because a full integration would be too costly in terms of multi-RAT measurements and interworking [16]. This means that a user-driven decision-making would be much more suitable particularly when the selection decision must be made fast for specific applications (e.g., delay-sensitive) or under particular circumstances (e.g., fast degradation of radio conditions). However, the limited amount of information that is typically available to user equipments (UEs) through beacons and pilot channels (e.g., signal strength and signal-to-noise-ratio (SNR)) is not enough to select the best RAT in a given context (e.g., application requirements, terminal capabilities and network constraints). This calls for a form of network assistance to inform UEs about all the relevant information that would be needed to perform an efficient decision.

To cope with the lack of information and uncertainty level associated with UEs, most proposals have relied on fuzzy logic to infer the best RAT out of the available pieces of information. In this respect, some works developed a single fuzzy logic controller (FLC) that is fed with all the relevant attributes of the available RATs [14][17]. The main limitation of these works is their lack of scalability. As a matter of fact, when more RATs and attributes are considered, the number of inference rules increases exponentially. Other works proposed a single FLC per RAT that combines the set of available radio parameters with the various components of the context [18][19]. Such approach improves scalability, but is not flexible. First, it assumes that all data is fuzzy, while some attributes may be obtained precisely (e.g., QoS requirements). Second, using fuzzy logic rules to combine both QoS- and context-related attributes does not offer flexibility in adjusting the importance (i.e., weight) of each attribute depending on the operating conditions.

To offer higher degree of flexibility to fuzzy logic, few proposals have combined it with multiple attribute decision making (MADM) that is known for its ability to efficiently combine various heterogeneous attributes [20]-[22]. Such combinations make no distinction between the radio parameters that are directly related to the achievable QoS (e.g., signal strength, bandwidth and SNR) and the various components of the context (e.g., velocity, battery consumption and price) whose importance varies from case to case. It follows that the suitability level of each RAT to meet the set of QoS requirements cannot be explicitly assessed before considering the various contextual information, which means that, in practice, the selected RAT may not meet the target QoS level. Another key limitation of these schemes is that they all assumed that the decision-maker has full access to all required information (e.g., MADM attributes and weights). Therefore, they are not valid

to tackle the unique issues and constraints associated with a user-driven mode of operation, e.g., which attributes can be in practice obtained by UEs or how the network may adjust the controlling parameters to achieve a target strategy. Clearly, all the architectural constraints should be considered early in the design of any feasible user-driven decision-making.

Therefore, the first main contribution of this piece of work is to construct a novel functional architecture to enable context-aware user-driven operation in multi-RAT environments. The proposed architecture relies on a UE connection manager (CM) that selects the best RAT per a policy that is remotely adjusted by a network policy designer. Correspondingly, the second contribution is to develop a feasible fuzzy MADM implementation of the CM to select the best RAT for a set of heterogeneous applications. Fuzzy logic is used to first estimate the out-of-context suitability level of each RAT to support the various QoS requirements. Second, an MADM component combines these estimates with the various components of the context (e.g., user preferences and operator policies) to derive the in-context suitability levels of the considered RATs.

5.2 System model

An ultra-dense environment is considered, where a set of *K* available RATs ($\{RAT_k\}_{1\le k\le K}$) are exploited by various UEs to establish a set of *L* applications ($\{A_l\}_{1\le l\le L}$). Each UE is assumed to belong to one of *S* subscription profiles ($\{P_s\}_{1\le s\le S}$). Without loss of generality, *L=2* applications are considered, namely interactive (i.e., delay-sensitive) video with stringent QoE requirements and FTP (i.e., delay-tolerant) transfer with loose QoS requirements, together with *S=2* subscription profiles, namely *Bronze* and *Gold* with limited and unlimited credits, respectively.

To support the considered applications, a set of LTE small-cells are deployed on top of some existing WLAN APs. The various small-cells are assumed to be dual-access i.e., they can jointly use licensed and unlicensed bands. Therefore, the following *K=3* RATs are considered:

- *Licensed/LTE*: use the newly deployed small-cells in licensed bands,
- *Unlicensed/WLAN*: offload part of the traffic to the existing WLAN APs,
- *Unlicensed/LBT*: use the newly deployed small-cells to share unlicensed bands with the existing WLAN. To assess the unique potentials offered by unlicensed bands, a stand-alone muLTEfire RAT [23] is assumed.

In the remainder of this chapter, the band will be dropped from the notation for the sake of simplicity. Therefore, the above combinations will be denoted as LTE, WLAN and LBT, respectively.

5.2.1 Target behaviour

Access to licensed bands (i.e., LTE) is assumed to be paid (i.e., consumes some units of the credit), while access to unlicensed bands (i.e., WLAN and LBT) is free of charge. The aim is to maximize usage of unlicensed bands provided that the application requirements are met. This means that the FTP transfer associated with loose QoS requirements can be always established on unlicensed bands, while interactive video sessions need to maximize usage of unlicensed bands as long as the associated requirements are met.

5.2.2 Functional architecture

To efficiently tackle the considered problem of network selection in multi-RAT environments, the generic framework previously proposed in [10] will be instantiated. According to its functional architecture described in Figure 7, a CM is introduced at the UE to implement a given decision-

making policy. To this end, it exploits the relevant components of the context available locally (e.g., velocity and battery level) and a radio characterisation of each available RAT in terms of a set of short-term attributes (e.g., SNR and load) obtained, for example through beacons and some medium- and long-term attributes (e.g., cost and regulation rules) stored in a policy repository together with all the policy-related data. The content of the policy repository may be retrieved in practice from a local instance following a pull or push mode using, for example the Open Mobile Alliance-Device Management (OMA-DM) protocol [24]. To offer higher flexibility to the network manager, a policy designer entity builds and updates the policy repository content based on measurement reports collected from the various UEs and some potential network-level constraints (e.g., operator strategy and regulation rules). Finally, the CMs of different UEs may collaborate to further improve their individual performances.

To achieve the target behaviour set out in Section 5.2.1, a fuzzy MADM implementation of the CM is developed in the next section.

5.3 Connection manager (CM): fuzzy MADM decision-making

This section describes the functional architecture depicted in Figure 7, to perform a context-aware exploitation of all available RATs to support each of the considered applications. The implementation instantiates the generic fuzzy multiple attribute decision making (MADM) methodology proposed in [10].

5.3.1 Out-of-context suitability levels

In accordance with the guidelines given in [10], a separate FLC is designed to estimate the suitability of each RAT to support the Quality of Experience (QoE) requirements of the interactive video application:

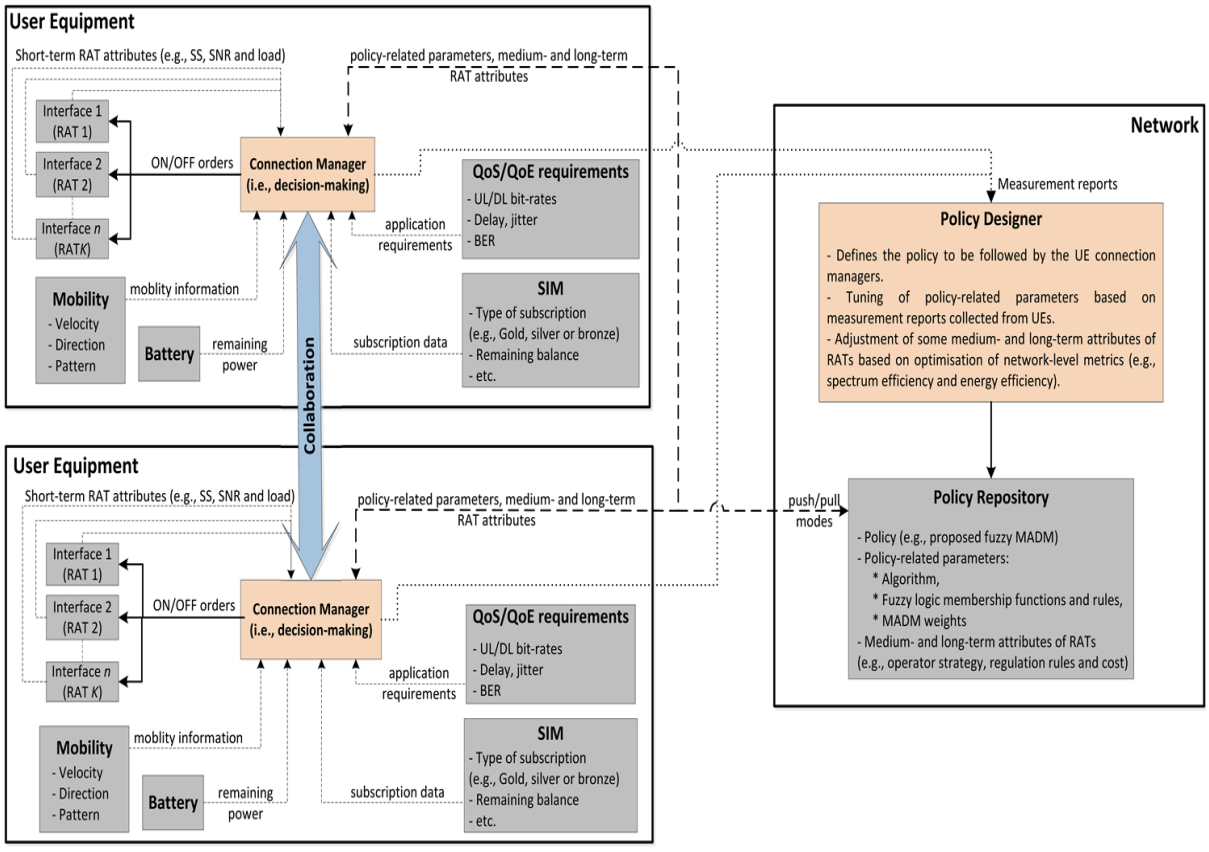


Figure 7: Functional architecture of the proposed context-aware user-driven framework

- LTE: The designed FLC and corresponding membership functions are described in Figure 8. In particular, the following input parameters are considered:

  1) *RSRQ*: the reference symbol received quality that captures the radio and interference conditions.

  2) *T_Sched*: the average time each packet received from upper layers waits before being scheduled. It reflects the load condition on the eNodeB and may be broadcasted in one of its system information blocks (SIBs). A non QoS-aware scheduler (e.g., proportional fair (PF)) is initially assumed, which means that all packets are treated equally.

- WLAN: The proposed FLC and associated membership functions are described in Figure 9. Specifically, the set of input parameters is designed as follows:

  1) *SINR*: the signal-to-interference-and-noise-ratio of the AP beacon that reflects the radio and interference conditions.

  2) *T_ACK*: the delay between receiving a packet from higher layers till successfully transmitting it (i.e., receiving the corresponding ACK from the receiver). It jointly captures the impact of the load served by the AP and channel contention (e.g., retransmissions due to collision). No 802.11e QoS support is considered initially, which means that all traffic types share the same MAC queue.

  3) *Drop_R*: the rate of dropped packets. A packet is dropped whenever the MAC queue is full or the maximum retransmission limit is reached for an unacknowledged packet.

- LBT: This option inherits some features (e.g., MAC scheduler) from LTE and others (e.g., contention-based access) from WLAN. In this respect, the designed FLC described in Figure 10 is fed with the following inputs:

  4) *RSRQ:* the reference symbol received quality that captures the radio and interference conditions.

  5) *T_Access:* the average time each packet received from upper layers waits before being granted access to the channel. Compared to *T_Sched*, this metric includes the additional delay (e.g., back-off) that may be introduced by the LBT procedure after a packet gets scheduled.

  6) *NACK_R:* the ratio of NACKs out of the HARQ-ACK feedback values available for the first sub-frame of the latest data burst. It reflects the likelihood of a collision as reflected by the rule selected by 3GPP to update the contention window size [25].

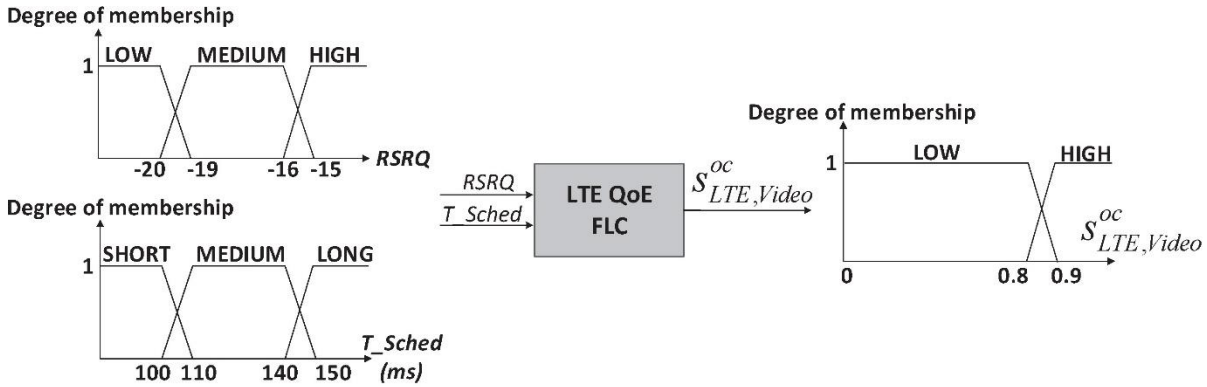


Figure 8: LTE/Video FLC

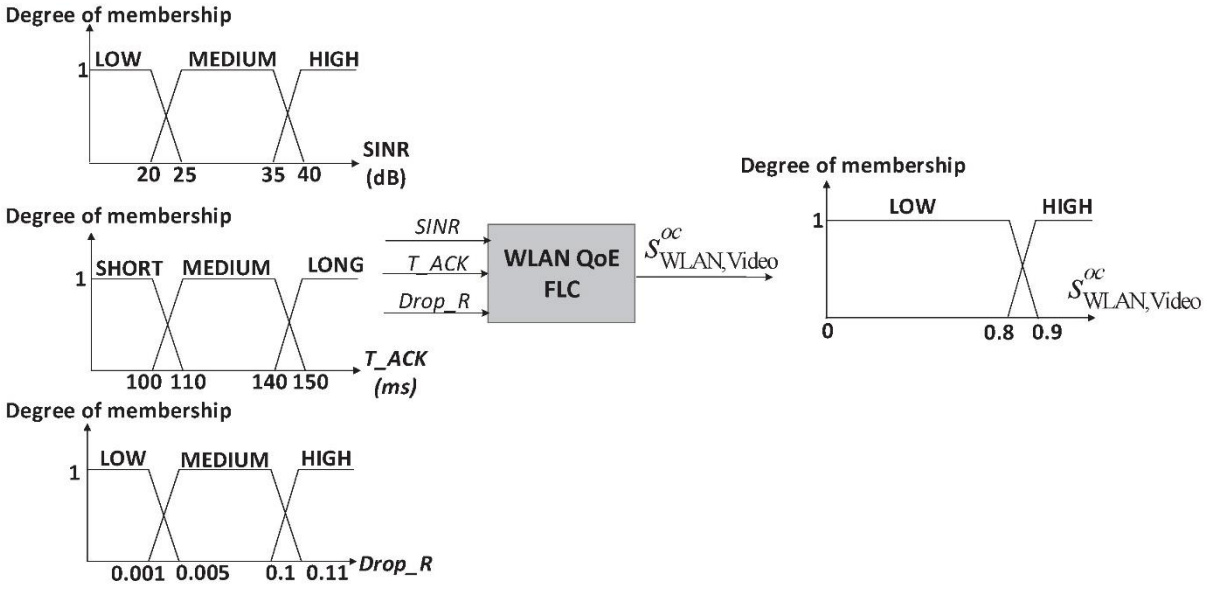


Figure 9: WLAN/Video FLC

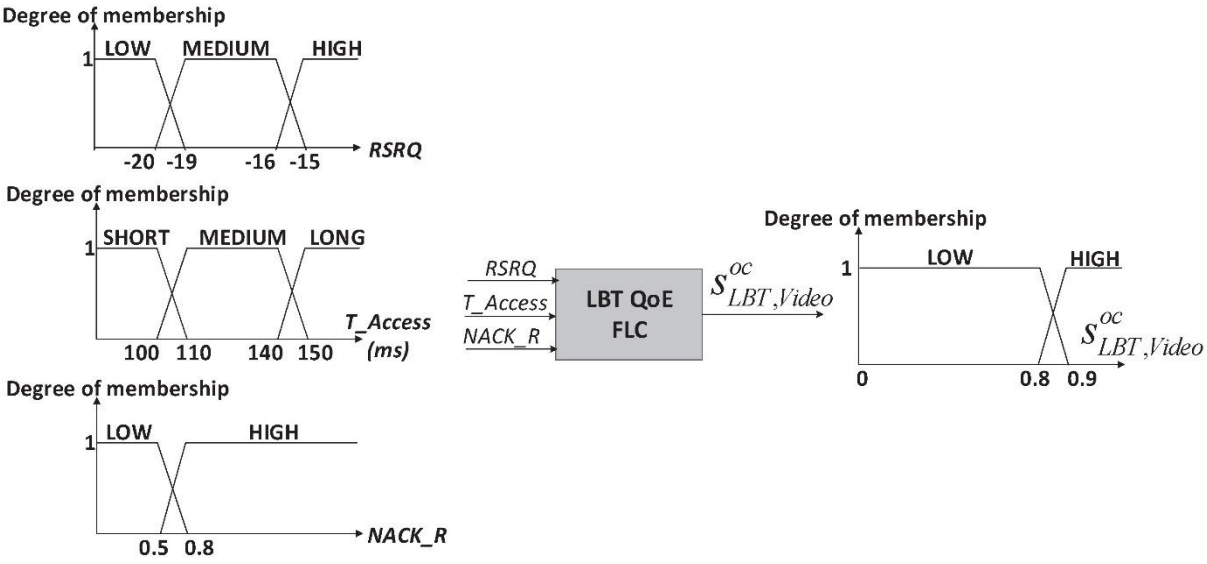


Figure 10: LBT/Video FLC

For all FLCs, the required inferences rules have been designed based on a sensitivity analysis to the various combinations of the input parameters, which is omitted for the sake of brevity. This mimics the adjustment performed by the policy designer of Figure 7 based on the actual performance

measurements collected from the various UEs. Finally, the defuzzification process is based on the commonly used centroid method for its accuracy [26].

Finally, it is worth pointing out that, for the other considered application (i.e., FTP file download), there is no need to develop separate FLCs. Both WLAN and LTE are assumed to meet the associated loose QoS requirements as long as the corresponding UEs are associated/attached.

5.3.2 In-context suitability levels

In this section, the previous estimates are combined with the components of the context to derive the in-context suitability levels. To particularly cope with the heterogeneity of the context components, a methodology is developed based on MADM [27].

In this respect, for each $k \in \{1,\ldots,K\}$, RAT_k is characterized in terms of the following $M=4$ attributes:

- ✓ $s_{k,l}^{oc}$: the out-of-context suitability to meet the set of application requirements. Recall that this is the output of the previous sub-section.
- ✓ $cost_k$: the monetary cost of RAT_k.
- ✓ $power_k$: the power consumption level when using RAT_k.
- ✓ $range_k$: an assessment of the range to reflect the appropriateness from the UE velocity perspective.

Therefore, for each application A_l, the RATs can be fully characterized in terms of a KxM decision matrix $\mathbf{D_l}$ whose element $d_{k,m}^{l}$ denotes the performance of RAT_k in terms of the m-th attribute:

$$\mathbf{D_{Video}} = \begin{array}{c} \\ \mathbf{LTE:} \\ \mathbf{WLAN:} \\ \mathbf{LBT:} \end{array} \begin{array}{c} \mathbf{QoE} \\ \mathbf{s}_{LTE,Video}^{oc} \\ \mathbf{s}_{WLAN,Video}^{oc} \\ \mathbf{s}_{LBT,Video}^{oc} \end{array} \begin{array}{c} \mathbf{cost} \\ HIGH \\ LOW \\ LOW \end{array} \begin{array}{c} \mathbf{power} \\ HIGH \\ MEDIUM \\ HIGH \end{array} \begin{array}{c} \mathbf{range} \\ LARGE \\ SMALL \\ SMALL \end{array} \quad (1)$$

$$\mathbf{D_{FTP}} = \begin{array}{c} \\ \mathbf{LTE:} \\ \mathbf{WLAN:} \\ \mathbf{LBT:} \end{array} \begin{array}{c} \mathbf{QoS} \\ HIGH \\ HIGH \\ HIGH \end{array} \begin{array}{c} \mathbf{cost} \\ HIGH \\ LOW \\ LOW \end{array} \begin{array}{c} \mathbf{power} \\ HIGH \\ MEDIUM \\ HIGH \end{array} \begin{array}{c} \mathbf{range} \\ LARGE \\ SMALL \\ SMALL \end{array} \quad (2)$$

Note that, compared to licensed access (i.e., LTE), unlicensed RATs (i.e., WLAN and LBT) are qualified as cheaper and smaller for both applications. In turn, the LTE-based options (i.e., LTE and LBT) are judged as more power consuming than WLAN. Recall that, for all RATs, the first attribute associated with video (i.e., $s_{k,Video}^{oc}$) is the output of each of the FLCs designed in the last sub-section and is, therefore, a real number. Note that the out-of-context suitability level to support FTP transfer (i.e., $s_{k,FTP}^{oc}$) is always set to HIGH to reflect the loose QoS requirements.

To adjust the relative importance of the various attributes, a vector $\mathbf{w_l^s}$ of M weights ($\{w_{l,m}^{s}\}_{1\le m\le M}$) is introduced for each l-th application and s-th subscription profile:

$$\mathbf{w_{Video}^{B}} = \mathbf{w_{FTP}^{B}} = \begin{pmatrix} HIGH \\ HIGH \\ LOW \\ LOW \end{pmatrix} \quad (3)$$

$$\mathbf{w}^{\mathbf{G}}_{\mathbf{Video}} = \mathbf{w}^{\mathbf{G}}_{\mathbf{FTP}} = \begin{pmatrix} HIGH \\ LOW \\ LOW \\ LOW \end{pmatrix} \quad (4)$$

where B and G stand for the *Bronze* and *Gold* subscription profiles, respectively.

Note that, for both applications, the cost attribute is more relevant for the *Bronze* user (i.e., $w^{B}_{l,cost} = HIGH$). In turn, the velocity and power attributes are not initially considered for the sake of simplicity (i.e., $w^{s}_{l,power} = w^{s}_{l,range} = LOW$).

Finally, the vector $\mathbf{s}^{\mathbf{ic,s}}_{\mathbf{l}}$ of in-context suitability levels ($\left\{s^{ic,s}_{k,l}\right\}_{1\le k\le K}$) is obtained by combining the various attributes and weights as follows:

$$\mathbf{s}^{\mathbf{ic,s}}_{\mathbf{l}} = \begin{bmatrix} s^{ic,s}_{1,l} \\ \vdots \\ s^{ic,s}_{k,l} \\ \vdots \\ s^{ic,s}_{K,l} \end{bmatrix} = \overline{\mathbf{D}_{\mathbf{l}}} \cdot \mathbf{w}^{\mathbf{s}}_{\mathbf{l}} \quad (5)$$

where $\overline{\mathbf{D}_{\mathbf{l}}}$ is the matrix of normalized attributes $\overline{d^{l}_{k,m}}$ that are calculated as $\overline{d^{l}_{k,m}} = d^{l}_{k,m} / \max_{k}(d^{l}_{k,m})$ for benefit attributes (i.e., QoS/QoE and range) and $\overline{d^{l}_{k,m}} = \min_{k}(d^{l}_{k,m}) / d^{l}_{k,m}$ for cost attributes (i.e., cost and power).

5.3.3 Decision-making

Based on the previous sub-section, the RAT that maximises the in-context suitability level is selected for the *l*-th application and *s*-th subscription profile:

$$k^{*}(l,s) = \underset{k\in\{1,\ldots,K\}}{\operatorname{argmax}} (s^{ic,s}_{k,l}) \quad (6)$$

To track the variability in the various attributes (e.g., radio conditions and contextual information), the CM implements the following functionalities based on the above criterion:

- **Spectrum selection (SS):** the best RAT is selected at the time of establishing each of the considered applications.
- **Spectrum mobility (SM):** a hand over (HO) to the best RAT is performed during sessions. Two types of HO are considered, namely event-triggered emergency HOs (i.e., whenever any of the application requirements is not met) and periodic comfort HOs triggered each *ΔT*. In both cases, if a better RAT_{k^*} is identified (i.e., $s^{ic,s}_{k^*,l} > s^{ic,s}_{serving,l}$), the UE is reconfigured to use it.

5.3.4 Considered environment

To gain insight into the relevance of the proposed approach in supporting the considered applications, a set of system-level simulations have been performed using the ns-3 simulator [28].

- To model ultra-dense deployments, a single LTE macro-cell overlaid by a set of buildings is considered. Each building is structured according to the dual-stripe layout [29] i.e., as two stripes of rooms with a corridor in-between, which corresponds in practice to e.g., the set of stores inside a shopping mall. The various propagation losses (i.e., indoor-indoor, outdoor-outdoor, indoor-to-outdoor and vice versa) are modelled using the hybrid building model that combines several well-known propagation loss models [30].

- A set of small-cells are dropped randomly inside each building and allowed to operate in both licensed and unlicensed bands. The licensed band is shared with the macro-cell according to a co-channel deployment, while the unlicensed band is shared with a WLAN 802.11n AP placed inside the same room. As an illustrative example, Figure 11 describes the signal-to-interference-and-noise-ratio (SINR) map obtained in the licensed band when two 20-room buildings are considered with one small-cell placed in each room.

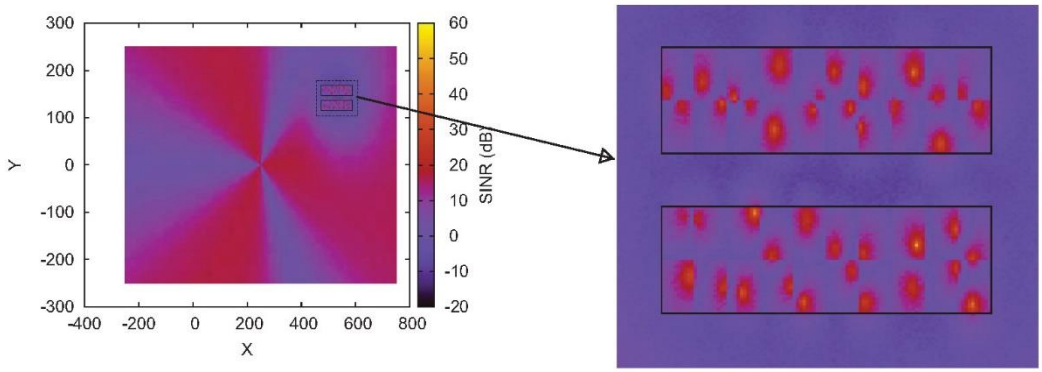


Figure 11: Illustrative example of SINR map, licensed band

5.3.5 Traffic model

The *L=2* applications considered in Section 5.2 are modelled as follows:

- *Interactive video:* During an interactive (e.g., live streaming) session, the UE receives a given video sequence from a remote host over a UDP transport session. The associated set of requirements is characterized in terms of a maximum end-to-end delay of D_{max}=100 ms and frame loss ratio of L_{max}=*0.1%*. In this respect, the video receiver accepts only in-sequence frames whose end-to-end delay does not exceed D_{max}. Any other frame is dropped with no subsequent retransmission. An approach to simulate the actual video transmission and assess the perceived QoE will be presented in the next sub-section.

- *FTP transfer*: an ON/OFF model is used to model FTP download sessions (i.e., ON periods) and the inactivity intervals in-between (i.e., OFF periods). Both ON/OFF periods are exponentially distributed with rate λ. To control the load generated by FTP sessions, each ON period is assumed to consume a percentage ρ of the capacity of the in-use radio link with loose QoS requirements.

5.3.6 Evaluation of video QoE

Given the high cost incurred by subjective tests, this section proposes to perform an objective assessment of the video QoE. In this respect, the Evalvid framework [31], whose simplified architecture is described in Figure 12, has been integrated.

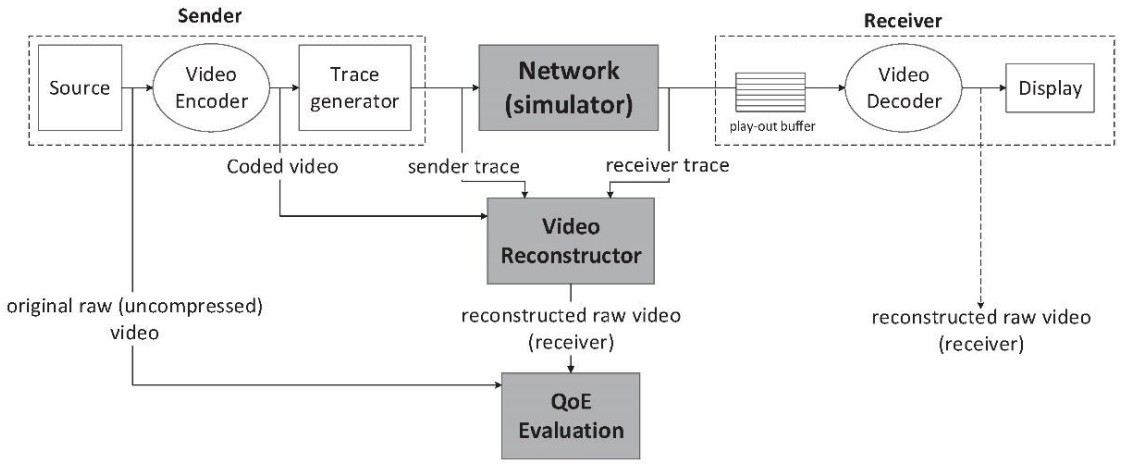


Figure 12: A simplified architecture of the Evalvid framework

On the sender side, a digital video sequence is encoded, packetized and transmitted over a simulated network. On the receiver side, a play-out buffer is optionally used for jitter reduction before the received sequence is decoded and displayed. Both the sender and receiver keep track of the time-stamp and type of each sent/received packet in separate trace files. These trace files are combined with the original encoded video to reconstruct the uncompressed raw video as it would be perceived by the receiver. Based on a comparison between the original and reconstructed raw sequences, the QoE is evaluated based on the following metrics:

- The peak signal to noise ratio (PSNR) which has been extensively used for its simplicity. It provides a measure of the similarity between the distorted video sequence and its original counterpart based on a frame-by-frame comparison [31].

- The structural similarity (SSIM) index which focuses on the structural information loss to which the human visual system is strongly sensitive. It computes the mean, variance and covariance of small patches inside each frame and combines them into a distortion map [32].

To reflect the current trends in high-end video entertainment, the test video is the popular 2-min Big Buck Bunny sequence with e.g., 16+ million views on YouTube. Its uncompressed raw sequence is downloaded from [33] and encoded with H264 (Main Profile, L4) at 1080p @24fps, which generates a variable bit-rate stream with an average of 2.5 Mbps.

5.3.7 Benchmarking

To assess the efficiency of the proposed framework in selecting the best RAT, the fuzzy MADM strategy developed in section 5.3 is applied in the following scenarios:

- *Offloading*: LTE small-cells are allowed to operate only in licensed bands. To exploit unlicensed bands, part of the traffic may be offloaded to WLAN.

- *Sharing:* Small-cells can operate in licensed and unlicensed modes. The unlicensed band is shared with the existing WLAN based on the LBT procedure.

- *Offloading+Sharing*: This scheme combines the previous two options. Therefore, unlicensed bands can be jointly exploited through offloading and sharing.

Additionally, to benchmark the performance of the proposed approach, the following baseline is considered:

- *Fixed*: A traditional scheme that always assigns delay-sensitive traffic (i.e., interactive video) to licensed bands (i.e., LTE) and delay-tolerant traffic (i.e., FTP transfer) to unlicensed bands (i.e., WLAN or LBT).

5.3.8 Simulation parameters

The proposed fuzzy MADM strategy approach has been evaluated under the set of simulation parameters defined in WP5 [34]. The additional/specific simulation parameters used in this chapter are summarized in Table 7.

Parameters	Values
Common	
Simulation environment	ns-3
Layout	Dual-stripe (see section 5.3.4)
Propagation model	The hybrid building model [30]
K (number of RATs)	3 RATs, namely: - LTE (licensed) - muLTEfire (unlicensed) - WLAN (unlicensed)
LTE-based RATs (i.e., LTE and muLTEfire)	
Number of LTE macrocells	1
Duplex method	FDD
Carrier frequency of LTE macrocells	2.12 GHz
Number of dual-access small-cells	1/room
Operating (Carrier) Frequency	2.12 GHz (licensed); 5.18 GHz (unlicensed)
System bandwidth per carrier	20 MHz (licensed and unlicensed)
DL MAC Scheduler	Proportional fair (PF)
WLAN	
Number of WLAN APs	1/room
WLAN standard	802.11n
WLAN channel	36 (i.e., 5.180 GHz)
Chanel bandwidth	20 MHz
Applications	
L (number of applications)	2 (i.e., interactive video and FTP)
Video sequence/format	*Big Buck Bunny* animation encoded with H264 (Main Profile, L4) at 1080p @24fps, which generates a variable bit-rate with an average of 2.5 Mbps
FTP traffic model	ON/OFF with rate λ
S (Number of subscription profiles)	2, namely:

	- *Bronze* associated with limited credit of 500 Mbits. - *Gold* associated with unlimited credit.
Number of video applications/subscription profile	1
Number of FTP transfer sessions/room	1
Direction of traffic	downlink-only
Simulation time	650 s

Table 7: Simulation assumptions and parameters for evaluating the fuzzy MADM strategy

5.3.9 Performance evaluation

This section evaluates the effectiveness of the proposed fuzzy MADM strategy in performing a context-aware exploitation of all available bands in the considered environment. To this end, the QoE metrics introduced in section 5.3.6 are evaluated.

A single-room scenario of the dual-stripe layout described in section 5.3.4 is initially considered with a specific focus on the interactive video application. In this respect, two *Bronze* and *Gold* video sessions are established in the same room, where the *Bronze* subscription is associated with a limited credit of 500 Mbits. To model the existing load on the unlicensed band, a single FTP transfer with $1/\lambda$=2 s is established and maintained over WLAN. Therefore, the percentage of used capacity (i.e., ρ) defined in section 5.3.5 can be used as an indicator of the existing load on WLAN.

Figure 13 shows the average PSNR perceived during video sessions for each of the proposed variants and baseline defined in section 5.3.7 as a function of the WLAN load when comfort HOs are triggered each Δt=200 ms. Note that the performance achieved by the *Bronze* user is shown separately for each scheme. In turn, the performance of the *Gold* user is shown only once as it does not depend on the used scheme. Finally, it has been checked that the QoE performance in terms of SSIM exhibits a similar behaviour and is therefore omitted for the sake of brevity.

The results first show that the performance of the *Gold* user remains unchanged for all considered schemes. This is because according to the MADM settings of section 5.3.2, the cost attribute is irrelevant for the *Gold* subscription profile, and thus the paid well-regulated LTE can be used all the time regardless of the WLAN load. Note that the baseline *Fixed* achieves an equal performance as it statically assigns interactive video sessions to LTE.

Next, the QoE perceived by the *Bronze* user is analysed. To better understand the relative performance achieved by each scheme, Figure 14(a) and Figure 14(b) plot the evolution of the instantaneous (i.e., per frame) PSNR perceived by the *Bronze* user when the WLAN load is low (i.e., 40%) and high (i.e., 100%), respectively.

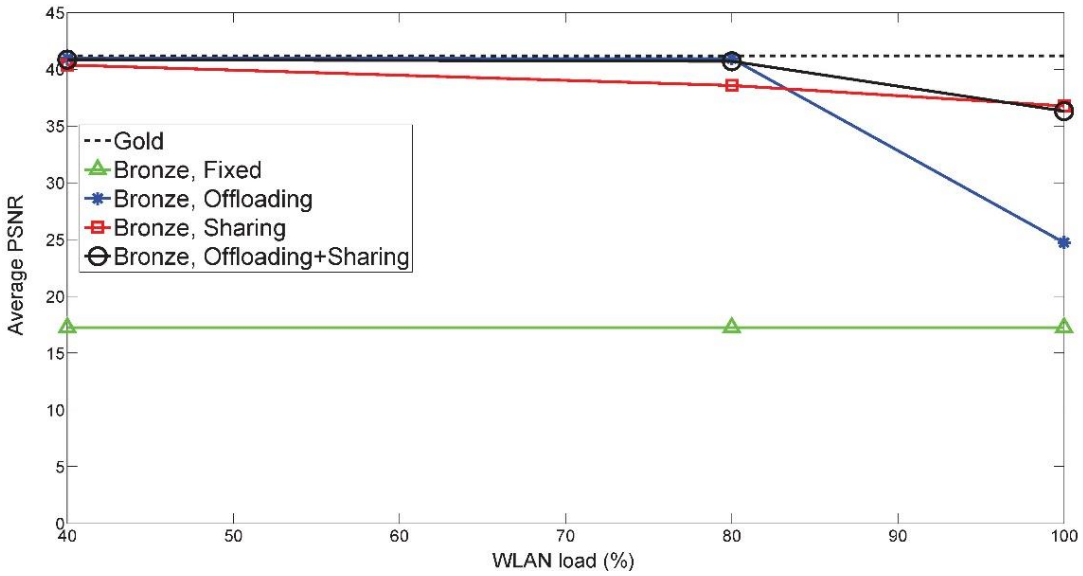
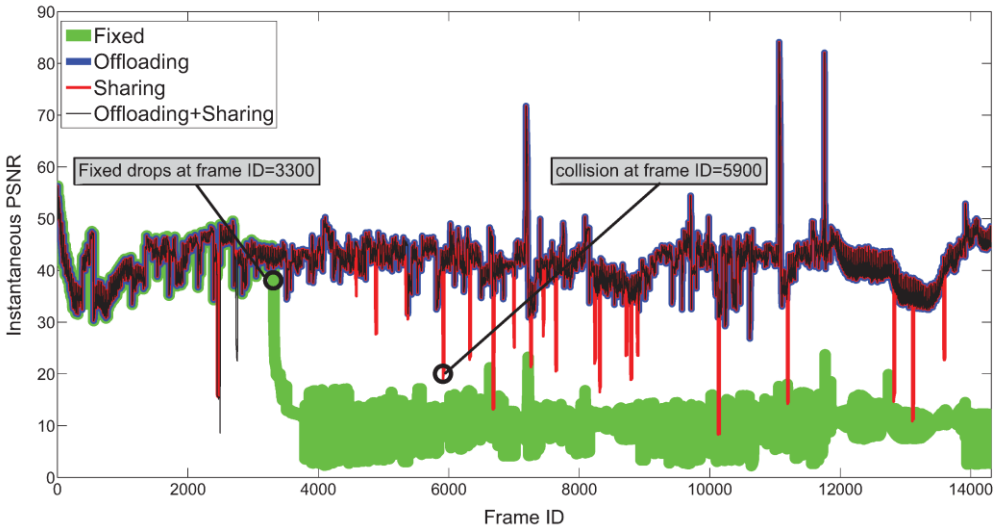
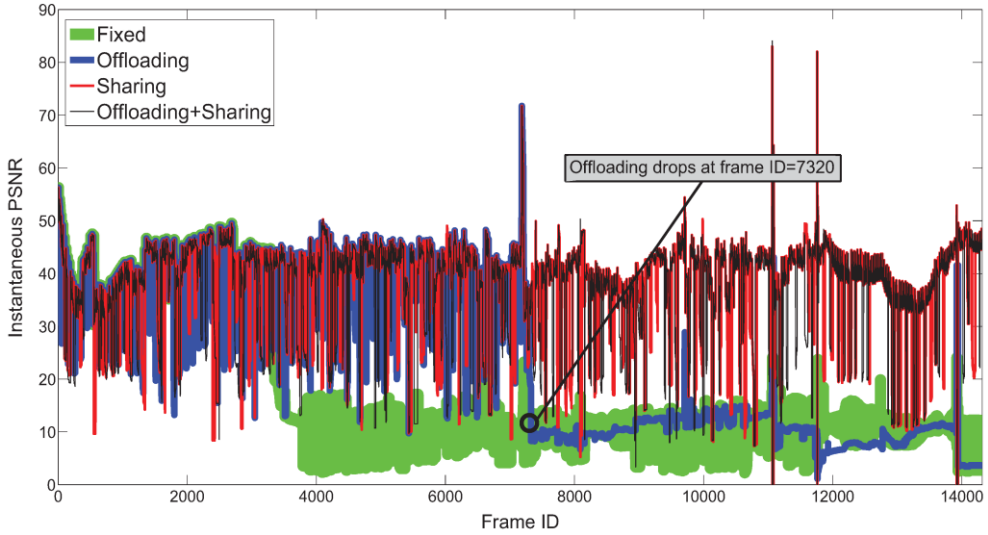


Figure 13: Impact of WLAN load on the perceived QoE in terms of PSNR



(a) WLAN load=40%



(b) WLAN load=100%

Figure 14: Evolution of the instantaneous PSNR of the Bronze user

(a) Offloading *(b) Sharing* *(c) Offloading+Sharing*

Figure 15: The actual frame perceived by the Bronze user at frame ID=5900, WLAN load=40%

The first key observation in Figure 13 is that the improvement introduced by the unlicensed access is significant (i.e., ranging from 100% to 135%) for all considered options (i.e., *Offloading*, *Sharing* and *Offloading+Sharing*). This is because the limited credit of the *Bronze* user is not enough to deliver the whole video stream over the paid LTE access. Therefore, when only the licensed band is exploited (i.e., *Fixed*), the video session shortly gets dropped after the credit is depleted at frame ID=3300 as highlighted in Figure 14 (a). The additional exploitation of the unlicensed band saves some of the valuable credit units and helps to maintain the video session till the end, which significantly improves the average PSNR.

When comparing the various options to exploit unlicensed bands, it can be observed in Figure 13 that the best option strongly depends on the existing load on WLAN. At low loads, the unused WLAN capacity is enough to accommodate the video traffic, which justifies the good performance maintained by *Offloading* in Figure 14(a). On the contrary, contending with WLAN for the unlicensed band results in some collisions as identified by the sharp PSNR drops experienced by *Sharing* in Figure 14(a). To assess the impact of these degradations on the end-user perception, Figure 15 shows the actual frame perceived by the *Bronze* user at the position indicated in Figure 14 (a) (i.e., frame ID=5900). It can be seen that unlike the other schemes, the reconstructed frame by *Sharing* is strongly degraded. As the load increases, *Offloading* is forced to often use the licensed band, which results in an ”out-of-credit” drop at frame ID=7320 as highlighted in Figure 14 (b). In turn, *Sharing* manages to maintain the video session till the end, which justifies the achieved gain in Figure 13 with respect to *Offloading* for high loads. When *offloading* and *sharing* are jointly considered (i.e., *Offloading+Sharing*), the combined approach efficiently switches between both options depending on the operating conditions as can be observed in Figure 13.

5.4 Summary

This chapter proposes a novel context-aware user-driven strategy to efficiently exploit all available bands and licensing regimes in ultra-dense environments without any prior knowledge about each combination. It relies first on fuzzy logic to estimate the suitability of each available RAT to support various applications subject to the uncertainty level associated with UEs. Then, a fuzzy MADM approach is developed to combine these estimates with the heterogeneous components of the context to assess the in-context suitability levels. Based on this metric, a spectrum management strategy combining two spectrum selection (SS) and spectrum mobility (SM) functionalities is developed to select the best RAT in a given context. As an illustrative use case, the proposed strategy is applied to support interactive video sessions for a set of *Bronze* (i.e., limited-credit) and *Gold* (i.e., flat-rate) subscription profiles. The results reveal that the proposed approach always assigns *Gold* users to the well-regulated licensed band, while switches *Bronze* users between licensed and unlicensed bands depending on the operating conditions of each RAT. This results in a significant improvement (i.e., 100% to 135%) of the achieved QoE compared to a traditional scheme that exploits only licensed bands. Then, a comparative study is conducted between the two available options to exploit unlicensed bands, namely *Offloading* and *Sharing*. The results show that the best

option strongly depends on the existing load on WLAN. At low loads, *Sharing* suffers from few collisions, while *Offloading* runs ”out-of-credit” at high loads. Therefore, a combined approach is proposed to efficiently switch between both options, which achieves the best QoE for all considered loads.

6 Resource allocation and aggregation for mixed traffic in HetNets with WLAN

In deliverable D4.2, we presented an algorithm for radio resource allocation and aggregation for UEs capable of dual connectivity to the cellular and the Wi-Fi network in a heterogeneous network. Focusing on the expected significant growth of real-time traffic (e.g., almost 70% of mobile data traffic will be from the real-time traffic by 2021 [47]), a utility function approach is used to model the different QoS of heterogeneous traffic types in the algorithm. While the primary carrier is allocated from the cellular spectrum to all UEs, secondary carriers can be allocated to UEs requiring high data rates. In the system model presented in D4.2, the Wi-Fi link could be also utilized for the supplemental carriers to boost UEs' data rate. For allocation of secondary carriers, the user mobility is estimated based on relevant user information and the access index is calculated by integrating the user mobility, the Wi-Fi signal strength and the traffic type. While UEs to access the Wi-Fi network are selected, UEs of lower mobility, stronger Wi-Fi signal strength, and inelastic traffic are likely to be chosen as more proper UEs for the Wi-Fi network. For the rest of UEs (i.e., unselected for the Wi-Fi link), the secondary carriers are allocated from the cellular spectrum based on the utility function. This proposed approach is to intend the efficient resource utilization to transmit heterogeneous traffic across multiple RATs. In D4.2, with our initial simulation results, it was shown that the proposed algorithm can contribute to better QoS satisfaction to UEs of different traffic types and better utilization of the Wi-Fi network for moving UEs. In this chapter, additional results are included based on the new reference algorithm presented in [45] to describe the effects of the proposed approach integrating the utility function and the user mobility estimation on the system performance.

6.1 Simulation assumptions

To assess the performance of the proposed approach, we consider the scenario with UEs capable of the dual connectivity for the cellular and Wi-Fi networks in a heterogeneous network with WLAN. While the primary and secondary carriers can be allocated to UEs, UEs can access multiple carriers (including the Wi-Fi link) with carrier aggregation capability. The proposed resource allocation and aggregation approach has been evaluated in MATLAB under the set of simulation parameters summarized in Table 7.

Parameters	Values
ISD	500m
Duplex method	FDD
Direction of traffic	downlink-only
Carrier frequency of cellular networks	800 MHz, 2 GHz
System bandwidth per carrier	20 MHz (10 x 2MHz @ 800 MHz), 40 MHz (10 x 4MHz @ 2 GHz)
Max Transmit power	24 dBm (for 20 MHz)
Path Loss model	800 MHz: 119.6 + 37.2 log(d) [km] [dB] 2 GHz : 128.1 + 37.6 log(d) [km] [dB]
Number of WLAN APs	2
Chanel bandwidth	20 MHz
Max Transmit power	20 dBm

Path Loss model	2.4 GHz: 140.7+37.6 log(d)[km]+ 21 log(2.4/2)[dB]
# of WiFi devices	Varying [1,6]
# of UEs	Varying [4, 10]
Traffic type	Mixed (Elastic, Inelastic)
UEs' Mobility	Varying {0, 3, 10, 30, 60} [km/h]

Table 8: Simulation assumptions and parameters for the Wi-Fi coexisted heterogeneous network

6.2 Performance evaluation

This section evaluates the effectiveness of the proposed resource allocation and aggregation algorithm for the multi-RATs heterogeneous network supporting mixed traffic.

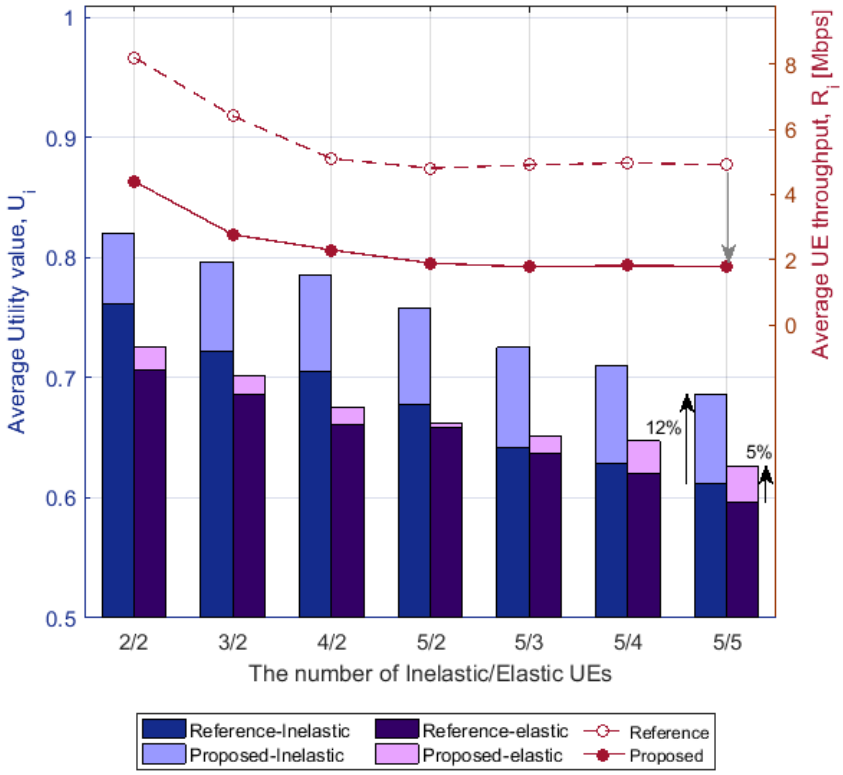


Figure 16: Comparison of performance of the proposed and reference algorithm (No mobility): 1) left axis for the bar graph–application utility values of inelastic/elastic UEs, 2) right axis for the curve graph-UE's average throughput

Firstly, in order to focus on the support of heterogeneous traffics, UEs are assumed to be static without mobility. Figure 16 shows the performance of the proposed and reference algorithm in terms of UE's average application utility value and the average UE's throughput. While the bar graph in the left axis presents the average utility level of inelastic/elastic traffic UEs, it is shown that the proposed approach can provide better QoS to both inelastic (~12%) and elastic UEs (~5%) compared to the reference regardless of the number of inelastic/elastic UEs. However, in the aspect of average UEs' data rate, the reference strategy shows superior performance. While the reference approach gives weight to the traffic load of RATs and the channel qualities to allocate resource to UEs, it

provides higher throughput to UEs. Nevertheless, since it does not consider the heterogeneous requirements of different traffic types, it could provide data rates to inelastic traffic much higher than to the required rate while it provides rates to elastic UEs much lower than the requirements. Its capability to support higher rate to UEs would not be properly linked to support better QoS of traffic types. Differently from the reference one, while the proposed algorithm considers different requirement of mixed traffic, it could support better QoS with sacrifice of data rates.

In multi-RAT heterogeneous networks, since UEs can move, their mobility should be considered in allocating resources for better resource utilization. In D4.2, we introduce the access index to identify proper UEs to access Wi-Fi networks. By utilizing the access index for resource allocation, high-speed mobility UEs are less likely to access the Wi-Fi networks. Figure 17 shows the result of resource allocation in a case that high-speed mobility (60km/h) UEs become to be closely located to Wi-Fi APs at a certain time (*T1*). The mobility speed of UEs is selected in a set {3, 10, 60} [km/h] for six users and the average speed is 24.3 km/h. While the resource allocation algorithm is operated at every ms, it updates the allocation of the cellular spectrum every ms. However, proper UEs to access the Wi-Fi links are identified only every 10 ms (e.g., *T1* and *T2* in Figure 17).

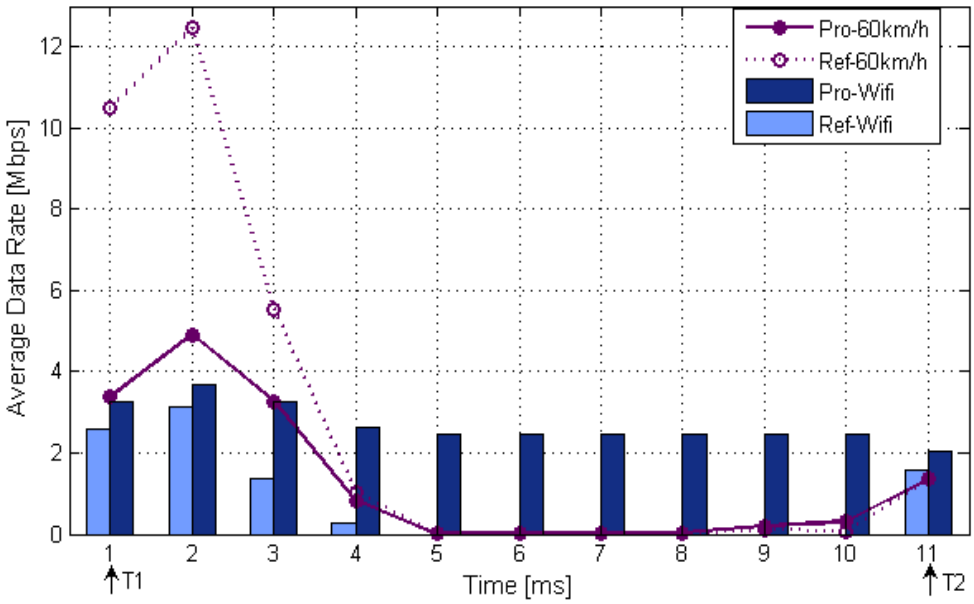


Figure 17: Comparison of performance of the proposed and reference scheme (with mobility) 1) 1) bar graph: average data rate achievable from the Wi-Fi network, 2) the curve graph: average data rate of high-speed mobility (60 km/h) UEs.

When the resource allocation algorithm is operated and proper UEs are selected for the Wi-Fi network at *T1*, for the reference scheme, the high-speed mobility (60 km/h) UEs are selected for the Wi-Fi networks. It is because UEs of stronger signal strength are likely to be chosen for the Wi-Fi networks. As the high-speed mobility UEs move, the Wi-Fi signal strength tends to get weaker and the UEs finally move beyond the Wi-Fi network coverage (at Time=5). At *T2*, the resource allocation algorithm is operated and proper UEs for the Wi-Fi network are chosen again. Then, the data rate of the high-speed mobility UEs becomes better. In the proposed scheme, user mobility is estimated based on the calculation of the past Wi-Fi signal strength fluctuation. Then, it is utilized as well as the signal strength to choose UEs for the Wi-Fi networks. Based on this, the low-speed mobility UEs are more likely to be selected for the Wi-Fi networks than high-speed mobility UEs. It leads to better utilization of the Wi-Fi link resource. In Figure 17, it is observed that the proposed scheme utilizes the Wi-Fi network better than the reference one over time.

In Figure 17, it can be observed that the proposed algorithm can better utilize the resource if it takes into account the estimated user mobility. In Figure 18, the performance of the proposed and the reference schemes is compared for different mobility level of UEs in terms of average UEs' utility and fairness index. For fairness, Jain's fairness index of each UE's utility value is calculated. The mobility speed of each UE is selected in a set {0, 3, 10, 30, 60} [km/h] and each UE's speed is not changed during simulation. The UE average speed is calculated as a mean of selected speeds of each UE.

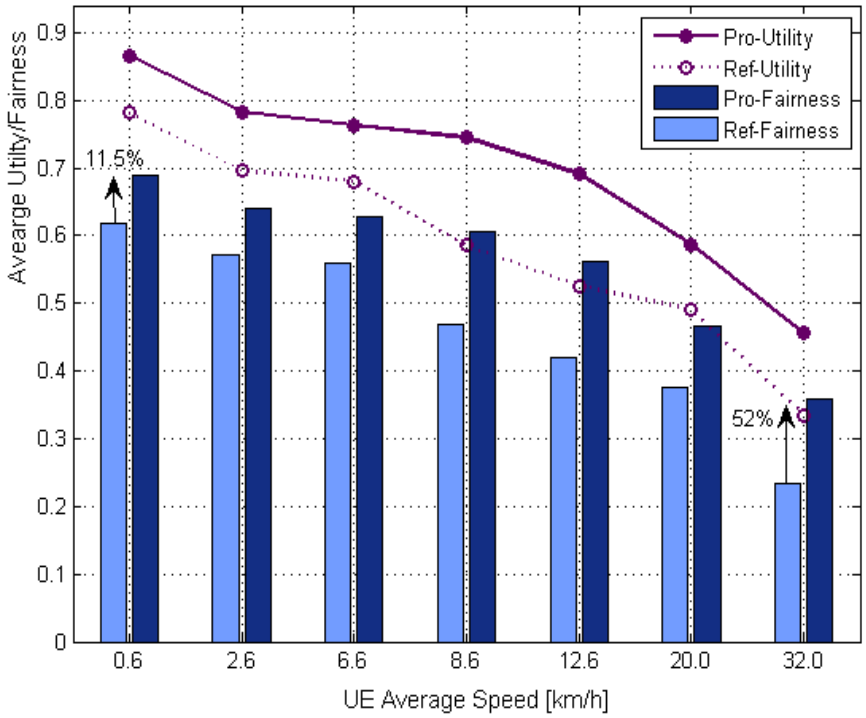


Figure 18: Comparison of performance of the proposed and reference scheme for different UE mobility speed: 1) 1) bar graph–average UEs' application utility values, 2) the curve graph- Fairness for different UEs' utility

From Figure 18, it can be observed that the UEs' average application utility value becomes lower as UE's mobility level increases. As explained in D4.2, high mobility UEs could move beyond the coverage of Wi-Fi networks shortly after being allocated Wi-Fi network resources to boost up their data rate. While the proposed algorithm utilizes the resource in multi-RATs better and considers heterogeneity of different traffic types, it could provide better QoS support for any cases of different UE's speed. Actually, the performance gap between the proposed and the reference one increases with the mobility speed level of UEs. While UEs move with low mobility of 0.6 km/h, the proposed scheme produces 11.5% improved utility value compared to the reference one. For UEs' high speed mobility of 32 km/h, 52% better performance can be achieved. In addition, in terms of fairness, the proposed algorithm outperforms rival algorithms.

In D4.2, the required signalling message sequence chart (MSC) was presented. Based on it, the signalling overhead was then evaluated for different user mobility rates. While there are 10 UEs, the channel quality of each carrier and the Wi-Fi links is assumed to be sent by all UEs at every ms. By having available user information, the proposed algorithm allocates resource to UEs every ms. Only selection of UEs to access the Wi-Fi networks is carried out every 10 ms. As expected, UEs selected for the Wi-Fi networks can be changed frequently as UEs move faster in Figure 19. Such change leads to increase the signalling overhead to reconfigure UEs to access the Wi-Fi networks (denoted as 'traffic steering'). However, since the amount of signalling overhead occurring every ms (due to CQI feedback and resource allocation for the cellular spectrum) is vast compared to the size of the signalling overhead for 'traffic steering', the overall signalling overhead does not increase significantly with the UEs' mobility level. Based on this, the signalling overhead performance is hardly influenced by the allocation algorithm strategy.

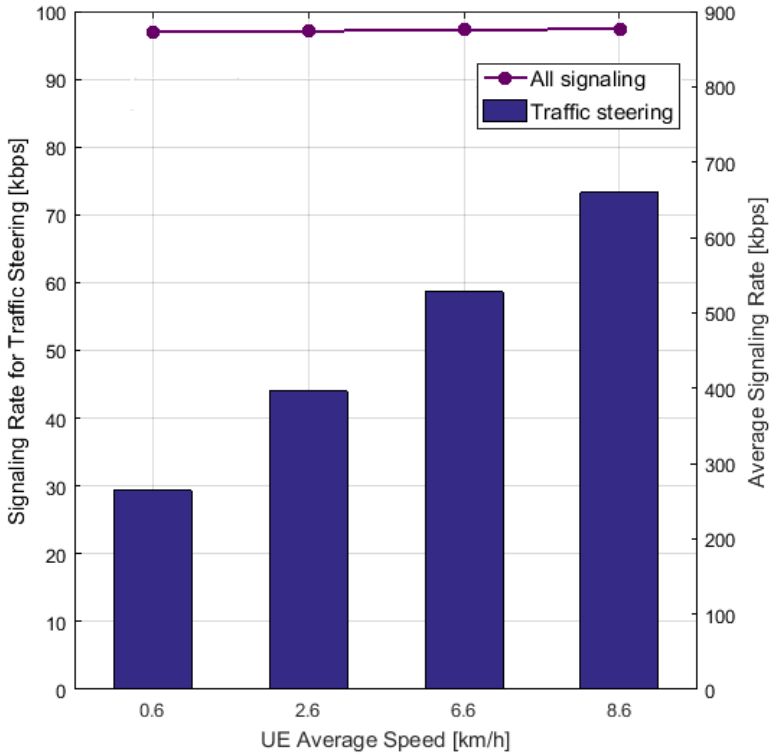


Figure 19: Signalling overhead for the proposed approach in HetNets with WLAN: 1) 1) bar graph–overall signalling overhead [kbps], 2) the curve graph- signalling overhead related to update UEs to access the Wi-Fi networks [kbps]

6.3 Summary

In this study, we have presented the radio resources allocation and aggregation algorithm to support different level of quality of service (QoS) of different traffic types in the cellular network where the WiFi network coexists. While the utility function approach is utilized to model different traffic types, the access index integrating the estimated user mobility and location information is calculated to identify more proper users for the specific network. Our simulation results show the proposed algorithm supports better QoS of different traffic types in terms of the application utility values with improved fairness. In addition, by estimating user mobility and using it to allocate resource, it is shown that the proposed one can also support QoS of high mobility users.

7 Management of traffic offload and channel selection in LAA systems

In deliverable D4.2 [11], we have presented an algorithm based on machine learning tools able to efficiently exploit the available licensed and unlicensed spectrum resources at Licensed-Assisted Access (LAA) small cells (see Figure 20). This algorithm is designed to support eMBB traffic services characterized by diverse Quality of Service (QoS) requirements, and our simulation results in D4.2 have highlighted that it provides 25% higher probability of successful access to the LAA channel, and improve the performance of the end users, especially those located at the cell edge. The message sequence chart of the proposed algorithm is described in Figure 21.

To complement and complete the work presented in D4.2, in this section we present and report evaluation results of mechanisms proposed for reducing the interference generated by neighbouring small cells on the unlicensed band through the usage of filter bank multicarrier (FBMC) modulation [9]. With FBMC the user data is parallelized and transmitted through a bank of modulated filters.

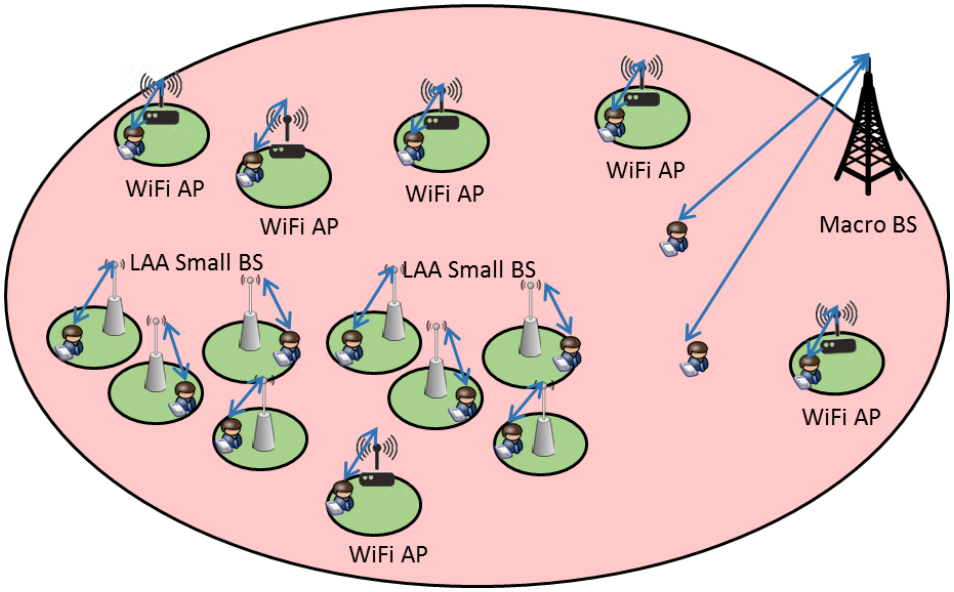


Figure 20: The heterogeneous network under investigation.

At the best of our knowledge, there is a lack of optimization framework for managing LAA spectrum in the case of FBMC transmission schemes. In addition, it is worth to underline that this study differs from the one presented in Chapter 9, where machine-learning tools are used to analyses and classifies the user traffic, while here reinforcement leaning enables the decision making mechanism.

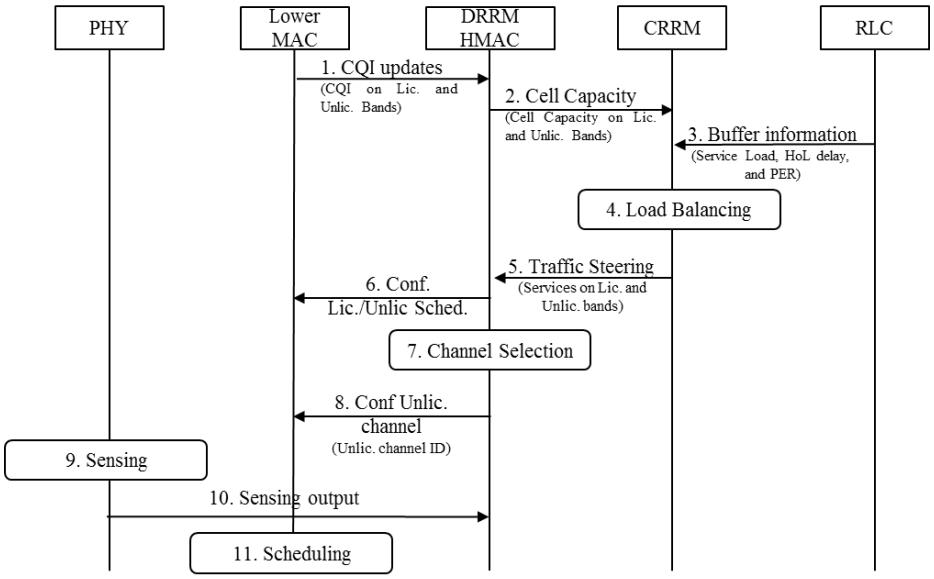


Figure 21: Message sequence chart for the efficient LAA operation in small cells, based on reinforcement learning.

Besides the relaxed synchronization constraints, which is a suitable feature for LAA communications, FBMC promises higher spectral efficiency and lower inter-cell interference with respect to the classic cyclic prefix orthogonal frequency division multiplexing (CP-OFDM). Therefore, our objective is to improve the user performance by efficiently combining FBMC and machine-learning technologies.

7.1 Comparison of FBMC and CP-OFDM

The major drawback of OFDM relies in its spectrum shape whose size is equal to the duration of the fast Fourier transform (FFT). Each rectangularly-shaped subcarrier in the time domain results in a sinc(*f*) function in the frequency domain, which when summed leads to high spectrum side-lobes. These side-lobes results in out-of band emissions which generate inter-cell interference when neighbouring small cells use adjacent channels.

	Max. power in 20 MHz	Max. power in 1RB (180 kHz)
OFDM	-37 dBc / 20 MHz	-59 dBc / 180 kHz
FBMC K = 2	-44 dBc/ 20 MHZ	-64.4 dBc/ 180 kHz
FBMC K = 3	none (< -90 dBc)	none
FBMC K = 4	none (< -100 dBc)	none

Table 9: OFDM and FBMC out-of band emissions.

Table 9 indicates the spectral density relative to the maximum spectral power density of the transmitted signal in the case of the OFDM and FBCM, for different values of the overlapping ratio (*K*), which is related to the duration of the impulse response of the prototype filter of the FBMC. The effect of these side-lobes is also described in Figure 22.

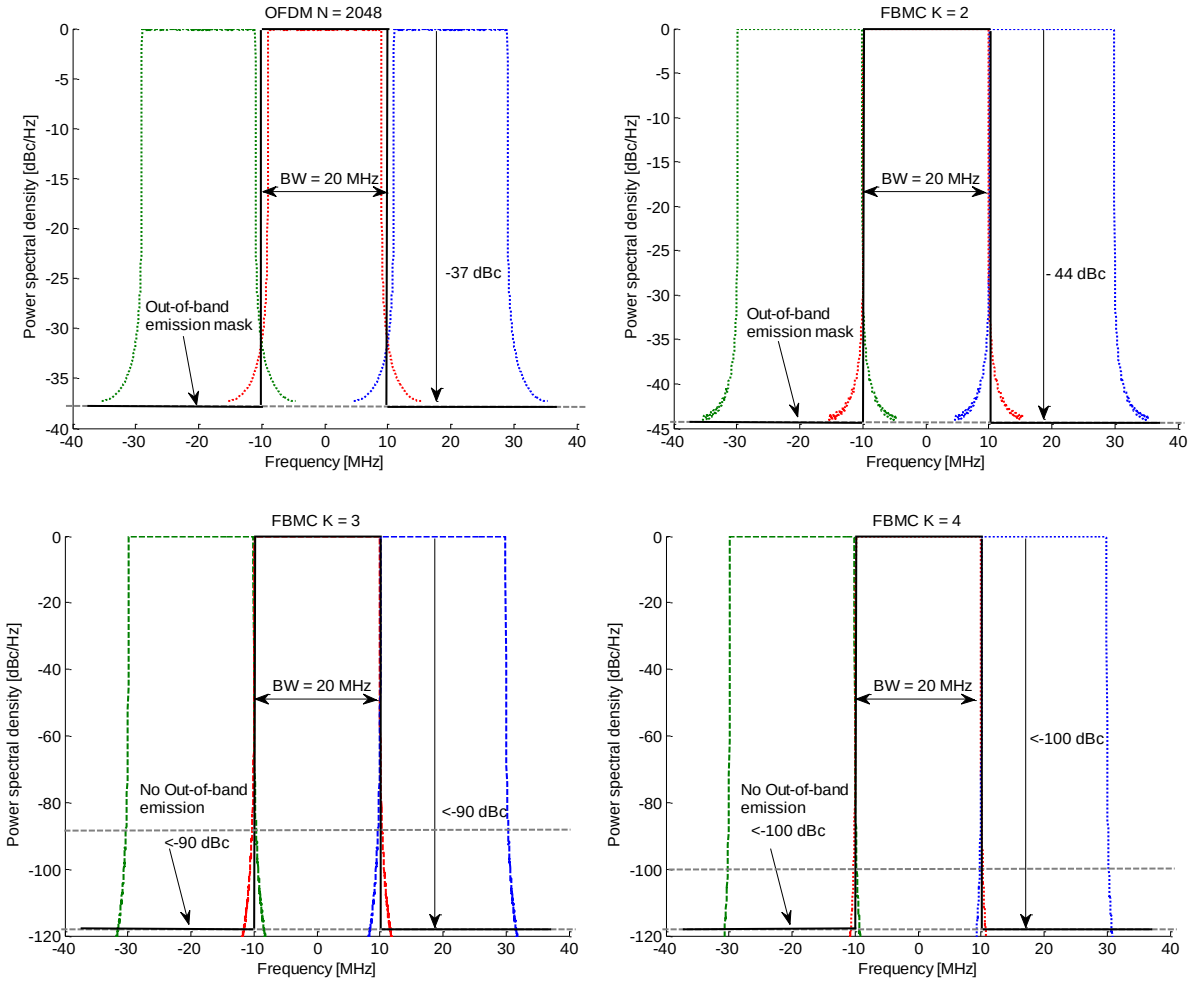


Figure 22: Spectral power density of a 20 MHz transmitted signal in the case of the OFDM and FBMC, with K=2, 3, and 4.

Besides, the sinc filter used with the OFDM transmissions reduces the system spectral efficiency with respect to FBMC where a better localized filter can be used. Moreover, with OFDM, at the receiver, perfect signal recovery is possible under ideal channel conditions thanks to the orthogonality of the subchannel filters. Nevertheless, under more realistic multipath channels a data rate loss is induced by the mandatory introduction of a CP, longer than the impulse response of the channel. The effect of the CP is a spectrum efficiency loss, which can be avoided when using FBMC technology [9].

Overall, if we consider that in a 20 MHz transmission, a CP size of 144 samples leads to 18 MHz of useful signal while our study show that 19.8 MHz can be exploited when using a FBMC modulation.

7.2 Mapping to scenarios

This study focuses on the eMBB use case where a massive deployment of small cells is put in place to provide a uniform broadband experience to the users demanding high data rate and limited latency for the provisioning of applications such as high resolution multimedia streaming, gaming, video calling, and cloud services. Small cells can operate in both co-channel and dedicated channel deployment; moreover, licensed/lightly licensed and unlicensed spectrum could also be used to boost the capacity, coverage and balance the traffic and meet the meet the 5G requirements.

The KPIs of primary importance, which characterize this use-case, are primarily the per-user throughput, and the E2E latency.

We consider a situation where multiple small cells are managed by a single network operator and they need to convey a mix of eMBB and experience varying interference that could degrade the QoS.

To deal with these challenges, the RRM selects a band from a combination of licensed and unlicensed spectrum available, depending on the context to meet some specified QoS requirement.

7.3 Simulation assumptions

To enable the global assessment of our solution, we investigate downlink communications in dense heterogeneous networks with UEs and small cells having LAA capabilities. More specifically, small cells can be deployed either in co-channel mode, using the same band as the macrocells (e.g., 2 GHz), or in dedicated channel mode (at 3.5 GHz). In both cases, the available licensed bandwidth is assumed to be 10 MHz. In addition, small cells can operate in the 5 GHz unlicensed band where eight channels of 20 MHz are available for shared access. All the small cells are assumed to be time-synchronized and unaware of the unlicensed band occupancy statistics (probability of being vacant, transition probability from vacant to busy, etc.). Table 10 shows the main simulation assumptions used to evaluate the proposed solution. Other system parameters are in line with the 3GPP specifications [48]. Since the focus is on small cell deployments with LAA capability, we only report on the performance of the small cell users.

Parameters	Values	Parameters	Values
ISD	500 m	# of SCs	4 per Macro sector
# of Macro eNB	19 three-sectorized	# of UEs	30 (2/3 of the UEs are randomly located in the hot-spots)
Available spectrum	8 x 20 MHz @ 5 GHz 1 x 10 MHz @ 3.5 GHz	UE Traffic type	NRTV (rate 4 Mbps, latency 100 ms latency) FTP (burst size 5 Mbit, average reading time 0.1 s, and latency 300 ms) CBR (rate 1Mbps)
LAA channel busy probability	(0.1; 0.2; 0.2; 0.8; 0.6; 0.4; 0.1; 0.3)		

Table 10: Simulation assumptions and parameters of proposed algorithm

We consider four traffic scenarios where we evaluate the performance of our channel selection scheme based on multi-armed bandit (MAB), when using a classic OFDM transmitter solution and a more advanced FBMC approach.

7.4 Performance evaluation

In all these scenarios, we can observe that the FBMC transmission scheme leads to notable throughput gain with respect to the OFDM solution. In the following plots, the dashed line and the full line respectively represent the user performance with the OFDM-based MAB and the FBMC solution. In the first case, all the users receive bursty traffic model by using the File Transfer Protocol (FTP). Figure 23 shows the empirical cumulative distribution function of the user throughput in this first traffic scenario. The dashed line and the full line respectively represent the user performance with the OFDM-based MAB and the FBMC solution. Overall, a 5.9% throughput gain can be observed.

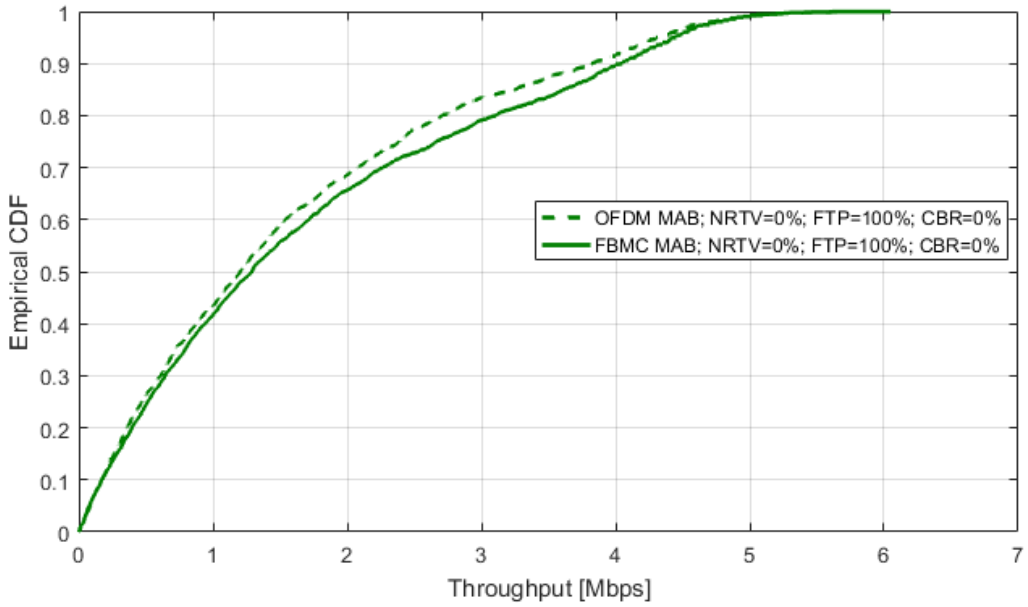


Figure 23: Empirical CDF of the user throughput in the first traffic scenario, when using the OFDM and the FBMC transmission scheme.

Figure 24 shows the empirical cumulative distribution function of the user throughput in second scenario. In this case, the user traffic user mix is composed by 30% of Near Real Time Video (NRTV) streaming, 50% of FTP, and 20% of Constant Bit Rate (CBR) traffic. Overall, a 6.3% throughput gain can be observed.

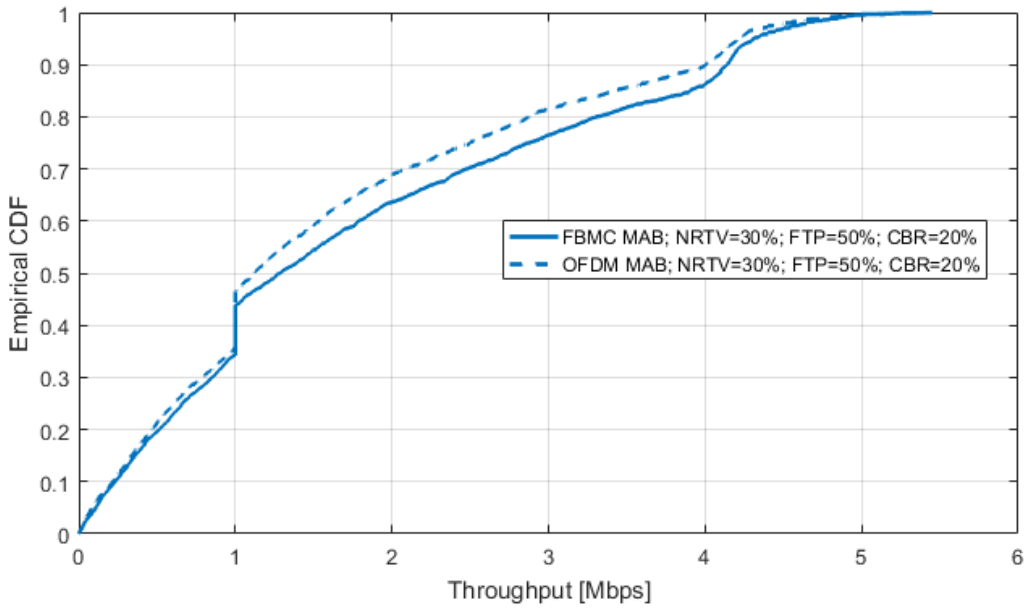


Figure 24: Empirical CDF of the user throughput in the second traffic scenario, when using the OFDM and the FBMC transmission scheme.

Figure 25 shows the empirical cumulative distribution function of the user throughput in third scenario. In this case, the user traffic user mix is composed by 50% of NRTV streaming and 50% of FTP. Overall, an 8.5% throughput gain can be observed.

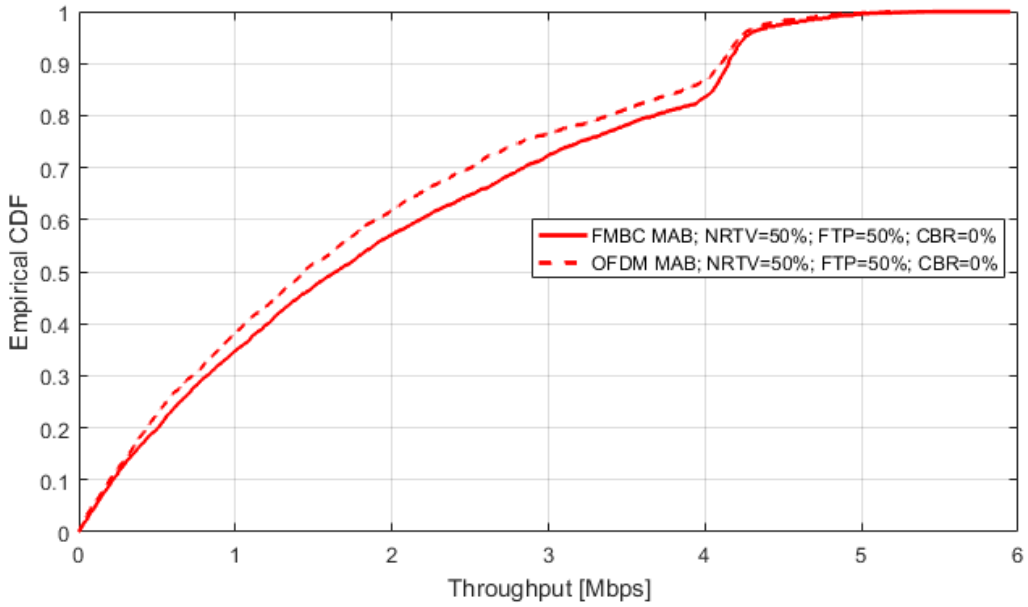


Figure 25: Empirical CDF of the user throughput in the third traffic scenario, when using the OFDM and the FBMC transmission scheme.

Figure 26 shows the empirical cumulative distribution function of the user throughput in third scenario. In this case, the user traffic user mix is composed by 50% of NRTV streaming, 30% of FTP, and 20% of CBR traffic. Overall, an 8.6% throughput gain can be observed.

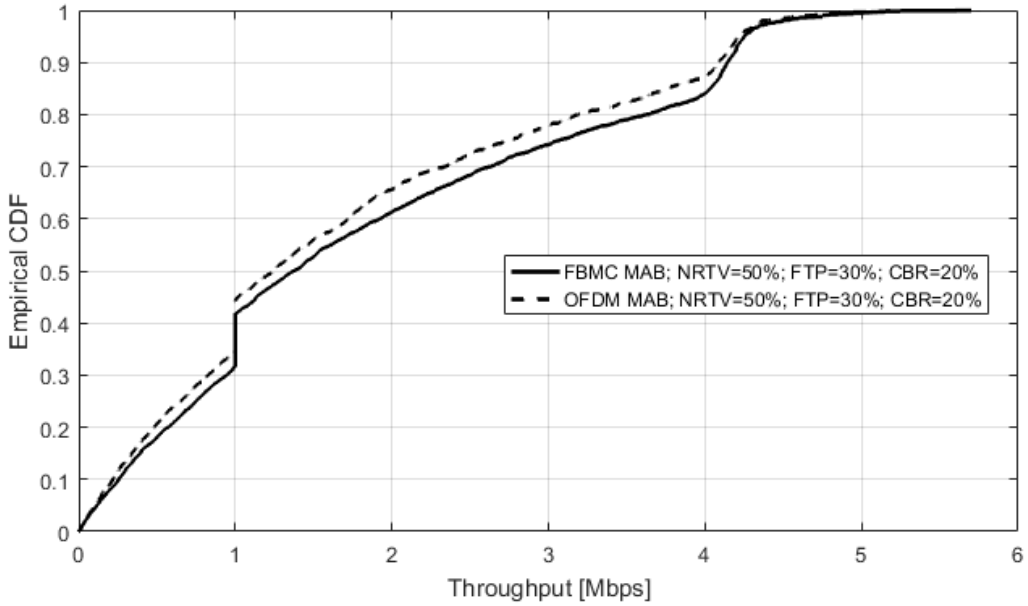


Figure 26: Empirical CDF of the user throughput in the fourth traffic scenario, when using the OFDM and the FBMC transmission scheme.

7.5 Summary

In this study, we have presented an algorithm based on machine learning tools able to efficiently exploit the available licensed and unlicensed spectrum resources at Licensed-Assisted Access (LAA) small cells. This solution is complemented with the integration of an FBMC transmission scheme that enables the reduction of interference generated by neighbouring small cells on the unlicensed band. Our results show that from the one hand reinforcement learning improves the spectral efficiency by learning the statistics of the usage of the unlicensed channel and accessing those channels where the expected interference is limited. Finally, it is worth noting that FBMC reduces the inter-cell interference when neighbouring small cells use adjacent channels.

8 Resource allocation for traffic offload in LTE-U

8.1 Introduction

For mobile operators, efficient spectrum utilization is essential. Therefore, mobile operators consider integrating the WiFi within their infrastructure as an additional supplementary downlink for offloading more data traffic [35]. The utilization of the WiFi unlicensed band offers a significant extension for mobile operators' resources to meet mobile traffic requirements. With a significant amount of unlicensed spectrum globally available in the 5 GHz band [36], the mobile operators and vendors are looking to use unlicensed spectrum to augment the capacity of licensed frequency carriers. In a 3GPP RAN plenary standards meeting in December 2013, the proponents, formally proposed "LTE-Unlicensed" (LTE-U) to utilize unlicensed spectrum to carry data traffic for mobile services with initial focus on the 5725-5850 MHz band for this use [36], [37].

The Federal Communications Commission (FCC) pioneered unused TV spectrum for "Super WiFi" hotspots with highly-efficient characteristics, many of which are owing to the comparatively low carrier frequencies of TV bands [38]-[40]. LTE-U stretches out LTE to the unlicensed spectrum and incorporate the unlicensed spectrum with the licensed spectrum based on existing Carrier aggregation technology continuous seamless data flow between licensed and unlicensed spectrum via a single Evolved Packet Core (EPC) network. For operators, the presence of LTE-U means synchronized integrated network management, same authentication procedures, more efficient resource utilization, all of which leading towards lower operational costs. For mobile users, LTE-U implies improved quality of experience, i.e., more data rates, uninterrupted service provisioning between licensed and unlicensed bands, ubiquitous mobility and improved reliability. However, it is observed that the coexistence of LTE-U and WiFi in the same frequency bands causes a significant degradation on the system performance. Currently, Wi-Fi systems adopt a contention based medium access control (MAC) protocol with random backoff mechanism [41]. If left unrestrained, unlicensed LTE transmissions can actively and aggressively occupy the channel (i.e. 5 GHz) and make the medium busy most of the time. This will generate continuous interference to Wi-Fi systems, resulting in "Busy Time Period" (BTP) of Wi-Fi nodes. This will not only degrade the Wi-Fi devices throughput, but also overall throughput of the system.

To overcome the above constraint, we propose a novel WiFi-Lic solution which can sense and access the spectrum hole of the LTE band during channel busy period. This solution not only solves the co-existence issues between WiFi and LTE-U, but also decreases the WiFi BTP of the channel to zero and efficiently uses the underutilized spectrum of LTE [42]. Lastly, WiFi-Lic also complements LTE-U solution to further enhance the overall system throughput. To the best of the authors' knowledge no work has been done so far on this topic.

8.1.1 LTE vs. WiFi

WiFi performs its transmission over the unlicensed band using the carrier-sense multiple-access with collision avoidance (CSMA/CA) feature, which is commonly referred to as listen-before-talk (LBT). This scheme enables the WiFi to perform carrier sensing prior to any potential transmissions. Once a channel is determined to be vacant after certain duration of sensing intervals, the WiFi transmits for certain time before retrieving to silence mode. If the medium determined to be busy, the WiFi backs off for certain period before resuming channel assessment. When the channel becomes idle, WiFi starts transmitting and the same procedure will be repeated without any guarantees for the quality of service (QoS). The simple WiFi transmission procedure reflects in temporally occupation and low utilization of unlicensed band.

The LTE transmission scheme supports various carrier bandwidths of 1.4, 3, 5, 10, 15, and 20 MHz and employ carrier aggregation technology that extends the LTE downlink bandwidth further by aggregating two or more component carriers to support high data rate transmissions [43]. LTE uses OFDMA signalling for downlink and the SC-FDMA signalling for uplink. Data and control signals are transferred

on different channels with lower spectral efficiency in the physical uplink shared channel (PUSCH) that carries user data. Moreover, LTE has better QoS management and control functions. Currently LTE supports eight different QoS classes with different performance requirements.

8.1.2 LTE-U

The principal objective of LTE-U is to incorporate carriers between licensed and unlicensed band. There are two types of carriers, those are: principal component carrier (PCC) in licensed band and secondary component carrier (SCC) in unlicensed band. Any user should be configured with one PCC and several SCCs. PCC is responsible for control-plane ad Layer 1 control signalling. On the other hand, user-plane data is carried through either by PCC or SCC. A technique called Listen-Before-Talk (LBT) [44] is used to incorporate WiFi systems with LTE systems. Recently, researchers have studied WiFi and LTE coexistence considering also TV white space. Several studies [43] propose CSMA/sensing-based modifications in LTE with features like Listen-before-Talk. In other studies, to enable WiFi/LTE coexistence, solutions like blank LTE subframes/LTE muting [43], carrier sensing adaptive transmission and interference-aware power control in LTE. All of the above proposals explain how to access WiFi unlicensed band or coexistence with it; however, so far no work has been proposed towards accessing the WiFi licensed band.

8.2 WiFi in licensed band (WiFi-LIC)

This section explains the WiFi-Lic technology in coexistence with LTE-U. All the scenarios and assumptions related to LTE-U [3] are also valid for WiFi-Lic, with the addition of the new concept on sensing the LTE licensed band during a busy period.

8.2.1 System design

Figure 27 shows the system design of WiFi Lic technology. It consists of operator-deployed WiFi and LTE small cell (SC) access points (AP) operating on the 5 GHz unlicensed and 2.6 GHz licensed band, respectively. Moreover, there is an LTE and WiFi device, which consist of three SDR radios (i.e. LTE, LTE-U and WiFi), with the WiFi user being complemented by cognitive radio functionality. Among many possible LTE-U deployment options, this section focuses on supplemental downlink (SDL) deployment in unlicensed band. In a SDL mode, the unlicensed spectrum is used to provide an additional link to be aggregated with the licensed downlink in order improve the overall downloading rate to the mobile user. For this work, we only consider deployments which target the regions without LBT requirements, such as the USA, but it will equally work for the region with LBT requirements such as Europe with slight modification in air interface of the LTE interface.

During the transmission interval, LTE users can access the licensed and unlicensed component carriers (PCC and SCC), simultaneously. However, the system configuration information including the SCC access on is only exchanged using the PCC carriers. The initial attach, authentications and security are performed on the PCC. All other applications sensitive to latency or jitter can be supported on PCC which have more predictable availability and QoS.

The unlicensed bands of SCC are only used for data transmissions due to the fact that data transmissions can be buffered when there are no free available channels. In contrast, control signals cannot be delayed due to the physical resource block (PRB) allocations and therefore have to be sent using the PCC. In our proposed scheme, once a WiFi channel becomes available, the small cell exchanges information with the LTE end user over the PCC. Upon reception, the LTE end user switches on its LTE-U interface to the assigned SDL channel. On the other hand, when WiFi user tries to access the channel (assuming all channels are currently occupied by LTE-U), the medium is sensed and transmission is postponed due to the detection of co-channel LTE-U. This will degrade WiFi users' throughput and increase the BTP. The main reason for this disproportionate drop in the WiFi throughput is due to the fact that LTE does not sense other transmissions before transmitting. In

contrast, Wi-Fi is designed to coexist with other networks as it senses the channel before any transmission.

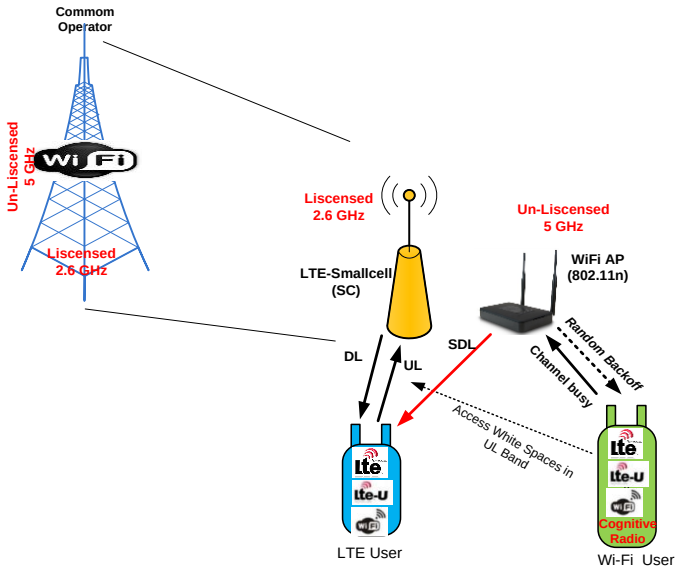


Figure 27: System model of WiFi -licensed spectrum

Therefore, in our novel solution as soon as WiFi user senses the medium is busy, it turns the cognitive radio to access the white spaces in the LTE band as shown in Figure 27. This solution will increase the WiFi throughput, as well as reducing the BTP to almost zero.

8.2.2 Channel selection for LTE-U

Before data transmission, LTE small cells scan the unlicensed band for the identification of the cleanest channels and inform the users of data transmission via cross-carrier scheduling. The measurements are performed at both the initial power-up stage and later periodically at the SDL operation stage. If interference is found in the operating channel and there is another channel available, the SDL transmission will be switched to the new channel using the LTE 3GPP Release 11 procedure. But there is the possibility that no clean channel can be found, and in such cases, LTE-U can share the channel with WiFi AP or another LTE-U system using Carrier-Sensing Adaptive Transmission, for which we assume that there is always a clean channel available for LTE-U.

8.2.3 Channel selection for WiFi-Lic

Our WiFi-Lic novel channel access mechanism is depicted in Figure 28. To better explain the operation of WiFi-Lic, we have divided it into four steps for sake of simplicity.

Step 1: During this step, LTE-U user accesses 5 GHz unlicensed channels (for example 15 channel) using SDL. SDL data can vary from channel to channel according to traffic demand and have a time duration of 1ms. Moreover, we assume that there is always a clean channel available for LTE-U in 5 GHz (no WiFi transmission).

Step 2: In this step, WiFi user tries to access one of the 15 channels which are currently in use by LTE-U. Hence, a WiFi user starts sensing all of the 15 channels and finds that none of them are free. BTP will automatically turn “on” the cognitive radio interface of WiFi for the whole duration of the SDL data (1 ms).

Step 3: During cognitive radio activation, WiFi-Lic user attempts to access the LTE channel. The WiFi-Lic conducts channel assessment only at preassigned periodic time intervals compatible to the LTE PRBs; we refer to these intervals as “WiFi-Lic access opportunities”. We denote the WiFi-Lic predefined period time as T_{sensing} (μs). Once the WiFi-Lic identifies an access opportunity, transmission starts for a fixed

transmission duration $T_{\text{WiFi\_Lic}_{Tx}}$ (μs); otherwise, the WiFi-Lic user starts sensing again, unless the channel is idle.

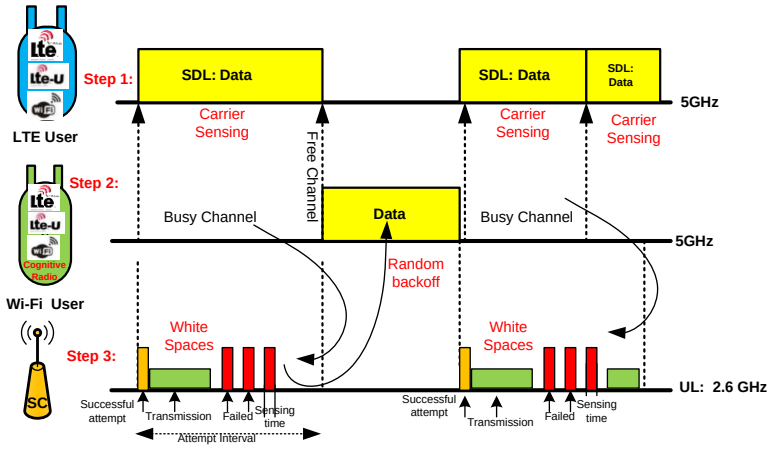


Figure 28: Channel access mechanism of WiFi-Lic

Step 4: Step 2 and step 3 have occurred for the duration of 1ms, which we called "attempt interval $\left(T_{\text{WiFi-Lic}_{\text{attempt}}}\right)$". As soon as attempt interval finishes, WiFi-Lic automatically switches off cognitive radio interface and switches on its WiFi interface to sense the carrier again. The WiFi performs the normal channel assessment and starts transmitting if a channel becomes available or backs off by a random number of time slots and then resumes transmission attempts. Otherwise, it goes to step 3.

We identify three key performance metrics to model the WiFi-Lic proposed channel access mechanism are as follows:

Attempt Interval $\left(T_{\text{WiFi-Lic}_{\text{attempt}}}\right)$ is the period of access opportunities in LTE band. $T_{\text{WiFi-Lic}_{\text{attempt}}}$ is used to control how frequent WiFi-Lic accesses licensed band and duration of 1ms.

Transmission duration $\left(T_{\text{WiFi-Lic}_{\text{Tx}}}\right)$ is the maximum duration that a WiFi-Lic can occupy the channel during transmission period. At the end of $T_{\text{WiFi-Lic}_{\text{Tx}}}$, the WiFi-Lic user has to move to the next available channel assuming that the licensed user returns to access this channel. So licensed users do not starve, as they are the primary users of LTE band.

Channel Sensing Duration $\left(T_{\text{sensing}}\right)$ is the predefined time interval during which the WiFi-Lic examines a selected licensed channel for a potential transmission attempt.

8.2.4 Theoretical analysis

Please refer to Appendix A.1 for a detailed theoretical analysis.

8.2.5 Simulation assumptions

To check the validity of our proposed approach, we have performed computer simulations. We considered the system design as described in section 8.2.1, coupled with the following parameters as shown in Table 11.

Parameter	Value	Parameter	Value
WiFi Parameters			
WiFi Type	802.11n	MAC	DCF
Payload size	1470 bytes	MPDUs	4

Parameter	Value	Parameter	Value
WiFi Parameters			
MAC+PHY hdr	24+06 bytes	ACK	16 bytes
CCA Threshold	-62 dBm	Header rate	6.5 Mbps
Channel rate	(13, 26, 39, 52, 78, 104, 117, 130) Mbps		
Required SINR	(5,7,9,13,17,20,22,23) dB		
ACK frame rate	Max(6.5, 13, 26) Mbps <= Ch_Rate		
Traffic Model	3GPP FTP Traffic model	Slot time	9 μs
Bandwidth	20 MHz	DL/UL Tx power	23 dBm
LTE Parameters			
LTE	TDD	Bandwidth	20 MHz
Scheduling	Proportional Fairness	DL Tx power	23 dBm
TTI	1ms	Control overhea	20%
SINR w.r.t CQI	(1.95, 4,6,8,11,14,17,19,21,23,25,27,29)		

Table 11: Simulation parameters

In our simulation, we consider a SDL mode in two-tier cellular network similar to [36]. The network consists of macrocells and outdoor picocells that share a bandwidth of 20 MHz in the licensed band. The inter-site distance between macrocells is set as 500 m. In each macrocell domain, we employ 5 picocells and 10 operator-deployed WiFi APs. Then, we examine network model scenarios, i.e. cellular/WiFi internetworking and conventional HetNets (only with licensed access). In the cellular/WiFi scenario, the WiFi replaces internetworking picocells. In HetNets, there is no WiFi and the network only operates in licensed band.

8.3 Performance evaluation

Figure 29 shows the results in terms of average user throughput of licensed, unlicensed and WiFi-Lic for LTE small cells. The actual user throughput of operator deployed scenario in case 1 is 4.526 Mbps. The other throughput values are all normalized according to case 1 baseline value.

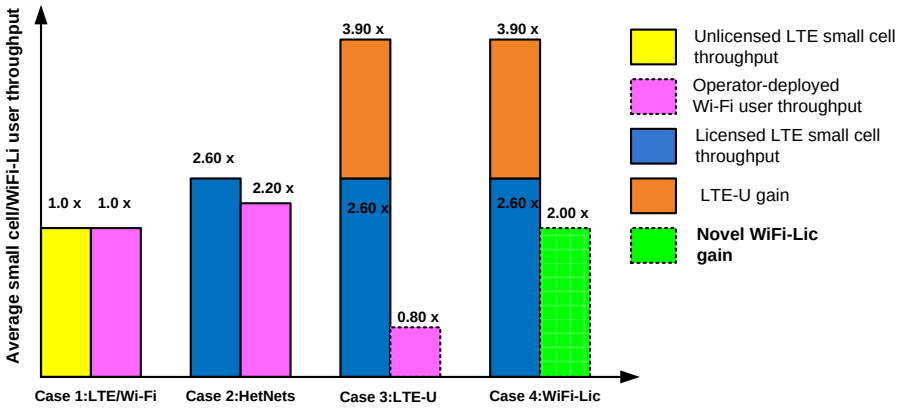


Figure 29: Results in terms of average user throughput of licensed, unlicensed and WiFi-Lic for LTE small cells.

Figure 29 shows the performance of our scheme against the state of the art schemes and it is divided into four cases. Case 1: the throughput of both networks (LTE & Wi-Fi) are the same and this is due to

the fact that both networks are using unlicensed band, equal number of users and access nodes densities. Case 2: the throughput for HetNets is enhanced almost double than case 1. This is due to the fact that users from small cell access the licensed spectrum of LTE using centralized MAC, which is more spectrally efficient than DCF. Eventually the WiFi user throughput is increased by a factor of 2 since LTE small cells use orthogonal spectrum with WiFi, which causes no interference from LTE to WiFi users. Case 3: instead of deploying HetNets small cell in licensed spectrum as in case 2 (blue bar), the introduction of LTE-U (small cell in 5 GHz unlicensed band) proves to be a better throughput enhancement solution (brown bar) for future 5G systems. However, this solution degraded operator deployed WiFi throughput severely (0.80x) as shown in case 3, without any coexistence mechanism. Case 4: When we apply our novel solution, WiFi-Lic throughput increases to 2.00x. This is due to the fact that during busy times WiFi-Lic will access the white space in LTE band, according to the procedure explained in section IV. Moreover, it is important to note that we need to do down conversion of WiFi 5 GHz carrier frequency to LTE 2.6 GHz. WiFi may lose some part of bandwidth but that loss has not much impact, for best effort traffic.

Figure 30 shows the performance of WiFi-Lic during a busy period (LTE access the 5 GHz and no channel is available). WiFi-Lic performance is much better than the normal WiFi mode during busy time. Normal WiFi waiting period varies randomly during channel busy time, as compared to WiFi-Lic solution which access the LTE white spaces during busy time and spends much less time in busy mode, which directly translates into higher overall throughput of the system.

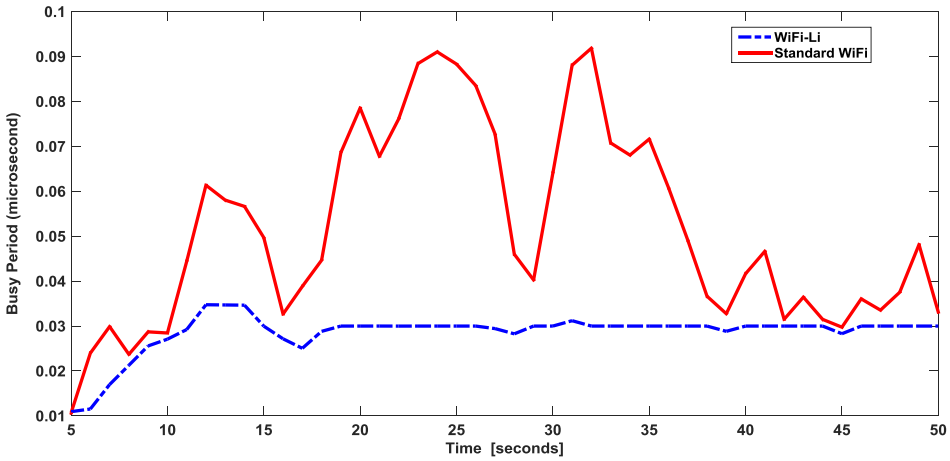


Figure 30: WiFi Lic Busy channel performance

9 RAT, spectrum, and channel selection based on hierarchical machine learning

9.1 Overall algorithm description

In SPEED-5G D4.2 Chapter 3, there was an initial description and documentation of a method of RAT, spectrum, and channel selection based on hierarchical machine learning.

In the following, *incumbent access* users are users with absolute priority in spectrum resources. These users will be protected from harmful interference from *Priority Access Licences* (PAL) and *General Authorized Access* (GAA) users. PAL users are in the second priority level and can utilize certain channels up to three years by acquiring the necessary licences. Finally, GAA users do not have to obtain an individual spectrum licence and use spectrum resources opportunistically (when incumbent and PAL users are not transmitting in the area).

The proposed solution of channel selection based on the 3.5 GHz SAS model included the following entities:

- GAA Users could use learning mechanism to get:
    - Channel usage of PAL Users
    - Channel usage of neighbouring GAA Users

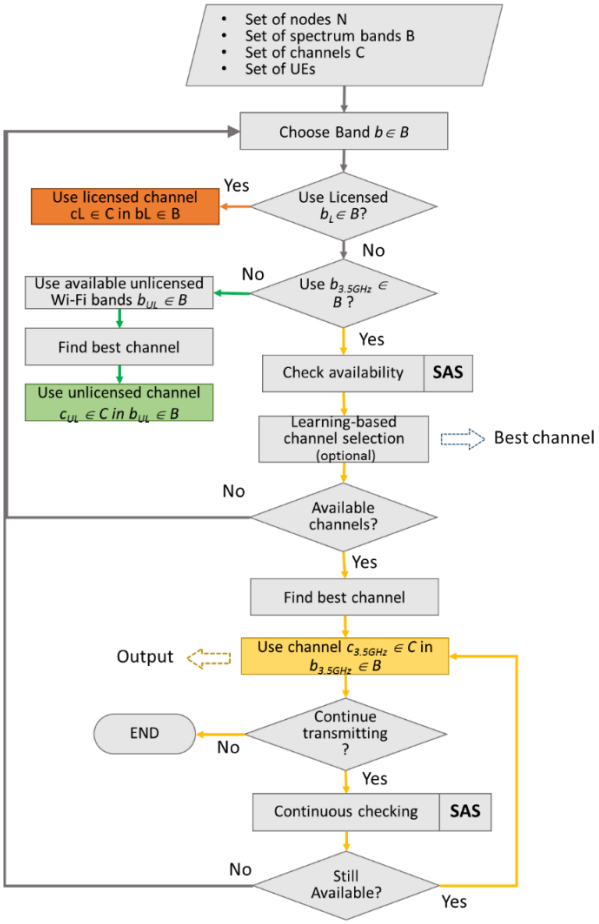


Figure 31: Flowchart of algorithm with learning capabilities as an option (source: SPEED-5G D4.2, Chapter 3)

Figure 31 illustrates a flowchart of the proposed algorithm and the procedure of learning and predicting the conditions of the channels from the viewpoint of the GAA users. As such, parameters such as band/channel availability, duration of channel usage, recurrence of channel usage (for example, every day, every week etc.) and location are needed in order to be able to select a specific

cell/channel faster and in a more reliable manner based on prediction for the utilization by incumbent and PAL users.

The flowchart illustrates the steps of the algorithmic solution for utilizing the 3.5 GHz band in our system which focuses on the SAS three-tier lever mechanism that addresses how the different licensed users transmit at the specific band. Specifically, Figure 31 highlights how the proposed algorithm will work with learning capabilities as an option when employed jointly with the SAS system. By involving the block of learning algorithm the system will be able to acquire information about the conditions (e.g., throughput, SINR, etc.) of the channels at a future time, which the SAS system is not capable of providing. Also, as illustrated in the Figure 31, the "Learning-based channel selection" box introduces a learning mechanism which will be able to provide information and predictions about the characteristics of the channel and more specifically the quality of each channel. In general, by applying statistical learning techniques will be easier to automatically identify patterns in data that can be used to make more accurate predictions. At this part of the main algorithm, it has to be determined whether a channel is good or bad. In machine-learning terms, a classification task will be introduced to categorize data points and for example, based on the maximum transmitting throughput, the channel that reaches high throughput should be classified as the best channel and the one to be chosen. Acquired data can be split into two parts: the training data which will be used to train the model of the algorithm, and the test data which will be used to test the model's performance on that new data that have never been applied to the algorithm before.

9.2 Simulation assumptions

For the evaluation of the proposed algorithm, system-level simulations have been conducted. The implementation of our suggested solution was performed under a proprietary system-level simulation tool which is developed in Java. The simulator takes into account various parameters such as traffic level, available infrastructure elements, available channels and evaluates the various test cases. The calibration state of the proprietary simulator has been checked against the reference results of the 3GPP LTE calibration campaign [36.814]. As a result, the Cumulative Distribution Function (CDF) of coupling loss and downlink SINR have been checked in order to calibrate the tool with leading operators and vendors such as Nokia, Ericsson, DoCoMo, Huawei, Telecom Italia, etc. The configuration is fully customizable so as to include various types of cells (i.e., macro and small cells). Specifically, it is possible to customize the following: the size of the simulation domain; the area type (e.g., dense urban etc.); the number and position of macro base stations and their inter-site distances (ISDs); the number and position of small cells per macro base station; the number and position of end-user devices in the domain; the mobility of the end-user devices; the number of available channels and many more. In addition, the path loss models for macro cells at 2 GHz band is set to L = 128.1 + 37.6 log10(R), R in km, and for small cells is set to L= 140.7+36.7 log10(R), R in km.

In order to evaluate the proposed algorithm in a heterogeneous system, the proprietary simulator shall be utilized with all the above parameters implemented. Therefore different scenarios and test cases are introduced to the simulation environment and will be further described and analysed in detail. In general, test cases with variable traffic loads based on the frequency of the arrival requests will be used in the system. Additionally we will experiment with various traffic mix situations, specifically the percentage of the different licensed users in order to obtain an even broader knowledge of the algorithm capabilities, and overall performance for the specific network environment that will be introduced to the simulator.

Initially, our algorithm is deployed to the system-level simulator that utilises a number of macro and small cells at 2 GHz and 3.5 GHz bands respectively with a number of different channels. Specifically, the parameters imported to the simulator are; 7 macro base stations (BS) each with three cells and also 30 small cells per BS, giving us a number of 210 small cells and a total of 231 cells throughout the network. In addition we have utilized 15 channels at 10 MHz bandwidth for every cell and 10 of them can be assigned to the PAL users which is the 2/3 of the total channels. The incumbent users are able to use any of the 15 channels of the system and of course are able to transmit to all of them at the

same time if needed. At last the GAA users are able to use all of the available spectrum but are permitted into using only 80 MHz of the 150 MHz at each time. Thus only 8 channels can be assigned to the GAA users of the 15 total that is available but they can choose which of them will be utilized dynamically. The total traffic is shared among the macro BSs and small cells and ranges between 21-22% for the macro cells and 78-79% for the small cells in our test cases.

In order to assess our algorithm, a number of test cases and scenarios have been created. Table 11 presents the simulation parameters of the system level simulation with the different loads of the traffic expressed with a variety of inter-arrival requests from users ranging from 5.5 up to 10 seconds. Also, Table 13 introduces four different scenarios for the appearance of the three-tier users to the system. Particularly, for the licensed incumbent users the percentage ranges from 5 to 50% of the total users meaning that they send requests to the system with low or high frequency. This was introduced to our tests in order to simulate better the presence of this category of user, which is responsible of creating the most problems for the other two user types.

Parameter	Value
Macro BSs (with 3 cell each)	7
Inter-site distance	500 m
Small BSs	210
Total RAT Devices	231
Total Number of UEs	5000
Request inter-arrival time per user (sec)	Exponential (5.5 - 10)
File size (MByte)	2.0
Traffic model	FTP download
Total Number of channels	15 (3550-3700 MHz)
PAL Number of channels	10 (3550-3650 MHz)
GAA Number of channels	8 (3550-3700 MHz)
Traffic Mix	4 scenarios
Number of resource blocks:	50 PRBs per channel
Transmission Time Interval (TTI) length	1ms
Simulation time	60s

Table 12: Summary of simulation parameters

Scenarios	Incumbent	PAL	GAA
1	5%	32%	63%
2	10%	30%	60%
3	20%	26%	54%
4	50%	17%	33%

Table 13: Simulation scenarios of the users' traffic mix

9.3 Performance evaluation

Distributed radio resource management deals with the notion of moving management decisions related to RAT, spectrum, and channel selection (in our example) closer to the node level. In the proposed distributed scheme radio information is exchanged between neighbouring nodes and

decision is made locally (compared to centralized schemes which aggregate information and make decisions centrally). In SPEED-5G, the impact of centralized versus distributed approaches for channel selection was compared in order to better understand how, when and which approach is better for the different environments and scenarios (e.g., MBB, massive IoT) of 5G networks. To this extent, we evaluated the user plane-related latency in the radio access of the distributed approach and compared it to the centralized approach. Of course, this is not the only variable that has to be studied. The figure below provides an indication of user plane-related average latency versus load per cell for each user category (either incumbent, Priority Access Licence (PAL), or General Authorized Access (GAA)). Incumbent users are users which have absolute priority in the use of radio resources. GAA can use opportunistically the available resources if are not used by incumbents and/or PAL users.

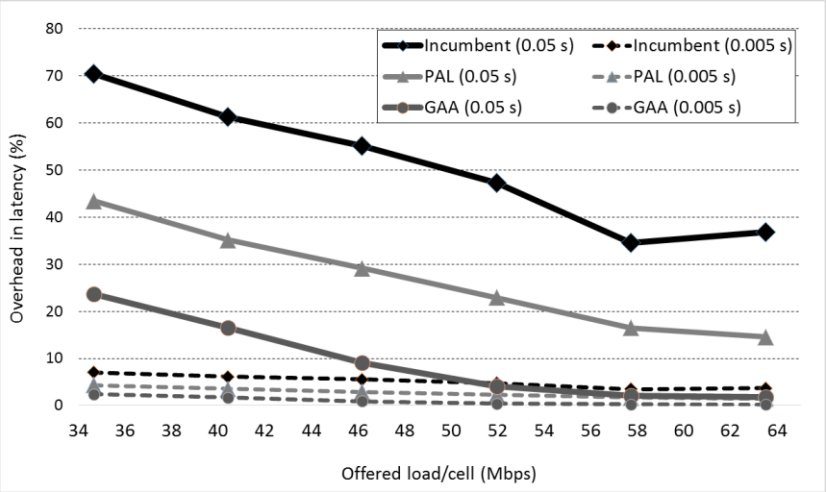


Figure 32: Air interface latency overhead in centralized approach compared to distributed as of load per cell

It can be seen that with increasing load also the latency rises and the difference between user categories is also greater in higher loaded cells compared to lower loads. Moreover, at the centralized approach, backhaul links between cells and the central entity are introduced to the system. The required overhead of backhaul channel knowledge feedback and the processing delay induced by the admission control algorithms need to be considered. In general, , the mentioned delay is composed of several factors; e.g., processing delay, serialization delay, queuing delay, etc. Backhaul characteristics will affect also the performance in terms of latency and thus have to be taken into account. As a result of the evaluation, a user-plane latency overhead with a range of 5 ms to 50 ms (0.05 and 0.005 s, respectively, as depicted in Figure 32) was obtained. It is observed that in higher loads it is better to have distributed decision making (in each cell) due to lower associated overhead latency. This is compared to a centralized decision making which assumes that various data would need to be transmitted to a central entity, which would execute the necessary actions for channel selection. A system with "good" backhaul characteristics will be able to handle both centralized and distributed resource management of the system. In contrast, if the backhaul conditions are not ideal (e.g. with high latency) – something that can easily occur and needs to be taken into consideration for wireless backhaul links – almost every user category experiences large overhead, which might reach 70% at low loads and almost 40% at high loads.

Figure 33 illustrates the relative average downlink throughput of a UE belonging to different access priorities. Specifically, it is shown that users (especially with higher priority, such as incumbents) can experience higher throughputs as packet arrival rate increases, but after a certain point PAL and GAA (lower priority users than incumbents) start to compete for radio resources and therefore their throughput drops. Moreover, relative packet transmission latency is better for higher priority users as the packet arrival rate increases, as depicted in Figure 34.

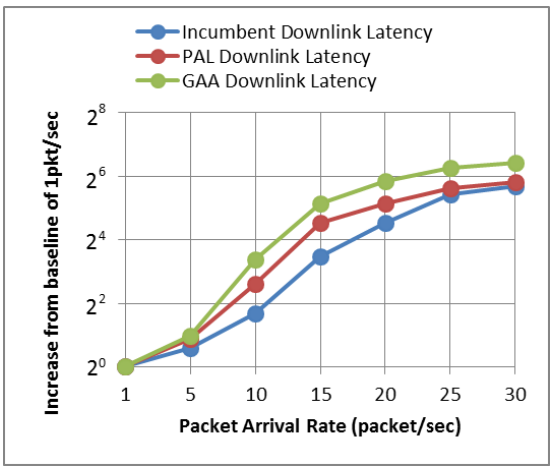


Figure 33: Relative increase of average downlink throughput and latency for different access priorities. Performance achieved at packet arrival rate of 1 packet per second (low load) is the baseline.

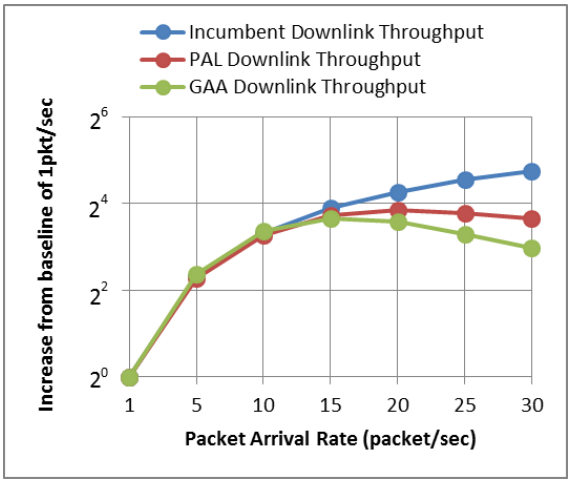


Figure 34: Relative increase of average packet transmission latency for different access priorities. Performance achieved at packet arrival rate of 1 packet per second (low load) is the baseline.

Figure 35 shows that for low system load the success ratio remains at 100%, but at 15 packets per second we can see an almost proportional reduction in the success rate. But this is not the case for traffic density of the whole area (Mbps/km^2) where we see that the increase in the throughput is not proportional.

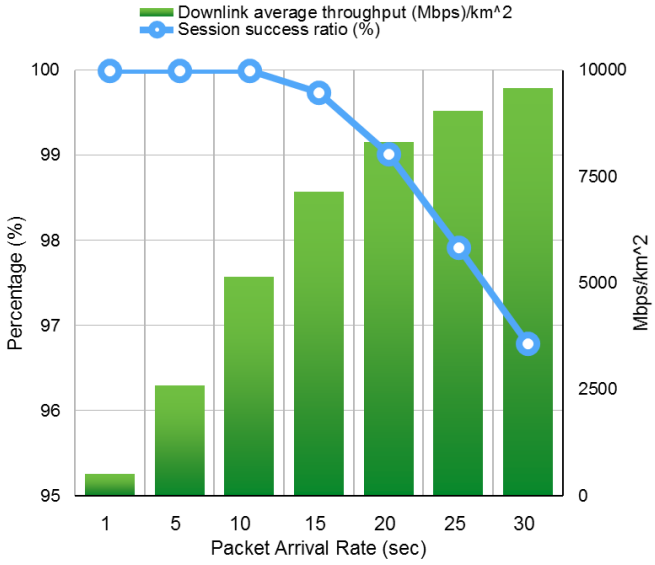


Figure 35: Session success ratio and downlink average throughput.

10 PtMP Wireless Backhaul Evolution

In order for the 5G vision to be realised, we need to consider all parts of the network. SPEED-5G targets also the backhaul, delivering key enhancements on capacity, latency and network infrastructure utilization. This section describes the evolution of a Point-to-Multipoint wireless system, namely WiBAS™ OSDR, towards 5G KPIs, with focus on data rate, latency, resource balancing and high availability. These enhancements are expected to lead to an evolved PtMP wireless system that can be used both in PtMP wireless backhaul - an application important to the operators due to its cost advantage over fibre and PtP wireless as well as its installation ease - and fixed wireless access, considered as one of the first expected applications of 5G.

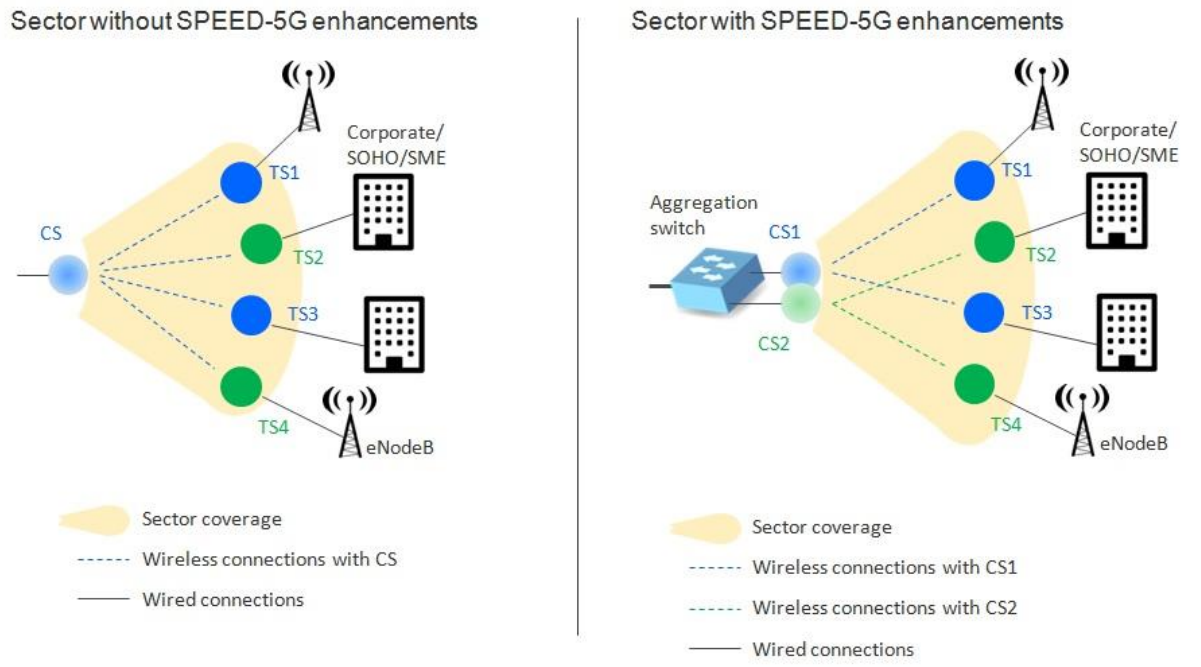


Figure 36: PtMP wireless backhaul and fixed wireless access use cases with and without SPEED-5G solutions

Figure 36 depicts the PtMP backhaul sector with and without SPEED-5G enhancements. Originally, the sector was aggregating traffic from N Terminal Stations (TS) to a single Central Station (CS). With SPEED-5G enhancements, a sector is aggregated by at least 2 CSs through an aggregation switch. It should be noted that all CSs "light up" the same area. A TS is wirelessly connected with exactly one CS, while a CS can have zero or more TSs (the exact number depends on resource balancing metrics). On the other end, a TS can service either an access Base Station (e.g. eNodeB), backhauling its cell traffic, or a building, providing high-speed access to its occupants.

The quantitative and qualitative targets under consideration are found in Table 14.

KPI	Description	SPEED-5G Target
Link data rate DL Link data rate UL	Downlink Data rate from a CS Uplink Data rate to a CS	≃ 1.0 Gbps ≃ 800 Mbps
Aggregate area capacity	DL and UL data rate between two CSs and their TSs	≃ 3.6 Gbps
One way latency	Latency between any CS and TS pair	~1.3ms DL ~0.8ms-1.4ms UL
Network availability	In case of a CS failure, service should still be provided over the working CS	Service restoration in < 5min
Resource balancing	Traffic load should be fairly distributed and network resources should be optimally assigned	During provisioning, BH should be able to auto assign a TS to the best CS

Table 14: SPEED-5G targets for the PtMP backhaul

10.1 Data rate increase

The increase of the aggregate area capacity is achieved by following two main strategies:

- Increasing the throughput per link
- Increasing the capacity per covered area

These solutions are described in the following sections.

Given that the current system has market-leading features (Table 15), it is expected that the evolved backhaul will constitute a significant step towards meeting 5G KPIs.

Throughput	SotA (Competition)	SotA (ICOM)	SPEED-5G Target
Downlink Data rate from a CS Uplink Data rate to a CS	300 Mbps 300 Mbps	540 Mbps 415 Mbps	≃ 1.0 Gbps ≃ 800 Mbps
DL and UL data rate between two CSs and their TSs	1.2 Gbps	1.9 Gbps	≃ 3.6 Gbps

Table 15: Throughput comparison with SotA and SPEED-5G evolution targets

10.1.1 Throughput per link increase

Throughput performance is upgraded by increasing the available channel bandwidth from 56 MHz to 112 MHz. Significant redesign was performed to all parts of the PtMP system WiBAS™ OSDR, i.e. the Baseband Unit, the RF Unit and the mechanical chassis that accommodates them. The main development steps taken were:

- Baseband modem FPGA redesign and clock frequency upscaling
- New and faster baseband FPGA part
- Redesigned RF Unit
- New Baseband board
- New enhanced power supply to meet the increased FPGA power needs

10.1.2 Capacity per area increase

In order to increase the area capacity, we use “aligned sector co-location”, where a second Central Station (CS2) is introduced, covering the same area. With this technique, the aggregate throughput is effectively doubled:

- In case the number of Terminal Stations is kept the same, we can split them to the two CSs, thereby allowing them to use higher rates.
- Otherwise, an increased number of Terminal Stations can now be deployed in the same area.

10.2 Latency

The one-way latency between any CS and TS pair is currently at ~1.9 ms at the DL and 1.2 ms to 1.8 ms for the UL. Based on a reduced frame duration, the target latency is in the area of 1.3 ms DL and 0.8 ms- to .4 ms UL.

10.3 Resource balancing

Resource balancing aims at the optimised association of TSs to the corresponding CS. The automatic TS entry at bootstrap is enhanced by an intelligent management application in order to select the best CS node, based on a performance metric such as number of CS connections or traffic load or processor load. An intelligent application running on an NMS (Network Management System) server, in cooperation with an AAA (authentication, authorization, and accounting) server, associates each TS to a CS, trying to satisfy the resource balancing requirement of the system.

Figure 37 shows a scenario where a fourth TS is inserted to a sector with three TSs already inserted and served by two CSs. The system uses the AAA server to validate TS4 and the NMS app to decide which one will “host” TS4. In the example, the traffic utilisation of the CSs is used as the resource balancing metric.

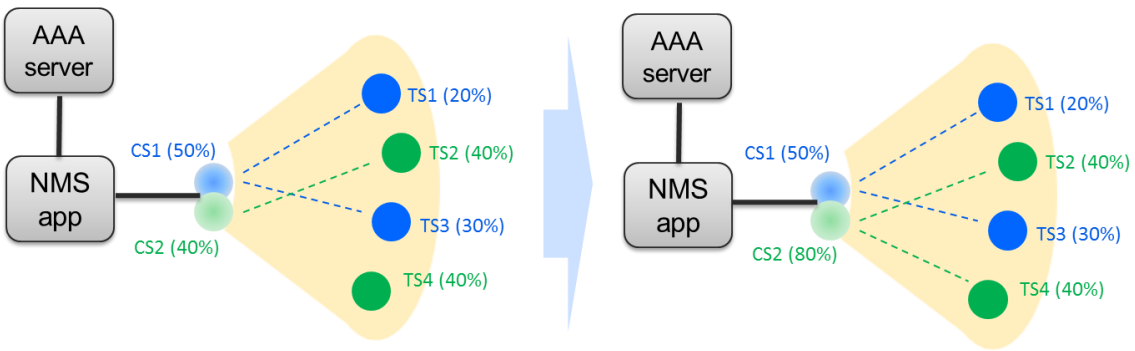


Figure 37: Resource balancing between 2 CSs

10.4 Network availability

In case of a CS failure, service should still be provided over the working CS. This is achieved by employing a scheme of redundancy at the CS (1:1 mode) and automatic frequency scanning at the TS. Both CSs are servicing their TSs, doubling sector's capacity and on a CS failure, all its TSs are switched to the other CS. Service should be restored in less than 5 minutes. In Figure 38, a scenario at which CS1 faces a failure is presented. In this, all TSs are eventually served by CS2, preserving the service.

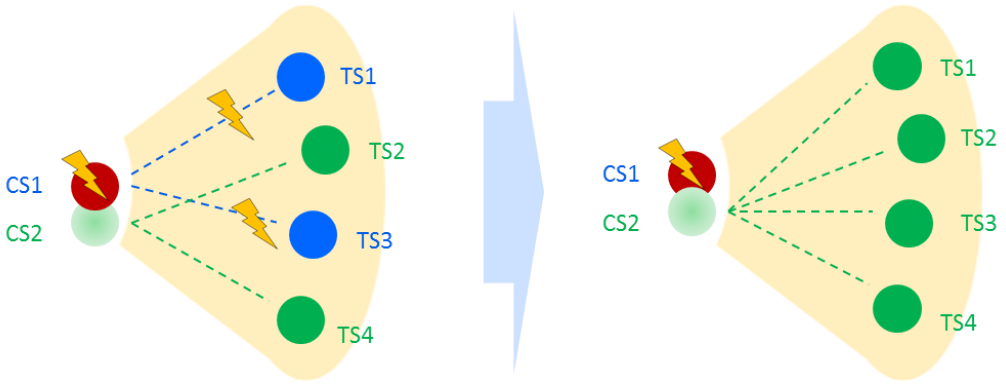


Figure 38: Increased availability with 2 CSs

References

[1] Anastasius Gavras (Editor), "D4.1 Metric definition and preliminary strategies and algorithms for RM", Speed-5G Project, Deliverable D4.1, Dec 2016. Available: https://bscw.5g-ppp.eu/sec/bscw.cgi/d135445/Speed5G-D4.1-v1.3_Metric_definition_and_preliminary_strategies_and_algorithms_for_RM.pdf.

[2] T. Sanguanpuak, S. Guruacharya, N. Rajatheva, and M. Latva-Aho, "Resource allocation for co-primary spectrum sharing in MIMO networks," in Proc. IEEE International Conference on Communication Workshop (ICCW), 2015, pp. 1083-1088.

[3] S. Hailu, A. A. Dowhuszko, and O. Tirkkonen, "Adaptive co-primary shared access between co-located radio access networks," in Proc. 9th Int. Conf. on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2014, pp. 131-135.

[4] J. Lindblom, and E. G. Larsson, "Does non-orthogonal spectrum sharing in the same cell improve the sum-rate of wireless operators?," in Proc. IEEE 13th Int. Workshop on Signal Processing Advances in Wireless Commun. (SPAWC), 2012, pp. 6-10.

[5] R. Litjens, H. Zhang, I. Noppen, L. Yu, E. Karipidis, and K. Borner, "System-Level Assessment of Non-Orthogonal Spectrum Sharing via Transmit Beamforming," in Proc. IEEE 77th Conf. on Vehicular Technology Conference (VTC Spring), 2013, pp. 1-6.

[6] P. Luoto, M. Bennis, P. Pirinen, S. Samarakoon, and M. Latva-aho, "Enhanced Co-Primary Spectrum Sharing Method for Multi-Operator Networks," in IEEE Transactions on Mobile Computing, in press.

[7] H. Holma, and A. Toskala, LTE for UMTS: OFDMA and SC-FDMA based radio access. Chichester, U.K.: John Wiley & Sons, 2009.

[8] http://www.hhi.fraunhofer.de/de/kompetenzfelder/imageprocessing/research-groups/image-video-coding/svc-extension-ofh264avc/jsvm-reference-software.html.

[9] R. Gerzaguet et al., The 5G candidate waveform race: a comparison of complexity and performance, EURASIP Journal on Wireless Communications and Networking December 2017.

[10] F. Bouali, K. Moessner, and M. Fitch, "A context-aware user-driven framework for network selection in 5G multi-RAT environments," in 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall), Sept 2016, pp. 1–7.

[11] Shahid Mumtaz (Editor), "D4.2 RM framework and modelling," Speed-5G Project, Deliverable D4.2, April 2017. Available: https://speed-5g.eu/wp-content/uploads/2017/06/speed5g-d4.2_rm-framework-and-modelling.pdf

[12] F. Bouali, K. Moessner, and M. Fitch, "'A Context-aware User-driven Strategy to Exploit Offloading and Sharing in Ultra-Dense Deployments," IEEE 2017 IEEE ICC Conference Proceedings, Paris, France, accepted.

[13] V. Marques, R. L. Aguiar, C. Garcia, J. I. Moreno, C. Beaujean, E. Melin, and M. Liebsch, "An IP-based QoS architecture for 4G operator scenarios," IEEE Wireless Communications, vol. 10, no. 3, pp. 54–62, June 2003.

[14] R. Agustí, O. Sallent, J. Pérez-Romero, and L. Giupponi, "A fuzzy-neural based approach for joint radio resource management in a beyond 3G framework," in QSHINE 2004, Oct 2004, pp. 216–224.

[15] X. Gelabert, O. Sallent, J. Pérez-Romero, and R. Agustí, "Radio access congestion in multiaccess/multiservice wireless networks," IEEE Transactions on Vehicular Technology, vol. 58, no. 8, pp. 4462–4475, Oct 2009.

[16] "5G Whitepaper: The Flat Distributed Cloud (FDC) 5G Architecture Revolution," Institute for

Communication Systems, 5G Innovation Centre, University of Surrey, Tech. Rep., January 2016.

[17] L. Giupponi, R. Agust´ı, J. Pérez-Romero, and O. Sallent, “A novel joint radio resource management approach with reinforcement learning mechanisms,” in PCCC 2005, April 2005, pp. 621–626.

[18] B. Ma, X. Liao, and X. Xie, “Vertical handoff algorithm based on type-2 fuzzy logic in heterogeneous networks.” JSW, vol. 8, no. 11, pp. 2936–2942, 2013.

[19] A. Kaloxylos, S. Barmpounakis, P. Spapis, and N. Alonistioti, “An efficient RAT selection mechanism for 5G cellular networks,” in 2014 International Wireless Communications and Mobile Computing Conference (IWCMC), Aug 2014, pp. 942–947.

[20] W. Zhang, “Handover decision using fuzzy MADM in heterogeneous networks,” in Wireless Communications and Networking Conference, 2004. WCNC. 2004 IEEE, vol. 2, March 2004, pp. 653–658 Vol.2.

[21] V. Ramkumar, A. Mihovska, N. Prasad, and R. Prasad, “Fuzzy-logic based call admission control for a heterogeneous radio environment,” in WPMC 2009, September 2009,” Invited paper.

[22] B. Z. Aymen, M. Ayadi, and S. Tabbane, “A fuzzy logic algorithm for RATs selection procedures,” in Networks, Computers and Communications, The 2014 International Symposium on, June 2014, pp. 1–5.

[23] Qualcomm news, introducing muLTEfire: LTE-like performance with Wi-Fi-like simplicity. [Online]. Available: https://www.qualcomm.com/news/onq/2015/06/11/introducing-multefire-lteperformance-wi-fi-simplicity.

[24] Open Mobile Alliance, “OMA Device Management Protocol,” Tech. Rep. version 1.2.1, June 2008.

[25] Intel et al., “WF on contention window adaptation based on HARQ ACK/NACK feedback,” 3GPP TSG RAN WG1 #82bis, Tech. Rep. R1-156332, October 2015.

[26]T. J. Ross, Fuzzy logic with engineering applications. Chichester, U.K. John Wiley, 2010.

[27] C.-L. Hwang and K. Yoon, Multiple Attribute Decision Making: Methods and Applications A State-of-the-Art Survey. Springer Berlin Heidelberg, 1981, ch. Methods for Multiple Attribute Decision Making, pp. 58–191.

[28] The network simulator-3 (ns-3). [Online]. Available: https://www.nsnam.org/

[29] 3GPP, “TSG RAN WG4 Meeting 51: Simulation assumptions and parameters for FDD HeNB RF requirements,” Tech. Rep. R4-092042, May 2009.

[30] “The ns-3 Model Library,” Release NS-3.26, October 2016. [Online]. Available: https://www.nsnam.org/docs/release/3.26/models/ns-3-model-library.pdf.

[31] J. Klaue, B. Rathke, and A. Wolisz, EvalVid – A Framework for Video Transmission and Quality Evaluation. Springer Berlin Heidelberg, 2003, pp. 255–272.

[32] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, April 2004.

[33] The Big Buck Bunny. [Online]. Available: https://peach.blender.org/

[34] Bismark Okyere (Editor), “MAC approaches with FBMC,” Speed-5G Project, Deliverable D5.2, June 2017. Available : https://speed-5g.eu/wp-content/uploads/2015/06/speed5g-d5.2_mac-approaches-with-fbmc-final.pdf

[35] Feilu Liu; Bala, E.; Erkip, E.; Rui Yang, "A framework for femtocells to access both licensed and unlicensed bands," in Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2011 International Symposium on , vol., no., pp.407-411, 9-13 May 2011

[36] https://www.lteuforum.org/index.html.

[37] Yujae Song; Ki Won Sung; Youngnam Han, "Coexistence of Wi-Fi and Cellular With Listen-Before-Talk in Unlicensed Spectrum," in Communications Letters, IEEE , vol.20, no.1, pp.161-164, Jan. 2016.

[38] Ratasuk, R.; Uusitalo, M.A.; Mangalvedhe, N.; Sorri, A.; Iraji, S.; Wijting, C.; Ghosh, A., "License-exempt LTE deployment in heterogeneous network," in Wireless Communication Systems (ISWCS), 2012 International Symposium on, vol., no., pp.246-250, 28-31 Aug. 2012.

[39] FCC White Paper: The Mobile Broadband Spectrum Challenge: International Comparisons, Feb 2013

[40] Zhang, Ran, Miao Wang, Lin X. Cai, Zhongming Zheng, Xuemin Shen, and Liang-Liang Xie. "LT Eunlicensed:the future of spectrum aggregation for cellular networks", IEEE Wireless Communications, 2015

[41] LTE-Unlicensed: The Future of Spectrum Aggregation for Cellular Networks.

[42] Almeida, E.; Cavalcante, A.M.; Paiva, R.C.D.; Chaves, F.S.; Abinader, F.M.; Vieira, R.D.; Choudhury, S.; Tuomaala, E.; Doppler, K., "Enabling LTE/WiFi coexistence by LTE blank subframe allocation," in Communications (ICC), 2013 IEEE International Conference on, vol., no., pp.5083-5088, 9-13 June 2013

[43] Nihtila, T.; Tykhomyrov, V.; Alanen, O.; Uusitalo, M.A.; Sorri, A.; Moisio, M.; Iraji, S.; Ratasuk, R.; Mangalvedhe, N., "System performance of LTE and IEEE 802.11 coexisting on a shared frequency band," in Wireless Communications and Networking Conference (WCNC), 2013 IEEE, vol., no., pp.1038-1043, 7-10 April 2013.

[44] Sagari, S.; Baysting, S.; Saha, D.; Seskar, I.; Trappe, W.; Raychaudhuri, D., "Coordinated dynamic spectrum management of LTE-U and Wi-Fi networks," in Dynamic Spectrum Access Networks (DySPAN), 2015 IEEE International Symposium on, vol., no., pp.209-220, Sept. 29 2015-Oct. 2 2015.

[45] S. Singh, S. P. Yeh, N. Himayat and S. Talwar, "Optimal traffic aggregation in multi-RAT heterogeneous wireless networks," Prof. 2016 IEEE Int'l Conf. on Comm. Workshops (ICC), Kuala Lumpur, 2016, pp. 626-631.

[46] Robert Mullins and Michael Barros (Eds) Draft 5GPPP White paper on Network Management and QoS, February 2016. Available from https://5g-ppp.eu/wp-content/uploads/2016/11/NetworkManagement_WhitePaper_1.0.pdf. Link checked September 2017.

[47] Ericsson, "Ericsson mobility report on the pulse of the networked society," White Paper, Nov. 2015. [Online] Available: https://www.ericsson.com/res/docs/2015/mobility-report/ ericsson-mobility-report-nov-2015.pdf

[48] 3GPP Technical Specification Group Radio Access Network, "TR 36.889, Study on Licensed Assisted Access to Unlicensed Spectrum, (Release 13)," V13.0.0, June 2015.

[49] Telemanagement Forum Business Process Framework (eTOM). Available at the website https://www.tmforum.org/business-process-framework/

Appendix A Theoretical analysis

The WiFi nodes are assumed to employ the clear channel assessment (CCA) mechanism to evaluate the channels availability. To this end, all WiFi nodes are considered to be able to detect surrounding nodes to avoid any collide in transmissions due to simultaneous channel access attempts. In LTE-U model, the WiFi considers a channel to be busy if energy level (ε_c) exceeds a certain threshold (CCA_T). Thus, the CCA mechanism is a key parameter to access the white spaces in LTE band. We assume transmit powers are denoted as $p_i\left(i\in\{\omega,\lambda\}\right)$ where ω and λ indices are to denote WiFi-Lic and LTE links respectively. We note that the maximum transmission power of an LTE small cell is comparable to that of the WiFi-Lic, and thus is consistent with regulations of unlicensed bands. The power received from transmitter j at a receiver i is given by $p_j\xi_{i,j}$ where $\xi_{i,j}\geq 0$ represents a channel gain which is inversely proportional to $\delta_{i,j}^{\gamma}$ where $\delta_{i,j}$ is the distance between i and j and γ is the path loss exponent. $\xi_{i,j}$ may also include antenna gain, cable loss and wall loss. SINR (Γ) of link i is given as

$$\Gamma_i=\frac{p_i\xi_{i,i}}{p_j\xi_{i,j}+\sigma_i},\ i,j\in\{\omega,\lambda\},\ i\neq j \tag{7}$$

where σ_i is noise power for receiver i. For the case of single WiFi-Lic and LTE, if i represents the WiFi-Lic links, then j is the LTE link, and vice versa.

A. SINR of WiFi Link

Therefore, SINR of WiFi link, i, $i\in\omega$, in the presence of LTE and no LTE is shown as

$$\Gamma_i=\begin{cases}\dfrac{p_i\xi_{i,i}}{\sigma_0} & \text{if no LTE;}\\ \dfrac{p_i\xi_{i,i}}{\sum_{j\in\lambda}p_j\xi_{i,j}+\sigma_0} & \text{if LTE,}\end{cases} \tag{8}$$

Here the term $\sum_{j\in\lambda}p_j\xi_{i,j}$ is the interference from all LTE networks at WiFi link i.

B. Throughput of WiFi Link

Firstly, for single WiFi link (in this case BoE model), CCA mechanism is originated as

$$R_\omega=\begin{cases}0 & \text{if}\quad \varepsilon_c\geq CCA_T;\ \text{(WiFi operation)}\\ f\left(\Gamma_{\omega(p)}\right) & \text{if}\quad \varepsilon_c\geq CCA_T;\ \text{(WiFi-Lic operation)}\\ f\left(\Gamma\right) & \text{if}\quad \varepsilon_c< CCA_T;\ \text{(Normal operation)}\end{cases}$$

and, if $\varepsilon_c < CCA_T$, the throughput is expressed as follows:

$$\begin{aligned}&T_S=f\left(R_\omega\right),\ \Lambda=f\left(R_\omega\right),\mathrm{E}[S]=T_E+T_S+T_C,\\ &\eta_E=\frac{T_E}{\mathrm{E}[S]},\ \eta_S=\frac{T_S}{\mathrm{E}[S]},\ \eta_C=\frac{T_C}{\mathrm{E}[S]},\ \Gamma=\frac{P(S)\Lambda}{\mathrm{E}[S]},\end{aligned} \tag{9}$$

Here $\mathrm{E}[S]$ is the WiFi-Lic expected time per packet transmission; T_E,T_S,T_C are the average times per $\mathrm{E}[S]$that the channel is empty due to random backoff, or busy due to successful transmission or packet collision (for multiple WiFi in the CSMA range), respectively. $P(S)$ is the AP successful probability of transmitting a packet in a given time slot . Λ is the average time spent transmitting the payload data.

If $\varepsilon_c\geq CCA_T$ (WiFi-Lic), the throughput is expressed as follows:

$$T_{LTE} = f(R_\omega),\ \Lambda = f(R_\omega), \mathrm{E}[S] = T_{sensing} + T_{\mathbf{WiFi\text{-}Li}_{\mathrm{Tx}}},$$
$$\eta_{sensing} = \frac{T_{sensing}}{\mathrm{E}[S]},\ \eta_{\mathbf{WiFi\text{-}Li}_{\mathrm{Tx}}} = \frac{T_{\mathbf{WiFi\text{-}Li}_{\mathrm{Tx}}}}{\mathrm{E}[S]},\ \Gamma = \frac{P(RB)\Lambda}{\mathrm{E}[S]} \quad (10)$$

$P(RB)$ is the successful probability of AP to transmit in a given resource block of LTE. Other parameter definitions can found in section 8.2.5.

C. SINR of LTE Link

The SINR of LTE link, $i, \forall i,$ in the presence of WiFi and no WiFi is shown as

$$\Gamma_i = \begin{cases} \dfrac{p_i \xi_{i,i}}{\sum_{j\in\lambda, j\neq i} p_j \xi_{i,j} + \sigma_0} & \text{if no WiFi;} \\ \dfrac{p_i \xi_{i,i}}{\sum_{j\in\lambda, j\neq i} p_j \xi_{i,j} + \sum_{\tau\in\omega} \alpha_\tau p_\tau \xi_{i,,\tau} + \sigma_0} & \text{if WiFi,} \end{cases} \quad (11)$$

Here the terms $\sum_{j\in\lambda, j\neq i} p_j \xi_{i,j}$ and $\sum_{\tau\in\omega} \alpha_\tau p_\tau \xi_{i,,\tau}$ indicate the interference contribution from other LTE links and WiFi links, (assuming all links in ω are active). For the τ^{th} WiFi link, $\forall \tau$, the interference is reduced by a factor α_τ to capture the fact that the τ^{th} WiFi is active approximately for only α_τ fraction of time due to the CSMA/CA protocol at WiFi.

D. Throughput of LTE Link

We are considering LTE TDD downlink mode. The LTE channel quality index (CQI) is used to calculate the peak throughout $\left(T_{LTE_{peak}}\right)$ of LTE which resembles to be eNB to User SINR. Moreover, CQI defines the modulation method (B_S → bits/symbol and C_R → coding rate) for data transmission.

Hence, we have

$$CQI = f(\Gamma);$$
$$B_S = f(CQI),\ C_R = f(CQI), \quad (12)$$
$$T_{LTE_{peak}}(bps) = \Psi(R_E B_S C_R)$$

Here R_E is the number of resource elements for a given bandwidth and Ψ is the control and signaling overhead for LTE.