Speed5G

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

D4.1: Metric definition and preliminary strategies and algorithms for RM

Version: V1.3

Deliverable type	R (Document, report)
Dissemination level	PU (Public)
Due date	30 June 2016 / *(revision 15/12/16)*
Submission date	15 July 2016 / *(revision 15/12/16)*
Lead editor	Anastasius Gavras, Eurescom GmbH
Authors	Keith Briggs, Michael Fitch, Federico Miatton, Ioannis-Prodromos Belikaidis, Voula Vassaki, Andreas Georgakopoulos, Evangelos Kosmatos, Evangelia Tzifa, Kostas Tsagkaris, Aikaterini Demesticha, Vera Stavroulaki, Panagiotis Vlacheas, Dimitrios Kelaidonis, Serafim Kotrotsos, Yiouli Kritikou, Athena Ropodi, Panagiotis Demestichas, Marinos Galiatsatos, Aristotelis Margaris, George Paitaris, Dimitrios Kardaris, Ioannis Kaffes, Shahid Mumtaz, Antonio de Domenico, Uwe Herzog, Adam Kapovits, Anastasius Gavras, Kazi Saidul Huq, Imran Khan, Paulo Marques, Jonathan Rodriquez, Dionysia Triantafyllopoulou, Faouzi Bouali, Seiamak Vahid
Reviewers	Thanasis Oikonomou, Federico Miatton
Work package, Task	WP4, T4.1
Keywords	Radio Resource Management, 5G Use Cases, Key Performance Indicators, Performance Metrics

Abstract

This deliverable describes the SPEED-5G relevant scenarios, requirements and metrics for integrating and assessing the proposed resource management enablers. Relevant state-of-the-art on non-continuous spectrum aggregation and multi-RAT (Radio Access Technologies) are detailed and a preliminary description of the proposed solutions is provided.

Document revision history

Version	Date	Description of change	List of contributor(s)
V0.1	26/04/2016	Initial version	Based on MS7 document
V0.2-v0.5	In May 2016	Incremental improvements	Anastasius Gavras
V0.6	31/05/2016	Restructuring of section 4	Anastasius Gavras
V0.11	04/07/2016	Incorporation of major contributions, executive summary, 2 missing analyses of white papers	WINGS, SISTEL, BT, EUR
V0.12	06/07/2016	Added conclusions, aligned metrics tables with D3.2 tables, introduced reference to D5.1	EUR
V0.12		Submitted for internal review	
V1.0	15/07/2016	Consideration of internal reviewer comments	
V1.1	25/11/2016	Updated version taking into account comments from period 1 review	Anastasius Gavras, Michael Fitch, Frederico Miatton
V1.2	28/11/2016	Final revised draft. Figure 6 update and polishing for delivery	Anastasius Gavras, Michael Fitch, Seiamak Vahid
V1.3	15/12/2016	Added references and mapping table to section 4	Michael Fitch

Disclaimer

Copyright notice

Acknowledgment

---

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

Executive Summary

This deliverable defines the goal of SPEED-5G in terms of Radio Resource Management (RRM), describes the relevant state-of-the-art (SoTA) and explains where the work so far goes beyond it. The report also provides descriptions of the SPEED-5G functions relating to proposed common RRM (cRRM) and discusses the architectural impacts with respect to the selected uses cases for demonstrating cRRM aspects.

The document briefly recaps the project use cases (Broadband Wireless, Massive IoT, Ultra-Reliable Communications and High Speed Mobility) that have been introduced in deliverable D3.1. Then it goes on to analyse the white papers provided by the vertical sectors – Energy, eHealth, Automotive, Factories of the Future, Media & Entertainment – and arrives at a new set of Key Performance Indicators (KPIs) and metrics for the use-cases. From these KPIs, conclusions are drawn about the wireless requirements, such as UL and DL bit-rates, UL / DL symmetry and latency. Guidelines are also provided to the MAC layer from this analysis, in terms of bit-rate range, packet size range, and mobility requirements.

The state-of-the-art section reviews RRM in other 5G PPP projects, in 3GPP and in the general field, and identifies the gaps and the areas where Speed-5G will focus. These areas are:

- Centralised RRM and the concept of inter-cRRM co-ordination

- The integration of RRM with higher level functions such as a spectrum manager

- The concept of monitoring and maintaining per-session KPIs

Moreover, the state-of-the-art on emerging technologies that will be used to implement the RRM functions have also been identified and include:

- Resource allocation in Device to Device (D2D)-based C-RAN
- Dynamic resource allocation algorithms for coexistence of LTE-U and WiFi
- Interference and QoS aware channel segregation for heterogeneous networks

The SPEED-5G approach in terms of RRM and its scope is presented through an introduction of Enhanced Dynamic Radio Access (eDSA) which is a project innovation that aims to increase the efficiency in the use of radio resources and transmit energy, while sustaining user quality-of-experience (QoE) .

The list of RRM functions that have been identified to be relevant to SPEED-5G are:

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

RRM functions are described in the scope of the three demonstration use cases, which are load balancing, channel selection and carrier aggregation. Architecture is elaborated in the context of each of these use cases, where the context includes indoor usage, interference and mobility issues.

Novel mathematical models are presented for spectrum sharing, D2D performance, co-existence of LTE and WiFi, interference and QoS aware channel segregation, and multiple attribute decision making (MADM) for channel selection using fuzzy logic techniques. Although the models presented in this report assume legacy i.e. LTE and WiFi RATs, they will be extended to address operational requirements of multi-RAT and multi-channel capable UEs and challenges associated with coexistence and inter-RAT cooperation. Results of the mathematical modelling will provided in the next deliverable, D4.2.

Table of Contents

List of Figures

List of Tables

Abbreviations

ABS	Almost Blank Sub-frame
CA	Carrier Aggregation
CC	Component Carriers
CMC	Connection Mobility Control
CoMP	Coordinated Multi-Point
CRE	Cell range extension
cRRM	Centralised RRM
CRS	Common Reference Signals
CSI	Channel State Information
D2D	Device-to-device
DC	Dual Connectivity
DER	Distributed Energy Resource
DRA	Dynamic Resource Allocation
EC	European Commission
EC	European Commission
eICIC	enhanced ICIC
eNB	Evolved NodeB
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FFR	Fractional Frequency Reuse
GTP	GPRS Tunnelling Protocol
HbbTV	Hybrid Broadcast Broadband TV
HDR or HDRI	High Dynamic Range Imaging
HeNB	Home eNodeB
HFR	High Frame Rate
HII	High Interference Indicator
ICIC	Inter-cell Interference Coordination
IP	Internet Protocol
IP	Internet Protocol
IRRRM	Inter-RAT Radio Resource Management
ITS	Intelligent Transportation System
LAA	Licensed-Assisted Access
LB	Load Balancing
LCS	Location Service Client
LOB	Live Outside Broadcast

LWA	LTE-WLAN aggregation
MAC	Medium access control
MBSFN	Multicast Broadcast Single Frequency Network
MME	Mobility Management Entity
OAM	Operations, Administration, Maintenance
OI	Overload Indicator
PDCP	Packet Data Convergence Protocol
PHY	Physical layer
PS	Packet Scheduling
RAC	Radio Admission Control
RAN	Radio Access Network
RAT	Radio Access Technology
RBC	Radio Bearer Control
RLC	Radio Link Control
RNTP	Relative Narrowband Transmit Power
RRC	Radio Resource Control
RRM	Radio Resource Management
RSRP	Received signal reference power
SCH	Shared Channel
TCO	Total Cost of Ownership
TCP	Transmission Control Protocol
UE	User equipment
UGC	User Generated Content
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
V2X	Vehicle to everything
WAVE	Wireless Access Vehicular Environment
WT	WLAN Termination Point

1 Key performance indicators

In SPEED-5G, we investigate the effectiveness of various RRM algorithms through simulation methods and, in some cases, also through measurements on testbeds. In order to compare RRM algorithms and methods, we need a procedure to measure the improvements that are achieved by each algorithm and method. The procedure has two basic parts, which are (i) the definition of a use case that sets up a deployment and traffic scenario, and (ii) the definition of Key Performance Indicators (KPIs) that are related to the use case and are a set of parameters with target values.

1.1   Use cases

The use cases of SPEED-5G have been described in deliverable D3.1. These are: Broadband Wireless, Massive IoT, Ultra-Reliable Communications and High Speed Mobility. The 5G PPP Phase 1 project METIS-II initiated a joint activity among the 5G projects within which the use cases defined by the individual 5G projects are collected and compared. This activity attempts to harmonise use cases and architecture across the projects within 5G PPP. It has been revealed that the use cases between projects differ to a varying extent. For example, the use cases of the METIS-II project are slightly different from the SPEED-5G ones in that they have different content but also in that they describe the scenarios in more detail. The mapping of the SPEED-5G use cases to the METIS-II use cases (or scenarios) is shown in Table 1.

Metis II use-cases

Speed-5G Use-case	Future Home Environment	Dense Urban Information Society	Virtual Reality Office	Suburban HetNet	Remarks
Broadband Wireless	Included as extreme broadband	Included as extreme broadband	Included as extreme broadband	Included as extreme broadband	Is included in all four Metis cases
Massive IoT	Included as massive MTC	Included as massive MTC		Included as massive MTC	Terminology
Ultra-reliable communications		Partially included as uMTC		Partially included as uMTC	Only MTC in overlap
High Speed Mobility					No mapping to Metis

Table 1. Mapping of SPEED-5G use cases to METIS-II use cases

The mapping has quite some overlap and the use cases of SPEED-5G that are distinct are:

(i) part of Ultra-Reliable Communications whereby SPEED-5G means to include all types of traffic (data, messaging etc.) and METIS-II only refers to uMTC and

(ii) High Speed Mobility is not addressed by METIS-II.

"Massive IoT" in the table above refers to low bit rate, low mass, small size and low power consumption.

1.2 Analysis of verticals' KPI requirements

1.2.1 Analysis of "5G and Energy" white paper

This section analyses the white paper[2] of the energy sector (version 30/09/15), whose use cases largely relate to the Massive IoT, but also to the Ultra-Reliable Communications use case of SPEED-5G.

1.2.1.1 Key transformations in the energy industry

- With technology evolution and growing concerns regarding the environment, the importance of renewable energy sources grew and consumers are increasingly turning into prosumers (energy producers)
- As a result, transformation of the power grid into a much more distributed model of generation and storage of power as well as micro-grids is underway, which also leads to the emergence of novel business models
- The problem is essentially a control problem that depends on communications with strict requirements
- The anticipated performance and flexibility of 5G will enable a communication infrastructure, which is able to support the emerging energy use cases of 2020 and beyond.

1.2.1.2 5G as a catalyser for energy efficiency

Communication networks for *smart grids*, *smart meters* and other use cases such as *electric vehicles (i.e. e-cars)* need to be able to serve the needs of various applications from automated meter reading and control of distributed energy resources to integration of e-cars to the energy systems. Taking smart grid as an example, the demand for efficient and reliable communication solutions is expected to grow due to the emergence of smart grids, and a lion share of the growth will take place in the medium-voltage and low-voltage domain towards secondary substations and distributed energy resources as well as between secondary substations and primary substation. These assets in a smart grid network currently have no communications or measurement equipment and 5G can provide economically viable wireless solutions, with respect to purely fibre-based communication systems, decentralising the energy networks with increased resilience compared to LTE. This way, future smart grids can be provided with an increased usage of protection, control and monitoring leading to improved power quality, fewer power outages, smaller power outage areas and easier grid deployments with less environmental impact in urban areas. Decentralised smart grids with increased resilience will also enable micro-grid solutions that will become important in blackout recovery.

1.2.1.3 Smart grid communication networks

Smart grid communication networks consist of multiple domains. Each of these domains serves a specific area e.g. a distribution network or location like a secondary substation (transformation between medium and low voltage) and has to meet individual requirements driven by the applications it serves. The communication network architecture, visualised in the figure below, is derived from and mapped to the power network it serves.

---

[2] https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White_Paper-on-Energy-Vertical-Sector.pdf
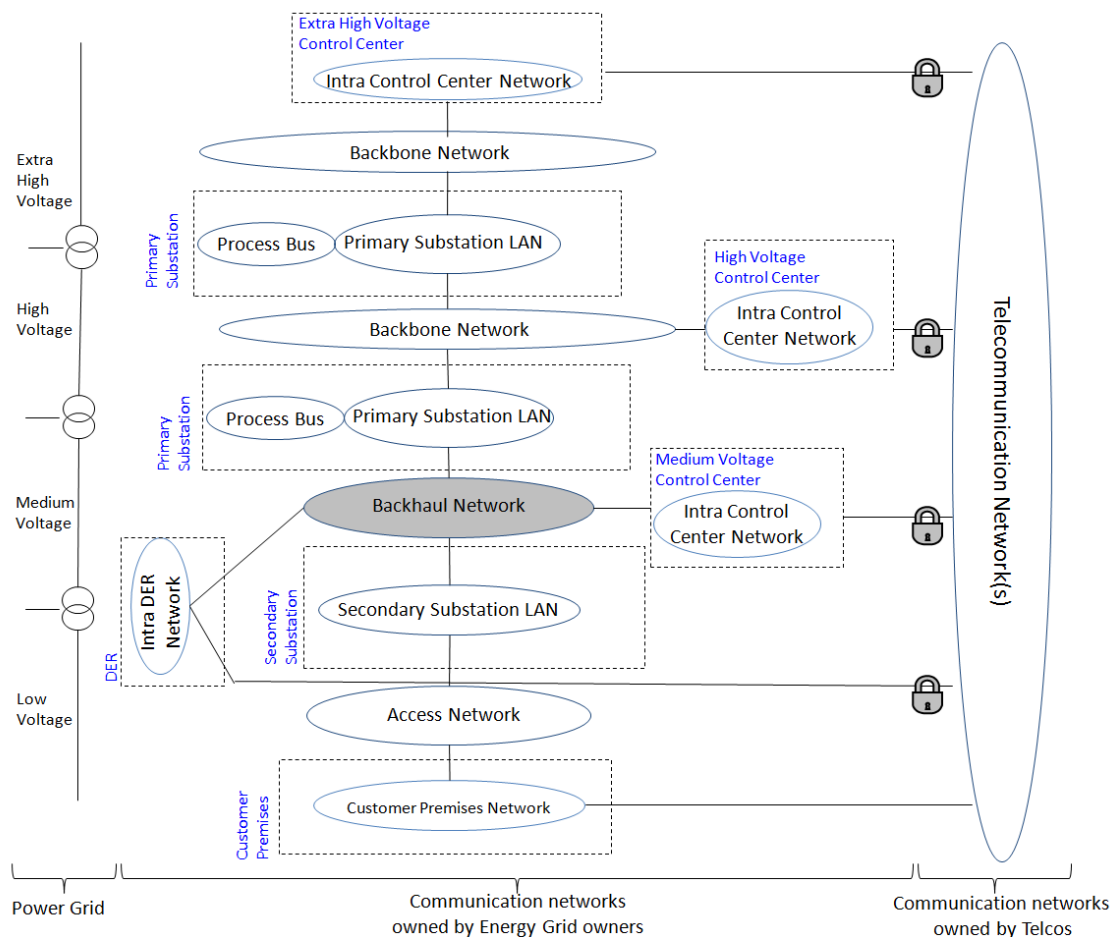
Figure 1. Communication network reference architecture in the energy sector

The border elements of these voltage levels are of major relevance for the Smart Grid communication domains as they have their own very specific requirements:

- *Primary Substation*: the medium voltage to high voltage transformation point, and the high voltage to extra high voltage transformation point
- *Secondary Substation*: low voltage to medium voltage transformation point

The following communication network domains can be distinguished:

- *Grid Backbone Communication Network*: communication network which connects the Primary Substation LANs with each other, with their regional control centres (often collocated) and with the central control centres
- *Primary Substation LAN*: a Primary Substation LAN is quite complex and requires its own communication infrastructure that distinguishes between a Process Bus and a Station Bus. It is mainly based on a Gigabit Ethernet infrastructure
- *Grid Backhaul Communication Network*: communication network which connects the Secondary Substation LANs with each other and with a control centre. This network domain might also connect to the respective Primary Substation LAN
- *Secondary Substation LAN*: network inside the secondary substation (today this network is quite trivial and may consist of just one single Ethernet switch / IP router). The Secondary Substation LAN is implemented in US-Style regions more in a distributed manner whereas in Europe the Secondary Substation LAN is very often located in an encapsulated enclosure
- *Grid Access Communication Network*: communication network which connects the customer premises or e.g. low voltage sensors to a specific Secondary Substation

- *Customer Premises LAN*: in-building communication network; a customer is characterised by consumption and production of energy (prosumer) and the customer can be residential, public or industrial prosumer
- *Intra DER (Distributed Energy Resource) Network*: for medium-sized DERs like wind and solar parks, a dedicated LAN is required for control, management and supervision purposes
- *Intra-Control-Centre Network*: LAN within a Distribution System Operator's or Transmission System Operator's control centre
- *External Network*: fixed or mobile network operator owned communication network which offers different connectivity services either via dedicated services or via the open Internet

The communication network domains in the Smart Grids context where 5G is immediately expected to play a significant role are the access networks (connecting the elements in the low voltage power grid) and the backhaul networks (connecting the elements in the medium voltage power grid). For the far future 5G might also play a role in other network domains like the backbone network domain (connecting the elements in the high voltage and extra high voltage power grid), which has the most stringent requirements in terms of real-time and reliability.

1.2.1.4   Technical requirements and KPIs for Smart Grid use cases

Application area	Grid access communication network domain	Grid backhaul communication network domain	Grid backbone communication network domain
Resilience (in terms of power independent operation in case of power outage) in hours	From 8-12 hours up to 72 hours for the most critical services and sites		
Region to be covered (diameter in km)	< 10km	< 100 km	< 1000km
Data rate	1 kbps (even in basements)	Several Mbps (between secondary substations and towards control centre)	Mbps - Gbps
End-to-end latency– guaranteed upper bound	< 1s (from control centre / meter data management centre / secondary substation to the smart meter)	< 50ms (between secondary substations and towards control centre)	< 5ms (between primary substations and towards the control centre)
Packet loss	No specific requirement as long as E2E latency requirement is covered (TCP-based communication is dominant)	$< 10^{-6}$ (E2E latency requirement has to be covered, as well as non-acknowledged status information distribution – e.g. IEC 61850 8-1-based – has to function properly)	$< 10^{-9}$ (E2E latency requirement has to be covered, as well as non-acknowledged status information distribution – e.g. IEC 61850-8-1 based – has to function properly, which is more demanding for high/extra high voltage than for medium voltage applications)

Application area	Grid access communication network domain	Grid backhaul communication network domain	Grid backbone communication network domain
Availability	99%, equalling 9h downtime p.a.	99,99%, equalling 50 min downtime p.a.	Better than 99,999%, equalling less than 5 min downtime p.a.
Failure convergence time	< 1s	< 1s	Seamless failover required, i.e. no loss of information in case of a failure while maintaining real-time delivery of information (within a small number of milliseconds)
Handling of crisis situations – surviving medium power down-times on a large scale, assuring black-start capability	Not required	Mandatory (can be realised by suitable battery backup or other 5G technical solutions that allow switching between different communication technologies which are guaranteed not to fail at the same time)	Mandatory
Summary	Less demanding and critical than in the other application domains, but some real-time capability is required	Stable, reliable, secure and real-time capable communication is required	Ultra-stable and reliable, secure and real-time capable communication is required

Table 2. Technical requirements and KPIs for Smart Grid use cases

The 5G and Energy white paper [3] concludes the following:

"In summary, Smart Grid services are often associated with stringent requirements in terms of reliability, tolerating only short packet dropouts. … To accommodate these under a common network topology, novel technology components are needed, as well as some disruptive techniques. The design of the new radio access network involves the interplay among various radio interfaces that are seamlessly integrated and results in a radical paradigm shift on the connectivity concept in the future 5G vision. Efficient integration of the 5G access technologies includes multi-connectivity approaches where the user equipment is simultaneously connected to several access technologies or frequency bands which could help to address the requirements in terms of crisis situation handling."

"Reliability-of-service will have to be orders of magnitude higher than in current wireless access networks, usually in combination with stringent E2E latency requirements, e.g. for the grid backbone communication network domain below 5 ms, while the acceptable downtime per year must not exceed 5 minutes, and data rates in the order of Mbps or even Gbps are required."

---

[3] https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White_Paper-on-Energy-Vertical-Sector.pdf

1.2.2 Analysis of "5G and Media & Entertainment" white paper

This section analyses the white paper[4] of the media and entertainment sector (version 8, 20/11/15).

The applications in the Media & Entertainment (M&E) sector (e.g. gaming, video streaming) match well with most of the scenarios in the SPEED-5G Broadband Wireless use case and the respective requirements from M&E could thus be considered for that.

1.2.2.1 Key transformations in the media and entertainment industry

5G shall enable at least **six main families of M&E use cases** in the 2020s with an overall user experience that well exceeds that of 4G and other legacy networks:

- Ultra-high fidelity media (3D, Full 4K, HDR, HFR, 8K)
- On-site live event Experience (large scale event sites, such as cinemas, stadiums and hall parks)
- User & machine generated content: people and objects will capture more and more content in order to share it with others in the cloud
- Immersive and integrated media: to provide ambient media consumption at home but also on the move, with content capable of following the users and adapt to his ambience for viewing
- Cooperative media production: content will be captured and shared immediately, by multiple people in multiple locations simultaneously
- Collaborative gaming: in a full immersive multi-sensorial environment

1.2.2.2 5G uses cases for the M&E industry

5G use cases for the media and entertainment sector include, among others:

- Publishing (books): security (DRM), end user context
- Film: security (DRM), end user context
- Music: security (DRM), end user context
- Audio/video: interactivity, security, end user context
- Advertising: location, business model, augmented reality, end user context
- Designers: interactivity, virtual reality
- Cultural heritage: storage, security, durability (never lost)
- Fashion: augmented reality, virtual reality
- Game: interactivity, end user context, virtual reality
- Programme Making and Special Events (PMSE): carriage of low latency audio/video as well as provision of office functions

1.2.2.3 Technical requirements and KPIs for M&E use cases

SPEED-5G KPIs	Media & Entertainment requirements (as extracted from the white paper)	Translation into numeric / specific figures and KPIs
Peak data rate DL	Full HD or 4K UHD (# of pixels 3840 × 2160 , i.e. 4x full HD), complementing linear TV and audio broadcasting. Video 360° will give a new	~20-35 Mbit/s for 4K[5] ~50-90 Mbit/s for 8K

---

[4]   https://5g-ppp.eu/wp-content/uploads/2016/02/5G-PPP-White-Paper-on-Media-Entertainment-Vertical-Sector.pdf

[5] https://en.wikipedia.org/wiki/Ultra-high-definition_television

SPEED-5G KPIs	Media & Entertainment requirements (as extracted from the white paper)	Translation into numeric / specific figures and KPIs
	immersive dimension to content consumption.	
Minimum guaranteed data rate DL	Anytime, at the busy hour without signs of congestion, excessive latency and delays, low error rates for video	Same as above for 4K / 8K ~5-8 Mbit/s for Full-HD
Peak data rate UL	Growth of User Generated Content (UGC) increases the need for high uploading network capabilities. Live Outside Broadcast (LOB) production	At least ~5-8 Mbit/s to enable Full-HD for UGC. LOB production might require 4K resolution (~20-35 Mbit/s)
Minimum guaranteed data rate UL		
Connection density	Significant peaks of simultaneous viewing during large live events; be it sporting events such as the Olympic Games or popular entertainment shows.	For e.g. 80,000 spectators in a stadium it would be 1.6m/km²
Traffic density	The QoS shall not depend on the size of the concurrent audience.	
Radio latency	Latency should be similar to current TV distribution mechanisms like satellite, DTT or IPTV. Low latency for gaming.	For live TV viewing: < 5s For gaming: 50ms E2E For LOB production: 1s
End-to-end latency	Consumers need a sufficient Quality of Experience which can be provided by a controlled throughput and latency. For LOB production, low latency and ruggedness of transmission are key requirements	
Mobility	-	
Availability	For LOB production, low latency and ruggedness of transmission are key requirements	99.999%
Energy consumption	-	
Fairness	-	
BLER	Ensure Quality of Service for M&E enjoyment, e.g. low error rates for video	Not quantifiable
Spectrum efficiency	A single frequency network configuration should be possible to allow efficient use of spectrum (for linear TV broadcast), avoiding the need of excessive spectrum for live TV delivery	The requirement from M&E is more of a regulatory / network frequency planning nature

SPEED-5G KPIs	Media & Entertainment requirements (as extracted from the white paper)	Translation into numeric / specific figures and KPIs
Energy efficiency	-	Not quantifiable
Cost efficiency	Ability to deliver live TV, on demand video and games at sustainable cost	Requirement is too unspecific to quantify
Misc.	Must operate indoor, deep indoor, portable outdoor and mobile	Not quantifiable

Table 3. Technical requirements and KPIs for M&E use cases

1.2.2.4   Further technical requirements

In addition to the above KPIs relevant for SPEED-5G, there are a number of further technical requirements expressed by the M&E sector. They are listed here just for information and include:

- Ability to provide content **free-to-air**
- Delivery of PSM content to the public **without blocking or filtering** of the service offer
- Content and service integrity - **no modification of the PSM content** or service by third parties; e.g. TV content and additional services (e.g. subtitles, HbbTV) must be displayed on screen unaltered and without unauthorised overlays
- **QoS** to be defined by the broadcaster/national regulator, including availability of a network, robustness, up-time, reliability
- QoS for each user shall be **independent of the size of the audience**
- PSM shall **not** be subject to **discrimination** compared to equivalent services
- **Geographical availability** of the service (e.g. national, regional, local)
- Support of **at least a minimum service offer** (e.g. a minimum number of programmes)
- Ease of use - straightforward accessibility and prominence of the PSM offer
- Low barrier for access to PSM content and services **for people with disabilities** (e.g. subtitles, audio description and signing)
- **Seamless handover** between different heterogeneous network entities
- Ability to reach audiences **in emergency** situations
- Dynamic **switching between unicast, multicast and broadcast** modes depending on the demand and traffic load in the network
- Possibility to use 100% of available spectrum resources in unicast, multicast or broadcast mode, respectively
- **Possibility to operate a 5G network in a standalone broadcast mode** for the distribution of linear TV and radio services
- Heterogeneous network topology, i.e. a **mixture of high-power-high-tower and low-power-low-tower infrastructure**, to adapt network coverage to requirements in an optimised manner

1.2.3   Analysis of "5G Automotive Vision" white paper

This section analyses the white paper[6] of the automotive sector (version 30/09/15). Notably, the following citations summarise well the requirements of the automotive industry.

> *"…automated vehicles will have to rely not only on their own sensors, but also on those of other vehicles, and will need to cooperate with each other, rather than make decisions on their own. These trends pose significant challenges to the underlying communication system…*

---

[6] https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White-Paper-on-Automotive-Vertical-Sectors.pdf

*…improved performance in terms of **reduced latency, increased reliability and higher throughput under higher mobility and connectivity density**…"*

1.2.3.1   Key transformations in the automotive industry

- Automated driving
- Road safety and traffic efficiency services
- Digitalisation of transport and logistics (information and data relating to goods, means of transport, authentication and access to ports or customs clearance information, increase of road transport efficiency)
- Intelligent navigation
- Information society on the road (vehicle dashboards providing appropriate Human Machine Interfaces (HMIs) for using leisure and entertainment or using the car as second office), will demand **high data rate** (e.g. HD / 4K video streaming) and **low latency** connectivity

1.2.3.2   Link to 5G objectives

Automotive industry has developed and standardised technologies for direct communication between vehicles (V2V), as well as between vehicles and roadside infrastructure (V2I) – ITS-G5 in Europe (5.9 GHz band) and WAVE in the United States. These technologies act locally, with low latency, and availability is guaranteed as long as the communication partners are within a certain communication range. In the mobile networking industry, supporting V2X over cellular networks is rapidly gaining interest. 3GPP RAN is currently working to enhance LTE (Long Term Evolution) in Release 13 and Release 14 (beginning of 5G) to fulfil requirements for V2X over licensed and unlicensed spectrum (3GPP TSG RAN Meeting #68, June 2015). IEEE 802.11-based V2X communication technology is a short-range ad hoc broadcast system developed for the exchange of object information and not for the exchange of sensor data.

There is need for a complementary communication technology for the exchange of cooperative information with higher bandwidth and improved reliability. Worth mentioning is in particular: (1) the exchange of sensor data for collective perception (e.g. video data), (2) the exchange of control information for platoons from very close driving vehicles (down to only 1 meter away) and (3) the exchange of vehicle trajectories to prevent collisions (cooperative decision making, very fast re-planning of vehicle trajectories).

Availability of wireless services is crucial for wide deployment of ITS services. Network coverage along roads and in low-density areas will be important. Where network coverage cannot be guaranteed, Device-to-Device (D2D) communication will be essential. Data has to be available practically everywhere, not necessarily stored but transferred fast and efficiently from central and local points to various destinations, in particular vehicles on the move.

1.2.3.3   Technical requirements and KPIs for V2X use cases

	Automated overtake	Cooperative collision avoidance	High density platooning	See-through	Vulnerable road user discovery	Bird's eye view
E2E latency	10 ms	1-5 ms	10 ms	50 ms		50 ms
Reliability	10^{-5}	10^{-5}	10^{-5}			
Data rate				10 Mbit/s		40 Mbit/s
Positioning accuracy	30 cm	30 cm	30 cm		10 cm	

Table 4. Technical requirements and KPIs for V2X use cases

1.2.3.4   Connectivity demands of future connected vehicles

	Negotiation with other road users (coop. manoeuvre)	Information to other road users (coop. perception)	Exchange with other road users (prediction of situation)	Sensors of other road users	Comm. to backend server
Data rate	Up to 1 Mbit/s	Up to 20 Mbit/s	Up to 1 Mbit/s	Up to 20 Mbit/s	A few Mbit/s
E2E latency	5-10 ms	10-50 ms	10 ms	10-50 ms	100 ms

Table 5. Connectivity demands of future connected vehicles

1.2.3.5   Digitalisation of transport and logistics

Remote sensing and control

- No stringent requirements in terms of latency (< 1s) or data rate (as for any IoT application)
- Nevertheless, must operate in challenging reception scenarios such as underground parking
- Energy efficiency: LTE Rel-12 specifies solutions to reach 10 years battery life with two AA batteries. 5G releases will aim at 15 years battery life and 30 dB coverage extension

Remote processing for vehicles

- Latency: may be very stringent or not, depending on the processes (down to a few ms)
- Bandwidth: also depending, e.g. for uploading all sensor data: up to 100 Mbit/s

Information society on the road

- high data rates (tens of Mbit/s)
- low latencies
- moving at velocities close to 200 km/h
- good coverage levels along the road infrastructure

1.2.3.6   Conditions under which the latency and reliability figures should be achieved

Scenario	Vehicle density (vehicles/km2)	Relative velocity (km/h)	Communication range (m)	Offered load (Mbit/s/vehicle) (average/peak)
Urban	1000-3000	0-100	50-100	1.0 / 10
Suburban	500-1000	0-200	100-200	0.5 / 10
Highway	100-500	0-500	200-1000	1.0 / 10

Table 6. Conditions for latency and reliability

1.2.4   Analysis of "5G and eHealth" white paper

The section analyses the white paper[7] of the eHealth sector (version 30/09/15), whose use cases largely relate to the Massive IoT, but also to the Ultra-Reliable Communications use cases of SPEED-5G. Specific cases relate to a combination of Broadband Wireless and Ultra-Reliable Communications.

---

[7] https://5g-ppp.eu/wp-content/uploads/2016/02/5G-PPP-White-Paper-on-eHealth-Vertical-Sector.pdf

1.2.4.1 Key transformations in the eHealth industry

The key transformations in the eHealth industry are driven by the following trends:

- Deliver treatment or care outside hospitals

  o in homes
  o general practitioners' premises
  o nursing homes
  o day surgeries / clinics
  o rehabilitation
  o over networks (for example the Internet)
  o rural areas

- Attention is being shifted to the root causes

  o lifestyle
  o wellness

- Increased attention to care for people with one or more chronic conditions

  o Prevention: action to reduce or eliminate the onset, causes, complications or recurrence of diseases
  o Compensation & support: action to reduce physical or cognitive impairments
  o Independent & active ageing: action to allow autonomous living, participation in social activities and remaining longer at work

Further trends with social context are as well driving the transformation

- Community services for patients and users allowing voluntary sharing of observations with peers
- Services for benchmarking and competitions in the scope of wellbeing/healthy lifestyle
- Personalised care
- Reducing the risk of accidents, like falls and burns
- Enabling informal carers and lesser qualified professionals to take on routine tasks of more qualified individuals to reduce the burden to social systems
- Self-care
- Social/crowd care (including community and rating services)
- Patient-controlled information sharing/access
- Education and behavioural change communication

1.2.4.2 Requirements on 5G-based on use cases

The requirements are in the sector are broadly categorised along the lines of "cared for" and "cared by", i.e. requirements related to Users / Data owners and Health / Insurance providers. Four use cases are used to illustrate the broad range of requirements of the eHealth vertical sector:

- Assets and interventions management in hospitals
- Robotics - remote surgery and service robotics for assisted living
- Remote monitoring of health or wellness data
- Smart medication

	Assets management	Surgery and service robotics	Remote monitoring	Smart medication
(Peak) Data Rate	Below 10 Mbps	300 Mbps for HD video streaming / augmented reality	300 Mbps for multiple video streams	1 Mbps

	Assets management	Surgery and service robotics	Remote monitoring	Smart medication
Mobility (speed)	500 km/h for intervention in helicopter	N/A	250km/h (at max. typical vehicle speeds)	250km/h (at max. typical vehicle speeds)
Density / Number of Devices	>100/m2 in hospitals	5-10 surgical robots per hospital, several 100s care robots per hospital	Up to 50 connected objects per home. In dense populated areas (hotspots) up to 10000/km2	Up to 10 smart medication per person. In dense populated areas (hotspots) up to 5000/km2
Reliability	99.999%	99.99999%	99.9%	99.999%
E2E Latency	10s-1min	10 ms (100 ms latency for haptic systems)	10s-1min	10s-1min
Coverage	Very deep indoor, rural areas, air and maritime for interventions	Very deep indoor	Very deep indoor, rural areas, reach >99.9% of population	Very deep indoor, rural areas, reach >99.9% of population
Positioning Accuracy / Location	30 cm indoor	N/A (done by the robot)	<1 m indoor and outdoor	<1 m indoor and outdoor
Data Volume	Not specified	Very large data volumes due to HD imaging	Not specified	Not specified
Energy Efficiency	Not specified	Not specified	Wearable and implants should be self-sufficient for treatment duration that could be in the range of years	Similar to remote monitoring

Table 7. Technical requirements and KPIs for eHealth use cases

1.2.5   Analysis of "5G and the Factories of the Future" white paper

This section analyses the white paper[8] of the Factories of the Future vertical sector (version 30/09/15), whose use cases largely relate to the Ultra-Reliable Communications and to some extend to the Massive IoT use cases of SPEED-5G.

1.2.5.1   Key transformations in the Factories of the Future industry

Two main trends in manufacturing are driving this transformation and will influence the future competitiveness:

---

[8]   https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White-Paper-on-Factories-of-the-Future-Vertical-Sector.pdf

- the increasing role of services in manufacturing and
- the growing importance of global value chains. It is estimated that, by 2025, manufacturers will get more revenue from services than from products.

This is a consequence of a trend called "servitisation of manufacturing", indicating a shift from solely selling produced goods to providing added value services together with either connected (smart) or non-connected goods. The growing importance of global value chains is a second trend that drives the demand for truly connected manufacturing ecosystems.

Regarding the 5G roadmap and future research, the white paper recommends focusing on:

- high throughput with research on ultra-reliable wireless deterministic communication
- high availability with research on proper security mechanisms and ubiquitous coverage and device-to-device traffic/service offloading mechanisms
- lowering the TCO (Total Cost of Ownership) with research on network capabilities to manage heterogeneity
- high flexibility with research on plug-and-produce capabilities by adopting internet technologies into industrial stacks
- new data-driven business models for SMEs and large companies, exploiting data-oriented services and network virtualisation concepts
- changes needed with respect to legislative framework, standards and social acceptance
- building a specific strategy for 5G and manufacturing SMEs

1.2.5.2   Requirements on 5G-based on use cases

With respect to manufacturing, five use case families have been identified to illustrate the requirements along supply chain and manufacturing networks for realizing the next generation connected factory. The table below shows the five use case families, illustrated with representative scenarios and highlights the potential impact on manufacturing.

Each use case family represents a different subset of stringent requirements along supply chain and manufacturing networks:

- Time-critical process optimisation inside factory
- Non time-critical optimisations inside factory
- Remote maintenance and control
- Seamless intra-/inter-enterprise communication
- Connected goods

	Time-critical process control	Non time-critical factory automation	Remote control	Intra/inter-enterprise communication	Connected goods
(Peak) Data rate	Mbps-Gpbs	Mbps-Gpbs	Mbps-Gpbs	Mbps-Gpbs	1-100kbps
Mobility (speed)	10m/s	10m/s	4km/h	4km/h	0-100km/h
Density / Number of Devices	10-100/m^2	10-1000/m^2	-	-	-
Reliability	99.999%	99.999%	99%	99.99%	99%
E2E Latency	100us-10ms	100ms-1s	50ms	100ms	minutes

Table 8. Technical requirements and KPIs for Factories of the Future use cases

1.3   Mapping of use cases to KPIs

The SPEED-5G KPIs have been mapped against the SPEED-5G use cases and the result is shown in the following tables. Table 9 contains Massive IoT requirements, Table 10 contains those for Broadband Wireless, Table 11 contains Ultra Reliability Communications and Table 12 addresses High Speed Mobility. This is novel work to this project.

KPI	General requirements from network operator / vendor estimates for smart wearables, sensor networks and automated home	Example vertical requirement: 5G and Energy white paper
Peak data rate DL*	–	- Grid access: 1Kbps (even in basements) - Grid backhaul: Several Mbps (between secondary substations and towards control centre) - Grid backbone: Mbps - Gbps
Minimum guaranteed data rate DL*	10Kbit/s to 1Mbit/s	
Peak data rate UL*	–	
Minimum guaranteed data rate UL*	10Kbit/s to 1Mbit/s	
Connection density*	200K devices per km^2	–
Traffic density*	Zero mobility: packet lengths 125 bytes (1250 optional), packet intergeneration time 1 min (optional), 5 min, 30 min, 1 hour. Limited mobility: packet length 125 byte, intergeneration time 5s (optional), 10s, 30s.	–
Radio latency*	100ms	–
End to end latency**	1 second	- Grid access: < 1s (from control centre / meter data management centre / secondary substation to smart meter) - Grid backhaul: < 50ms (between secondary substations and towards control centre) - Grid backbone: < 5ms (between primary substations towards the control centre)

Mobility*	Zero and limited, where zero is pedestrian (0 - 3km/h) and limited is 95% at zero and 5% at 0 - 120km/h [METIS Dense urban uses 3km/h for small cell and 30km/h for macro cell]	–
Availability***	–	–
Energy consumption	Batteries to last tens of years.	–
Fairness	Congestion and interference requires managing	–
BLER	–	- Grid access: No specific requirement as long as E2E latency requirement is covered (TCP-based communication is dominant) - Grid backhaul: < 10^{-6} (and latency!) - Grid backbone: < 10^{-9} (and latency!)
Spectrum efficiency	–	–
Energy efficiency	Batteries to last tens of years.	–
Cost efficiency	–	–
Misc.	Over the air programming essential. Coverage is critical for some IoT cases outside of urban areas, but these may be out of scope for SPEED-5G. Automated secure registration and bootstrapping needed. Privacy and security has impact on the radio network.	Resilience (in terms of power independent operation in case of power outage): From 8-12 hours up to 72 hours for the most critical services and sites.

*Source = NGMN, ** Source = FMCF, *** Source = 3GPP

Table 9. Massive IoT

KPIs	General requirements from network operator / vendor estimates for gaming, large and fast file downloads, video streaming, linear TV, TV on demand, business transactions	Example requirements from media and entertainment vertical sector
Peak data rate DL*	300Mbit/s (50Mbit/s for inside-out)	~20-35 Mbit/s for 4K video ~50-90 Mbit/s for 8K video
Minimum guaranteed data rate DL*	50Mbit/s (10Mbit/s for inside-out)	Same as above for 4K / 8K ~5-8 Mbit/s for Full-HD
Peak data rate UL*	50Mbit/s	At least ~5-8 Mbit/s to enable Full-HD for UGC. Live Outside Broadcast production might require 4K resolution (~20-35 Mbit/s)
Minimum guaranteed data rate UL*	–	–
Connection density*	200 - 3000 /km^2	For e.g. 80,000 spectators in a stadium it would be 1.6m/km² (1 device per spect.)
Traffic density*	Bursty users, NRT video, bursty applications to 3GPP TS36.814, RT video with inter-arrival time 35ms. Packet lengths: bursty users 20Mbyte, NRT video 5Mbyte, bursty application 2Mbyte file size. RT video 3 - 6kbytes uniformly distributed file size. Traffic density on DL: 750Gbit/s/km^2, UL: 125Gbit/s/km^2	–
Radio latency*	50ms - 5s	For live TV viewing: < 5s For gaming: 50ms E2E For live Outside Broadcast production: 1s
End to end latency**	10ms	–
Mobility*	80% with zero mobility (or 3km/h within buildings), 20% with 60km/h andup to 80km/h for inside out. [METIS Dense urban uses 3km/h for small cell and 30km/h for macro cell]	Not specified
Availability***	99%	99.999%

Energy consumption	15 watts maximum for residential base-station	*Not specified*
Fairness	–	*Not specified*
BLER	1 in 10^6 after MAC processes	*This should be the max. BLER for video transmission*
Spectrum efficiency	DL: 2.6bits/s/Hz average across cell with 0.25bits/s/Hz at cell edge. UL: 2bits/s/Hz average across cell with 0.1bits/s/Hz at cell edge	*The requirement from M&E is more of a regulatory / network frequency planning nature*
Energy efficiency	–	*Not specified*
Cost efficiency	–	*Requirement from M&E is too unspecific to quantify*
Misc.	–	Must operate indoor, deep indoor, portable outdoor and mobile

Table 10. Broadband Wireless

KPIs	General requirements from network operator / vendor estimates for healthcare, some IoT applications	Example requirements from vertical health sector
Peak data rate DL*	300Mbit/s	Multiple video streams, 300 Mbit/s in case of remote monitoring and care
Minimum guaranteed data rate DL*	–	10 Mbit/s
Peak data rate UL*	300Mbit/s	Multiple video streams, 300 Mbit/s in case of remote monitoring and care
Minimum guaranteed data rate UL*	–	10 Mbit/s
Connection density*	$100/m^2$	varies a lot, in a hospital there might be hotspots of $100/m^2$
Traffic density*	$>10Tbit/s/km^2$	$>10Tbit/s/km^2$ in hotspots, such as hospitals and care

		centres
Radio latency*	–	do not really care, what matters is the e2e
End to end latency**	100ms< but <1s	This is the generic case, for robotic operation it is much tighter, but there the robot is wired with fibre
Mobility*	up to 500km/h typically up to 250 km/h	up to 500km/h in case of intervention in a helicopter, otherwise typically up to 250 km/h in high speed vehicles
Availability***	99,999%	in case of robotics it is 99,99999%
Energy consumption	Varies	implants need to survive as long as the treatment lasts, up to several years, but these are low end devices, few recordings a day
Fairness	Congestion and interference requires managing	–
BLER	1 in 10^6 after MAC processes	for video
Spectrum efficiency	–	Should come out of the high density
Energy efficiency	Batteries to last tens of years.	–
Cost efficiency	–	–
Misc.	- very deep indoor, rural areas, reaching >99.9% of population, - assured integrity of information inside the channel	–

Table 11. Ultra Reliable Communications

KPIs	General requirements from network operator / vendor estimates for automotive and train	Example requirements from automotive sector
Peak data rate DL*	–	See-Through: 10 Mbit/s Bird's Eye View: 40 Mbit/s Information society on the road: tens of Mbit/s Uploading sensor data: up to 100 Mbit/s
Minimum guaranteed data rate DL*	50kbit/s to 40Mbit/s[9].	
Peak data rate UL*	–	Scenario: Offered load (UL/DL) (Mbit/s/vehicle) (average/peak) URBAN: 1.0 / 10 SUBURBAN: 0.5 / 10 HIGHWAY: 1.0 / 10
Minimum guaranteed data rate UL*	A few bits/s to 25Mbit/s[10].	
Connection density*	2000 / km	Scenario: Vehicle density URBAN: 1000-3000 /km^2 SUBURBAN: 500-1000 /km^2 HIGHWAY: 100-500 /km^2
Traffic density*	V2I: [DL: 100 Gbit/s / km^2 (25 Gbit/s per train, 50 Mbps per car) UL: 50 Gbit/s / km^2 (12.5 Gbit/s per train, 25 Mbps per car)] V2V: Potentially high	–
Radio latency*	V2I: 10ms [NGMN Use Case No6] V2V: 1ms [NGMN Use Case No12]	–
End to end latency**	5 - 100ms[11]	Automated Overtake: 10 ms

---

[9] Vehicle to Infrastructure (V2I): 50Mbit/s, Vehicle to Vehicle (V2V): 50kbit/s to 10Mbit/s . From 5G automotive vision white paper: V2X 10, 40Mbit/s for see-through and bird's eye view. Connectivity up to 20Mbit/s with other vehicles, up to a few Mbit/s to back end servers.

[10] Vehicle to Infrastructure (V2I): 25Mbit/s, Vehicle to Vehicle (V2V): few bit/s to 10Mbit/s

		Coop. Collision Avoidance: 1-5 ms High Density Platooning: 10 ms See-Through: 50 ms Bird's Eye View: 50 ms Not critical for digitalisation of transport and logistics use case
Mobility*	On demand, <140km/hr for cars, <300km/hr [500km/hr] for trains	Scenario: Relative velocity URBAN: 0-100 km/h SUBURBAN: 0-200 km/h HIGHWAY: 0-500 km/h
Availability***	Reliability 10^{-5} for automated overtake, cooperative collision avoidance, high density platooning	Reliability: 10^{-5}
Energy consumption	–	–
Fairness	–	–
BLER	–	–
Spectrum efficiency	–	–
Energy efficiency	DL: 4.4 bits/s/Hz maximum UL: 2 bits/s/Hz maximum. Mapping based on 3GPP 36.942 (Annex A, Page 97)	–
Energy efficiency	–	15 years battery life and 30 dB coverage extension comp. to LTE Rel-12 (T&L)
Cost efficiency	–	–

---

[11] Automated overtake 10ms, cooperative collision avoidance 1 - 5ms, High density platooning 10ms, See-through 50ms, bird's eye view 50ms. Negotiation with other road users 10ms, Information to other road users 10 - 50ms, Exchange with other road users 10ms, Sensors of other road users 10 - 50ms, Comms to server 100ms

Misc.	Positioning accuracy 10 - 30cm[12]	Positioning accuracy: 10..30 cm for V2X use cases good coverage levels along the road infrastructure (also in low-density areas) must operate in challenging reception scenarios such as underground parking (T&L)

Table 12. High speed mobility

Table 13 below contains the summary of the KPIs and their range of values, as derived from the above tables.

KPIs	Range of values	Comments
Peak data rate DL	10 – 300Mbit/s	Driven by video and broadband
Minimum guaranteed data rate DL	1Kbit/s to 50Mbit/s	Low limit is IoT even in basements, high limit is power grid and broadband
Peak data rate UL	10 – 300Mbit/s	The same as DL. Driven by video, the high limit is required by healthcare
Minimum guaranteed data rate UL	10Kbit/s to 25Mbit/s	Wearables, sensors and automated home. The high limit is required by V2I
Connection density	$200 – 3000/km^2$ for most applications, $200K/km^2$ for massive IoT and up to $100/m^2$ for hospitals	Hot spots can be 80,000 in stadium and there may be a density of $100/m^2$ in some hospital areas
Traffic density	$100Gbit/s/km^2$ to $10Tbit/s/km^2$	Lower limit is broadband suburban and trains, and upper limit is hospitals and healthcare centres. UL and DL are symmetrical
Radio latency	10ms to 100ms	Lower limit is for D2D for live voice links and car to car.

---

[12] 30cm for automated overtake, cooperative collision avoidance and high density platooning, 10cm for vulnerable road user discovery

End to end latency	5ms to 1 second	Lower limit is set by automotive, gaming requires 50ms. Higher limit is broadcasting and NRT video. Value contradicts the radio latency KPI.
Mobility	Zero, limited (3km/hr), urban to 80km/hr, vehicle to 120km/hr, trains and helicopters to 500km/hr	–
Traffic model	Packet sizes 125 bytes to 20Mbyte Generation intervals 20ms to 1 hour File sizes up to 1Gbyte	Small packets over long intervals from IoT, larger packets for video. Model also depends on mobility in IoT case.
Availability	Between 99.999% and 99.99999%	The high end is set by robotics
Reliability	10^{-5}	Automotive
Energy consumption	Batteries to last tens of years 15 watts maximum for residential base-station	Highest requirement by implants in eHealth
Fairness	No quantitative requirement	Must be better defined by the vertical
BLER	10^{-6} after MAC processes	For video
Misc.	Position accuracy to within 10 - 30 cm for vehicles Over the air programming to IoT Very deep indoor coverage.	–
Spectrum efficiency	4.4 bits/s/Hz minimum DL 2 bits/s/Hz minimum UL (based on 3GPP 36.942).	–
Energy efficiency	Small indoor cells to be 15w max. IoT to be low enough for several years battery life. Medical implants to last a treatment duration which can be several years.	–
Cost efficiency	No quantitative requirement	Must be better defined by the vertical

Table 13. Summary of KPIs

From Table 13 we can observe the following:

- UL and DL are symmetrical. This is actually following the trend from 2G / 3G / 4G and also on fixed broadband, where the uplink is increasingly becoming of the same order as the downlink
- Two of the KPIs, Fairness and Cost Efficiency, do not have quantitative measures as there have been no specific requirements from the vertical sectors, and thus SPEED-5G will need to define and develop these
- Traffic and connection density have in-building hotspots
- Low latency links tend to be short distance

The high level requirements on the MAC layer are:

- High reliability when needed, and especially on the uplink (e.g. for low bit rates from terminals deep inside buildings), using all variations of licensed spectrum. This applies to link robustness as well as availability
- Bit rates between 1Kbit/s and 300Mbit/s
- Packet sizes from 125 bytes up to 20 Mb with repetition periods of 20ms up to 1 hour
- To provide symmetrical uplink and downlink including with TDD spectrum
- To provide low latency when needed, down to 1ms for localised links
- To provide mobility support up to 500km/hr depending on use case
- Queuing and scheduling for multiple QoS streams on uplink and downlink

The RRM algorithms can be evaluated using a system utility function, which will address a number of KPIs listed in Table 9 which are relevant to the lower layers. The utility function can for example maximise throughput or spectral efficiency, cost efficiency, and energy efficiency across the system.

2 State of the art (SoTA) on RRM

2.1   Other 5G-PPP projects

FANTASTIC-5G has similar objectives as SPEED-5G. The project takes the approach of evaluating the performance of components using simulation. Its ambition is to deliver an accurate scenario for full-system modelling and plans to use the Bari LTE simulator. The high-level system diagram provided by FANTASTIC-5G, introduces an "advanced RRM" block without further details. The PHY and MAC characteristics are approximated by pre-computed curves.

Flex5Gware focusses on reconfigurable network hardware and software elements. It introduces a function called "RAT flexibility and node selection", but without further detailed description.

SELFNET focusses on autonomic network management and uses the scenarios self-protection, self-optimisation and self-healing to demonstrate its achievements. It works almost exclusively at the virtualisation layer. It does not address the MAC or PHY layers.

Xhaul address the integration of fronthaul and backhaul in the 5G context. It does not address the MAC or PHY layers

mmMAGIC (Millimetre-Wave Based Mobile Radio Access Network for Fifth Generation Integrated Communications) develops and designs new concepts for mobile radio access technology (RAT) for mm-wave band deployment, without addressing currently areas of interest to SPEED-5G.

At the time of the review, these projects were in the start-up phase, and none of them had developed specific ideas of relevance to SPEED-5G. Future review of the projects' work will be focused on the following top-level questions:

1.   Into how many layers is the RRM placed?

2.   At the highest layer, how many cells does the RRM have sight of?

3.   At the lower layers, how autonomous are the RRM elements?

4.   Is information sharing done by push or pull?

2.2   3GPP

According to the 3rd Generation Partnership Project (3GPP), the goal of the Radio Resource Management (RRM) is to realise the efficient usage of the available radio resources, in both single and multi-cell environments, and to provide solutions that enable the Evolved Universal Terrestrial Radio Access Network (E-UTRAN) to reach the target requirements [14]. In the technical report 36.300, 3GPP section 16 defines the functional blocks included in the RRM, which are Radio Bearer Control, Radio Admission Control, Connection Mobility Control, Dynamic Resource Allocation, Inter-cell interference co-ordination, Load Balancing, Inter-RAT RRM and Subscriber Profile ID.

In this section, we recall the goal of each of these blocks paying particular attention to those functions that are of interest to SPEED-5G (related references [14], [64], [59], [1] and [13]). In Speed-5G we consider ICIC to be a self-organising network (SON) function rather than an RRM function and so this is not included. The relationship between SON and RRM is that they operate side by side, but that RRM can steer the SON by setting or fine-tuning parameters and thresholds.

The report 36.300 also makes reference to RRM architecture in section 17, but only considers distributed RRM.

2.2.1   3GPP RRM functions

Radio Bearer Control (RBC)

The RBC is a function located in the eNB that is in charge of establish, manage and release the radio bearers, and to configure the associated radio resources, by taking into account the overall available resources and the QoS requirements of the different services provided by the eNB.

Radio Admission Control (RAC)

The RAC is a function located in the eNB that has the role to accept or reject the requests for new radio bearers. Its objective is to provide high radio resource utilisation efficiency by accepting requests as long as radio resources are available without affecting the QoS of on-going sessions.

Connection Mobility Control (CMC)

CMC is a function located in the eNB that manages radio resources during user mobility. In particular, it configures reporting and measurements procedures and manages handover based on this information as well as on cell load, traffic distribution, transport and hardware resources, and operator policies.

Dynamic Resource Allocation (DRA) - Packet Scheduling (PS)

The task of these functions is to allocate buffer, processing and radio resources to user and control plane packets. In particular, PS takes into account QoS requirements, channel quality measurements, buffer status, and interference. DRA considers specific resource preferences/restrictions related to inter-cell interference coordination mechanisms.

Load Balancing (LB)

Load balancing is a function located in the eNB that is in charge to handle uneven distribution of the traffic load over multiple cells such that resource utilisation efficiency is maximised. Since LB algorithms may result in handover or cell reselection decisions, it has to avoid negatively impacting the QoS of on-going services.

Inter-RAT Radio Resource Management

Inter-RAT RRM is related to the management of radio resources during inter-RAT mobility. In particular, inter-RAT handover/LB may take into account the involved RATs load as well as UE capabilities and operator policies.

Subscriber Profile ID for RAT/Frequency Priority

The RRM strategy may be based on user specific information (e.g. mobility profile, service usage profile) received by the eNB via the S1 interface or the X2.

2.2.2   3GPP functions of interest to Speed-5G in more detail

The establishment, maintenance and release of radio bearers involve the configuration of radio resources associated with them. When setting up a radio bearer for a service, Radio Bearer Control (RBC) takes into account the overall resource situation in E-UTRAN, the QoS requirements of in-progress sessions and the QoS requirement for the new service. RBC also deals with the maintenance of radio bearers of in-progress sessions at the change of the radio resource situation due to mobility or other reasons. RBC is involved in the release of radio resources associated with radio bearers at session termination, handover or at other occasions. RBC is located in the eNB.

Radio Admission Control (RAC)

The task of Radio Admission Control (RAC) is to admit or reject the establishment requests for new radio bearers. In order to do this, RAC takes into account the overall resource situation in E-UTRAN, the QoS requirements, the priority levels and the provided QoS of in-progress sessions, and the QoS requirement of the new radio bearer request. The goal of RAC is to ensure high radio resource

utilisation (by accepting radio bearer requests as long as radio resources available) and, at the same time, to ensure proper QoS for in-progress sessions (by rejecting radio bearer requests when they cannot be accommodated). RAC is located in the eNB.

Dynamic Resource Allocation (DRA) - Packet Scheduling (PS)

The task of Dynamic Resource Allocation (DRA) or Packet Scheduling (PS) is to allocate and de-allocate resources (including buffering and processing resources and resource blocks (i.e. chunks)) to user and control plane packets. DRA involves several sub-tasks, including the selection of radio bearers whose packets are to be scheduled and managing the necessary resources (e.g. the power levels or the specific resource blocks used). PS typically takes into account the QoS requirements associated with the radio bearers, the channel quality information for UEs, buffer status, interference situation, etc. DRA may also take into account restrictions or preferences on some of the available resource blocks or resource block sets due to inter-cell interference coordination considerations. DRA is located in the eNB.

Load Balancing (LB)

Load Balancing (LB) aims at making efficient use of the limited spectrum to deal with unequal loads in order to improve network reliability by reducing the congestion probability in hot spot areas of cellular networks. Load Balancing is one of the key use cases in SON [23]. Four standalone load balance policies (transmit power adjustment, antenna parameters adjustment, cell reselection, and handover parameters adjustment) are proposed in [109] and [113]. Load-based handover and cell reselection optimisation are also proposed by 3GPP [23] and NGMN [9].

Inter-RAT RRM

Inter-RAT Radio Resource Management (IRRRM) is primarily concerned with the management of radio resources in connection with inter-RAT mobility, notably inter-RAT handover. At inter-RAT handover, the handover decision may take into account the involved RATs resource situation as well as UE capabilities and operator policies. Inter-RAT RRM may also include functionality for inter-RAT load balancing for idle and connected mode UEs.

The main drivers for inter-RAT radio resource management are captured in 3GPP TS 36.300. For idle mode, the methods and parameters are specified in 3GPP TS 36.304. The general working assumption in the 3GPP, at the time of writing, is that idle mode inter-RAT management is based on absolute priorities. For connected mode, the details of inter-RAT handover managements will be covered by 3GPP TS 36.331.

Inter-RAT Carrier Aggregation

Carrier Aggregation (CA) is functionality introduced in LTE Release 10. Its purpose is to make up to 100MHz bandwidth available for the downlink users. This is achieved through the aggregation of portions of available spectrum bands, called Component Carriers (CC). Several flavours of CA are available: intra-band contiguous CA, intra-band non-contiguous CA and inter-band CA. This basic notion of CA will be extended in 5G with the introduction of Multi-site Carrier Aggregation and Heterogeneous Radio Access Technologies (RAT) Carrier Aggregation, also called inter-RAT Carrier Aggregation.

Multi-site CA denotes a solution where user equipment (UE) is capable of aggregating CCs transmitted from different base stations. Inter-RAT CA, instead, denotes the possibility of combining multiple different technologies to enable network access. The integrated use of multiple RATs will be essential to address the challenges and meet the demands of future 5G networks. As an example, the use of unlicensed WiFi bands can bring important gains in terms of increased networks capacity, improved user throughput and reduced interference to cellular networks. The focus of this section is Inter-RAT CA.

One fundamental question related to the design of a 5G air interface is how the different radio access technologies can be integrated (or aggregated) such that the complexity of the implementation is minimised and the performance of individual RAT is not sacrificed. This raises

further questions related to protocol harmonisation and on which layer aggregation should take place. Different approaches imply different benefits and challenges.

However, there seems to be a general consensus that a full harmonisation between different air interface variants may not be possible at the physical layer due to the different requirements and delays of the different baseband processing chains. Therefore, harmonisation should happen on MAC layer and above. An example of such approach is the multi-MAC aggregation presented in [114], where a new module introduced in the stack right above the air interface permits to aggregate resources at the MAC layer. This module acts as the main responsible to route IP packets across different RATs.

Additional challenges are related to the design and the definition of mechanisms to coordinate the different available resources optimally. Indeed, the definition of efficient Radio Resource Management (RRM) algorithms capable of addressing these challenges represents an important part of the research work conducted within the SPEED-5G project.

The current research work in the context of LTE-Advanced is briefly reviewed here. The functionality of Dual Connectivity (DC) has been standardised starting from LTE Release 12, in order to enable carrier aggregation of different base stations [8]. With the beginning of the work toward the definition of LTE Release 13, a new study item called Licensed-Assisted Access has been initiated [6], with the objective of investigating carrier aggregation of licensed and unlicensed bands within one base station. Also in LTE Release 13, another work item called LTE-WLAN (LWA) aggregation has been established, with the objective of providing aggregation of LTE with WiFi at the Radio Access Network (RAN) level.

Indeed, the aggregation of LTE and WiFi is one of the most important use cases that is currently considered, due to widespread availability of WiFi and the fact that it operates over unlicensed bands. One specific study group in 3GPP is devoted to studying the coordination of multiple RATs and the most recent outcomes can be found in [7]. One important result of this study group is the introduction in the architecture of a newly defined interface, denoted Xw, between the base station (eNB) and the WiFi access point (called WLAN Termination Point or WT in short).
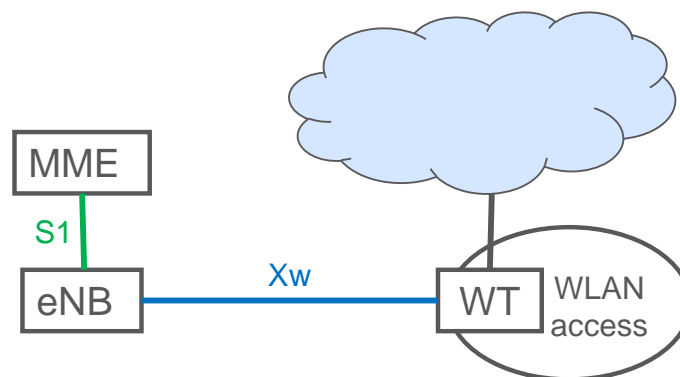
Figure 2. Terminating Xw between the eNB and the WLAN Termination, from [7].

The problem of the optimal resource allocation in multi-RAT heterogeneous networks has been addressed in literature in the past years. Solutions include both centralised and distributed algorithms. Examples of distributed algorithms that aim at maximizing the capacity of multi-RAT systems through the optimisation of the resource allocation are proposed is [127] and [74]. Recently, the topic of energy efficiency has attracted more interest, since it will play an important role in next-generation cellular networks. Therefore, several works have addressed this topic, such as [45] for the downlink and [46] for the uplink.

In conclusion, we can say that the problem of finding the optimal resource allocation in the case of a multi-RAT system composed of multi-tier heterogeneous networks subject to dynamically changing interference is a challenging task. The amount of feedback and coordination between users and cells, and among cells themselves is often limited, and algorithms are normally very susceptible to

outdated information. Additionally, the optimal multi-RAT resource allocation should take into account current interference variation.

2.3   Review of RRM work related to device to device communications

Devices are communicating with each other in D2D fashion without intervening by any mediator nodes. D2D communication utilises cellular spectrum (license band) that is aided by an infrastructure of cellular network and anticipates ternary kinds of benefit: (a) User Equipment (UEs) proximity may provide tremendously high bit rates, minimal delays, and high energy efficiency for consuming energy [29], [41]; (b) the reuse benefit possibly entails radio resources being applied by cellular system as well as D2D links concurrently; reducing the reuse factor so that the same spectral resource can be used more than once within the same cell [48], [61]; (c) finally, there is a gain from not having to use resource of a downlink (DL) and an uplink (UL) simultaneously, as is the scenario when conveying information in the cellular system through the access point. Furthermore, new categories of peer-to-peer wireless services are facilitated and cellular coverage might be expanded by D2D systems. A D2D system is also economical communication because it uses the same pre-existing cellular infrastructure which increases network efficiency. This increased network efficiency supports more services and improves current services and applications.

In the past, cellular operators did not consider D2D communication as a method to enhance the performance of cellular network because the effect of D2D communication is limited to local communication services. However, as mobile applications based on proximity of mobile devices are becoming increasingly popular, cellular operators are considering introducing D2D communication into the cellular networks. When operators enable D2D mode in their cellular infrastructure, they will be able to come across various gains in contrast to the typical infrastructure-based system; e.g. increased system throughput of whole network, enhanced energy efficiency and reduced network traffic load.

D2D communications will be a corner stone in future 5G networks, as manifested by the WiFi Direct specifications and proposals for LTE-A D2D standardisation, with exploitation in applications like emergency services. LTE-A, Qualcomm and IEEE 802.15.4g (SUN) are currently addressing the standardisation of D2D communication over licensed band. In June 2011, a 3GPP study item description on the radio aspects of D2D discovery and communication was submitted by Qualcomm. Meanwhile, a study item description on LTE Direct (LTE-D) was submitted to the 3GPP meeting held in August 2011, which proposed the study of the service requirement for direct over-the-air LTE D2D discovery and communication. Towards the finalisation of LTE Relelase11, 3GPP initiated the agenda for Release12 and beyond which started as a workshop in June 2012 [92], [117]. The study on ProSe includes two parts, namely D2D discovery and D2D communication. The main results on D2D use cases and potential requirements are captured in [5] and needed architectural enhancements to support ProSe in [124], [62]. The current work on LTE D2D device discovery and D2D communication mainly focuses on the technical details including discovery signal design, resource allocation and scheduling, synchronisation mechanism, channel models and D2D evaluation methodology. Complementing 3GPP, the IEEE 802.11 Infrastructure Service Discovery Study Group have done much work on proximate discovery and communication with low energy, long range (up to 500 metres) and large scale (up to 1000 mobile devices) for mobile social networks since 2010. Simultaneously, efforts in IEEE 802.11s, 802.11ac, 802.11ah may make D2D possible and attractive. IEEE 802.15.8 Peer Aware Communication Task Group defines the physical and medium access control (PHY and MAC) layers specifications and optimisations for infrastructure-less communications with fully distributed coordination in May 2012. Additionally, the D2D feature is also included in IEEE 802.16n.

D2D still has a lot of interesting research questions which need to be addressed before it is completely standardised. The current work on LTE D2D device discovery and D2D communication mainly focuses on the technical details, including discovery signal design, resource allocation and scheduling, synchronisation mechanism, etc. There are also several recent papers which address the multi-hop D2D communications as an underlay of an LTE-A network [123]. The feasibility and the

range of D2D communication and its impact on the power margins of cellular communications are studied in [115], [60] and [57] and references therein. For a scenario with D2D multicast communication integrated into a cellular network, a clustering concept is introduced in [110] and references therein. The work in [78] and [132] consider the case where D2D UEs use exclusive radio resource from cellular connected UEs and the same resource of cellular connected UEs. Many studies considered that D2D UEs reuse the radio resources of the eNB relaying system and tried to solve the interference problem caused by D2D communication to improve the scalability of D2D communications, distributed resource allocation methods were proposed in [53]. The authors in [33] proposed a mechanism for D2D communication session setup and management along with the interference coordination between two subsystems in the LTE system architecture evolution. A related coexistence between human-to-human and machine-to-machine communications in LTE-A is studied in [66]. In the D2D context, pairing refers to selecting the D2D pair(s) and at most one cellular UE that share (reuse) the same OFDM resource block, similarly to multiuser MIMO techniques [125], [70]. Pairing is a key technique to achieve high reuse gains [67], [80]. In [50], the authors explained the scheduling for D2D communication considering dynamic TDD slot structure for D2D.

2.4   Other RRM state of the art

According to [10], RRM requires contextual information, and this information is listed as SINR of wanted signal, system capabilities like edge of cell rate, bandwidth and availability rules, supported air interfaces and transmission modes, a list of operators and the traffic load per cell.

In [11], the role of the RRM is described as dealing with load balancing and interference management, and they focus on backhaul to small 5G cells using frequencies up to mm-wave.

A HP technical white paper [12] discusses RRM technology in unlicensed spectrum (WiFi), and describes it as having the following functions, performs dynamic channel selection and power control, monitors wireless channel interference, manages wireless signal coverage and balances load among APs.

Energy efficient RRM is discussed in [16] where they utilise fuzzy logic to make decisions on handover management combined with optimising energy efficiency. Choice of RATs is made according to a priority that is chosen from a number of criteria including QoS requirement and traffic class for which they use 3GPP classes.

2.5   Gaps in RRM SoTA and focus of Speed-5G

RRM SoTA assumes a distributed architecture, where the RRM functions reside in every Access Point or eNB. Speed-5G RRM is proposed to be mainly centralised, with certain functions devolved to base-stations when more efficient, a potential candidate for this is load balancing. A centralised RRM (cRRM) will be designed to influence a cluster of hundreds of cells, and in addition will provide inter-cRRM co-ordination to manage 'seams' in the coverage between clusters.

There is no coupling at present to higher layer entities, such as a spectrum manager and KPI collector, so that Speed-5G will introduce this concept. The notion is that one or more cRRMs will co-ordinate with a spectrum manager that contains a list of the spectrum bands in the portfolio and also the conditions and policies for its use. Such conditions and policies are those imposed by the national regulator (such as transmit power and duty cycle) and also by the network operator that owns the portfolio.

The SoTA includes the concept of KPIs to make decisions on spectrum and RAT selection, but Speed-5G will take this a step further by monitoring KPIs for each logical session, and re-balancing the radio resource to keep them within a given range.

3 The SPEED-5G approach in terms of RRM and its scope

3.1 Enhanced Dynamic Spectrum Access

Enhanced Dynamic Spectrum Access (eDSA) is an innovation of SPEED-5G, the aim of which is to increase the efficiency in the use of radio resources, especially spectrum but also transmit energy, while providing the high levels of user experience which have been defined as use cases and KPIs in section 1. SPEED-5G works with spectrum below 6GHz, and it is in these bands that harmonised spectrum is in particularly short supply. Pressure on the spectrum below 6GHz is expected to further increase with time, so that systems must be increasingly smart about how they use it.

eDSA consists of two major parts, which are the Radio Resource Manager (RRM) and the MAC layer of the base stations. The RRM is a centralised function, and in this respect it extends the work of 3GPP TS36.300 [68] which, although it explains the basic functions of RRM in release 13, it does not cover centralised RRM architecture. The SPEED-5G high level concept of eDSA and the mapping onto work packages is shown in Figure 3.
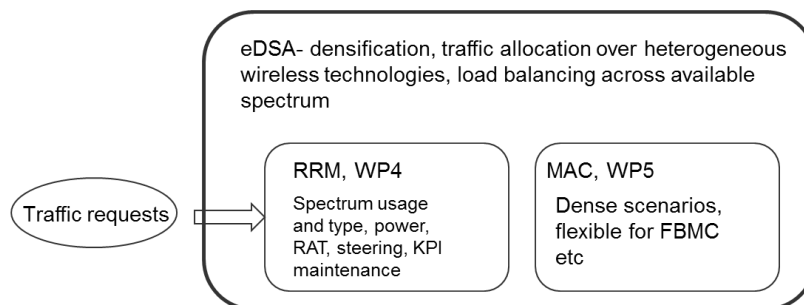
Figure 3. High level diagram of eDSA and mapping onto work packages

The following high-level requirements have been captured from the selected use cases, driving the overall design of the cRRM framework in SPEED-5G:

ID	(Level 1) L1-requirements
R1	The cRRM design shall support admission and prioritisation of traffic, for steering to base-stations and session flows
R2	The cRRM design shall support load-balancing across different RATs, across different licensing regimes and taking into account interference
R3	The cRRM design shall support spectrum / RAT resource aggregation and antenna selection, taking into account base-station and UE capabilities
R4	The cRRM design shall support channel selection, working on finer granularity than R3 to achieve better spectrum utility, delegated to MAC in the case of unlicensed spectrum
R5	The cRRM design shall support inter-RAT coordination, improving utility when different RATs are used in the same band

Table 14: Requirements list

The high-level functions of the RRM and the MAC and the interactions between them, and also between the RRM and some other system components, are shown in Figure 4. The functions may change or be elaborated as the modelling work proceeds. A brief description of the components is given below.

The Spectrum Manager is essentially a database that keeps track of the spectrum that is available to the network operator, and also the policies and restrictions on the use of the spectrum. It contains information on all types of spectrum, licensed, lightly-licensed and licence-exempt. In the case of licensed spectrum, it contains the licence conditions such as channel width, wanted and unwanted emissions masks, transmit powers and duty cycle, access mode (FDD/TDD) and any restrictions on technology type or dependency on environment (such as location accuracy exemptions for indoor

use of small cells). In addition, the conditions of use of licensed spectrum may be subject to harmonised European standards. In the case of the other two categories of spectrum, lightly-licensed and licence-exempt, the conditions are likely to be less onerous and subject only to harmonised European standards but there may also be certain national restrictions. There is only one Spectrum Manager per operator, whereas there is likely to be several RRMs, with each RRM being responsible for the supervision of hundreds or possibly thousands of cells.

The KPI block keeps a record of the target KPIs and also keeps track of current performance in terms of the KPIs.
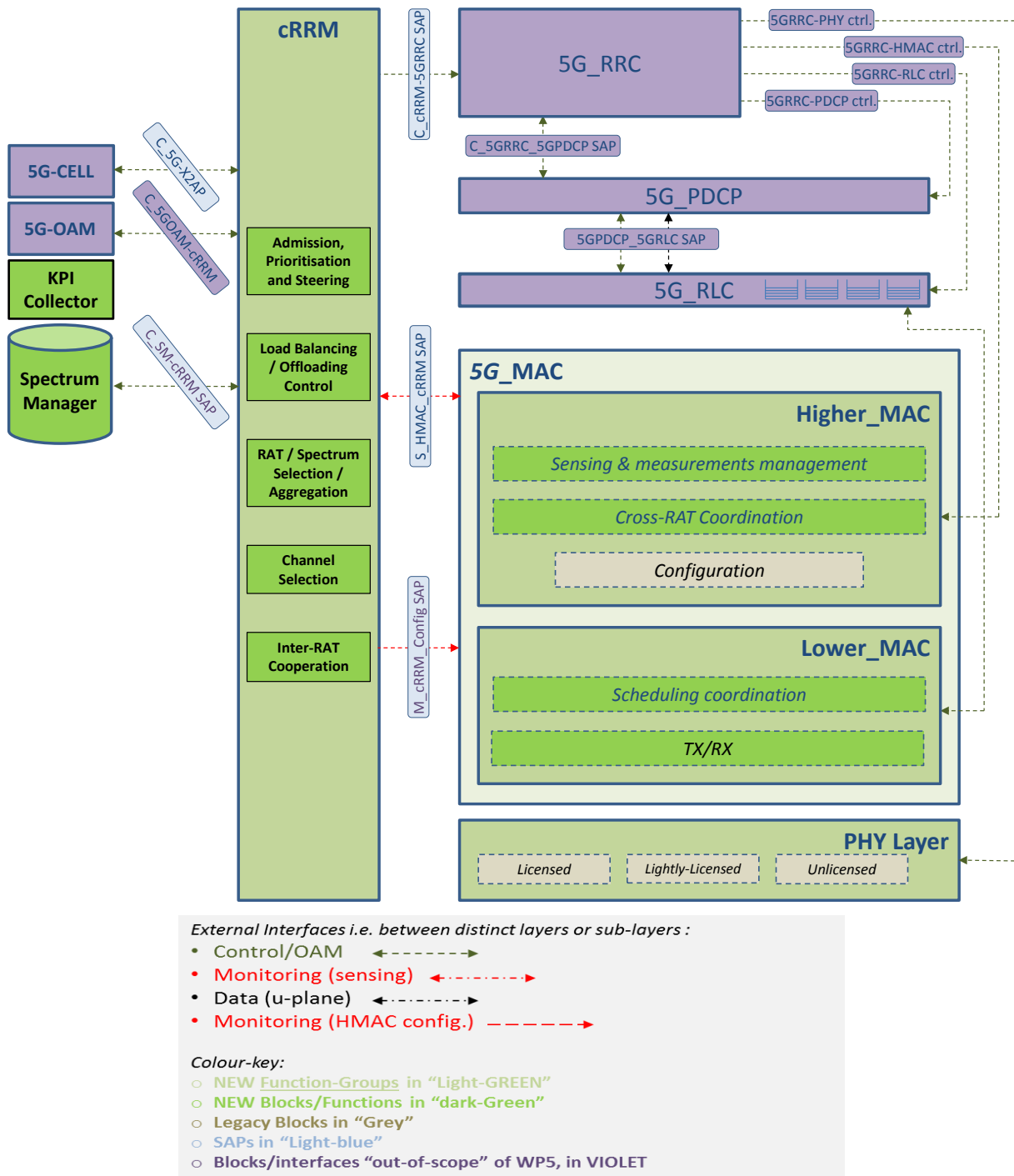

Figure 4. RRM and MAC high level functions

The Operations and Management (OAM) entity shown in the figure illustrates that the RRM must be connected to the OAM in order to access the user terminal registers for the purpose of learning the terminal abilities, such as what spectrum and RATs they are capable of handling. The RRM must also

know, as far as possible, the network topology. This would typically include the neighbour lists from the base stations, detected from REM scans and also from UE reports, and the base station locations, where available. The locations will be typically available for the macro and larger outdoor base stations but not for plug-and-play indoor cells.

The MAC functions will not be discussed in this report as they are the subject of WP5, but we need to clarify one or two of the MAC roles in order to make its interaction with the RRM clear. Looking at Figure 4, we see that the MAC is responsible for choosing the channel if the spectrum is licence-exempt (unlicensed), but the RRM is responsible for choosing the channel if the spectrum is licensed. The reason for this split is that the channel selection on unlicensed spectrum is likely to change in the short term as the interference level varies. And this is a good time to point out that the MAC works on a shorter timescale than the RRM, and this in turn works on a shorter timescale that the Spectrum Manager. The 'TIME' arrow on the left-hand side of the figure indicates that the functions higher up in the figure operate over longer timescales. It can be estimated that the Spectrum Manager would operate over several hours, the RRM over seconds or minutes, and the MAC at the millisecond (frame-length) timescale.

In Figure 4, the link between the RRM and MAC is bi-directional, and the towards-RRM direction provides feedback regarding channel monitoring. This is needed because the interference level will affect the RRM decision on the channel and channel aggregation patterns, when dealing with licensed spectrum. The MAC layer is responsible for allocating resource blocks, or their equivalent in any new 5G RATs, and this can include ICIC constraints where cooperation is needed between neighbouring base stations on co-channel. It can also include cooperation when CoMP or cooperative beamforming is used. These functions are typically managed using distributed SON functions, which aim to improve the performance of cell-edge users. There are also the centralised SON functions, which deal more with the configuration automation, such as PCI selection and downlink transmit power control, but here we are interested in the distributed SON. It is proposed that the RRM will work with distributed SON and can steer the SON mechanisms to further optimise features like ICIC and CoMP.

For unlicensed spectrum, the RRM will specify the frequency band, such as 2.4 or 5GHz, and it may also need to steer the MAC layer to a channel in certain situations, despite the earlier statement that the MAC is free to choose the unlicensed channel. For example, if LAA can be used from multiple base stations, a degree of coordination will be needed.

The eDSA architecture may well develop further as the project progresses, for example mapping to a practical network architecture that involves fronthaul and backhaul.

The SAPs/interfaces that are within the WP4 scope and will be specified in more detail in the next deliverable (e.g. description, primitives, etc.) are:

- o **M_HMAC_cRRM SAP** (monitoring interface between cRRM and HMAC [sensing & measurements management])
- o **M_cRRM_Config SAP** (monitoring interface between cRRM and HMAC [Configuration])
- o **C_SM-cRRM SAP** (control interface between cRRM and Spectrum Manager)
- o **C_cRRM_5GRRC SAP** (control interface between cRRM and 5G-RRC)
- o **C_5G-X2AP** (control interface between cRRM of once cell and cRRM of neighbour/another cell)

3.2   RRM Functions

In the following, a list of important functions that were identified for the SPEED-5G RRM is provided.

3.2.1   Admission / prioritisation / steering

This block makes a decision about whether to admit a new traffic flow, what priority level it should

have, and which base-station or base-stations to steer it to. The outputs will be to label the different types of traffic so that the MAC can steer to the appropriate RAT, and to keep a mapping between types of traffic and the available bands. The different types of traffic are labelled based on one or more criteria. Some pre-determined association rules could be established, depending on the requirements of each specific traffic type and depending on what spectrum bands or RATs are available.

For instance, traffic with the most stringent QoS requirements may be allocated only on licensed or lightly-licensed spectrum, which provide ways to ensure the provision of some end-to-end quality indicators. On the other hand, traffic with non-stringent error and latency requirements may be moved to unlicensed bands.

Additional criteria might be included in the traffic steering besides QoS and error requirements, such as the expected data rate or the expected coverage provided by a specific RAT or band. Low frequency bands might be more suitable for low-rate applications that require wide coverage areas. This is the typical use case for IoT-like type of traffic and low-end sensors operating indoor in factories that do not require high reliability. On the other hand, high-rate applications may not fit with low frequency bands.

3.2.2   Load balancing – offloading control

Load balancing (LB) aims at making efficient use of the limited spectrum to deal with unequal loads in order to improve network reliability by reducing the congestion probability in hot spot areas of cellular networks. Traditionally, in 3GPP, the focus of the LB function has been to handle uneven distribution of the traffic load over multiple cells such that resource utilisation efficiency is maximised. LB algorithms may therefore trigger handover or cell reselection procedures.

Since the small cells envisioned by SPEED-5G are able to handle and transmit traffic using different RATs and different spectrum bands, this functional block in SPEED-5G adds the capability to trigger the offload of traffic to unlicensed bands or to less loaded bands.

For instance, the traffic offload may be triggered if some level of acceptability cannot be provided anymore, e.g. when the QoS/QoE degrades excessively. This might be due to a too high interference experienced in the licensed bands or some other cause.

3.2.3   RAT/spectrum selection and aggregation

The objective of this functional block is to select a suitable band and RAT to be used by each type of traffic. It also selects the number of channels to be used within a band, if needed. This functional block determines the system spectral efficiency and therefore how much aggregation is needed. Additionally, it may suggest, for each band, the operational configuration per each band, for instance the maximum available bandwidth or maximum transmit power, among others. If needed, it may also suggest a MAC Frame configuration. This functional block takes its decisions based also on specific regulations inherent to each band.

3.2.4   Inter-RAT cooperation

This is an augmenting functionality of the RRM entity of SPEED-5G. It is augmenting in the sense that its operation is not fundamental for the correct operation of the entire system, but it can provide additional benefits.

Specifically, this block aims at improving the coexistence with other RATs in the same band. As an example, in the 5GHz unlicensed band where the data transmission must coexist with WiFi, this block may manage the transmission and reception of the Request-To-Send and Clear-To-Send (RTS/CTS) control messages as well as the Network Allocation Vector (NAV), which is used in WiFi as a form of virtual sensing.

3.2.5   Channel selection

The process of selecting the best available channel for transmission can be performed either in RRM or MAC layer, depending on the time scale and on the available information. It can work at a finer level of granularity than the spectrum selection. The dual blocks may interact with each other. For instance, in the case of autonomous small cell operation, channel selection may be performed autonomously in the MAC layer, and the results may be optionally verified by the RRM.

3.2.6   List of inputs/outputs of the identified RRM functions

Table 15 contains the list of functions and high level inputs and outputs

Functions/Functional Blocks	Inputs	Outputs	Comments
Admission / prioritisation / traffic steering	Traffic type(s), regulations, QoS requirement per traffic type	Suitable Band/RAT per traffic type	
Load Balancing- Offloading Control	Load, Interference level, SINR/CQI, KPIS	Decision whether performing traffic offloading or load balancing.	
RAT/Spectrum Selection/Aggregation	Set of available bands and RATs, KPIs, Set of traffic type(s), Load estimates, (Interference)	Mapping between traffic types and resources, RATs to be used for transmission, number of channels within band, potentially the details of operation per RAT/band	
Inter-RAT Cooperation	Set of Bands/RATs in use, Regulations per RAT/Band	Measures to ensure coexistence with other systems, depending on regulations and RATs in use	Incremental function
Channel Selection	Set of available bands and RATs, KPIs, Set of traffic type(s), Load estimates, (Interference)	The most suitable channel to be used within the selected band at the specific time	Can be at MAC layer for unlicensed spectrum

Table 15. RRM functions and functional blocks

3.2.7   Other functions external to RRM

The spectrum manager keeps a portfolio of available spectrum and conditions and policies for its use. It also can keep a record of spectrum quality in a particular area, from feedback on interference conditions. The KPI collector keeps a record of the target and achieved KPIs for each traffic stream. The OAM functionality is the northbound interface to the network operator network management system.

Spectrum sensing can be performed both at the cRRM and at the small cells, operating on a different time scale and collecting the KPIs on different time scales.

3.3   Demonstration use cases

The high level eDSA design just described needs further detail, depending on the use case that will be demonstrated in the SPEED-5G project. These demonstration use cases which are elaborated further below are:

- Load balancing
- Dynamic channel selection, and
- Capacity augmentation in small cells

3.3.1   Load balancing for interference management among a group of small cells

3.3.1.1   Scenario description

Let us consider a group of neighbouring small cells managed by the same operator and sharing the same licensed spectrum. These small cells are conveying a mix of traffic composed of broadband, ultra-reliable communications and IoT. Due to the high traffic load, the dense deployment of small cells and the overlay of the macro cell, which may operate on the same channel, the level of interference may be too high. This leads to an excessive degradation of the QoS/QoE.

A centralised RRM controller (located in the edge virtualised architecture) analyses the load on each small cell, the level of interference on the licensed band (e.g. by means of CQI reports, BLER measurements and/or NACKS). If an acceptability threshold is down-crossed, the controller decides to trigger a load balancing process. Based on the characteristics of the different available shared bands (bandwidth, central frequency, path loss and regulations), the default mapping of traffic types on spectrum resources (depending on the license regime, available bandwidth and regulation limitations) and the estimation of the interference of these bands, part of the data traffic is steered to some non-licensed bands.

In particular, the load balancing function is in charge of selecting the traffic types that can be handed over to a specific band. Furthermore, the RAT/spectrum selection will identify, for each band, the maximum bandwidth, suitable transmit power, RAT and default MAC frame configuration, depending on the regulation constraints (e.g. listen-before-talk maximum period) and the interference level. The control traffic is kept on the licensed spectrum.

Regarding the offloaded traffic, in the MAC layer, a suitable channel is selected for each traffic type and/or for each band. This can be done either by relying on prior available sensing measurements already available or by resorting on the sensing capabilities of the small cells. Additionally, a MAC frame format is configured, using either a default configuration provided by the RRM part or through an adaptive MAC frame configuration.

For bands requiring a listen-before-talk procedure, the MAC layer triggering for initiating the actual transmission of a frame is provided by the CCA function. This functional block receives the PHY measurements about channel occupancy (e.g. based on the energy level) and decides whether the channel is available and can be accessed. If the channel is available, the MAC triggers the scheduler to initiate the process of mapping resources to the frame format, allocating uplink and downlink traffic in physical resource blocks (time and frequency resources). Additionally, based on the PHY measurements (specifically the noise level in the channel), the CCA block may estimate a CQI value that can be provided to the scheduler as an additional updated information about the channel quality.

3.3.1.2   RRM functions

A brief description of the inputs and outputs of the functions identified at RRM is given in the following picture for load balancing.

	Functions	Inputs	Outputs	Comment
RRM	Load balancing / offload / traffic steering	Traffic load, interference level, SNIR / CQI, KPIs	Decision on offloading or balancing. Traffic list for offload to unlicensed spectrum, accompanied by a list of suitable bands based on association rules to Spectrum / RAT selection	Point of entry in the process, where decision is made, based on load, interference level and KPIs
	Spectrum / RAT selection	Traffic load, interference level, SNIR / CQI, KPIs, list of candidate bands	Mapping between traffic items and resource, RATs for offload. Number of channels per band -> MAC channel selection and to MAC frame design	
	Inter-RAT co-operation	RAT to use, regulatory constraints, interference levels	Measures for co-existence with other systems, update expected impact, output to MAC unlicensed controller	Function to enhance performance
Outside RRM	Radio environment sensing	Available bands, regulations per band, sensing configuration, detector type	Usage per RAT / band, interference and noise levels,	

Table 16. Load balancing: functions input/output for RRM

3.3.2   Dynamic channel selection

3.3.2.1   Scenario description

An initial situation is envisaged where multiple small cells are managed by a single network operator. The small cells are conveying a mix of traffic, that corresponds to the traffic in the SPEED-5G use cases that are extreme mobile broadband (xMBB), Ultra-Reliable Communications (URC), massive IoT and high speed mobility. There would be an assumed amount of interference that could be varied, degrading the QoS.

A decision is made by the RRM to select a band, from a combination of licensed and unlicensed spectrum available and possibly also to choose a different RAT depending on the context and the ability of the user terminal. The aims would be

- to meet some specified QoS requirement,
- to allow for cell selection as one mechanism,
- to allow for opportunistic sharing as an option, similar to TV Whitespace, and
- to be backwards compatible with existing LAA and LTE-U standards.

The decision as to which channel (i.e. sub-band) within the band is then delegated to the cell, which, for example, can use the exponential weighted randomised algorithm described below.

3.3.2.2   RRM functions

The RRM functionality shown in Figure 4 is high level and is designed to cope with all three demonstration use cases. It is necessary to refine the function processes for the case of Dynamic Channel Selection. Table 17 below is a list of all the RRM functions in Figure 4, with inputs, outputs and processes customised for this case.

Function name	Inputs	Outputs	Processes
Admission and prioritisation	Traffic scenario and QoS requirements. Target KPIs and operator policy.	Priority levels	Terminal discovery, decision on priority given existing load, available capacity and operator policy.
Load balancing: offloading control	Traffic scenario and its required QoS, operator policy, deployment topology, e.g. from neighbour lists	Split of traffic flows across licensed and unlicensed, required spectrum quality and channel widths	Optimise traffic flows across licensed and unlicensed spectrum
Load balancing: Traffic steering including base stations used in links	Traffic scenario and its required QoS, operator policy, deployment topology, e.g. from neighbour lists	Base stations involved in the sessions and throughput / delay allowances	Decide which base stations are involved in the sessions
Load balancing: Bearer establishment	Base stations involved in the sessions and throughput / delay allowances	Signalling to establish bearers on uplink and / or downlinks, tagging of packets	Establishes sessions or bearers, and determines their bandwidth / delay requirements
RAT/spectrum and channel selection / aggregation	Bandwidth / delay requirements, spectrum portfolio, spectrum quality, spectrum policy	Channel selection, per-channel spectrum (licensed)	Allocates licensed spectrum from available portfolio
Inter-RAT cooperation	Not needed for channel selection use case	X	-

Table 17. Refinement of RRM functions for the case of Dynamic Channel Selection

The RRM function that does the heavy lifting is the RAT/spectrum and channel selection/ aggregation function. The algorithm that it follows selects a channel for a cell to use, based on certain criteria. Each node can gather information directly from its neighbours by REM, but cannot gather from any other nodes. The information is collected when a node wakes up at a random time with a specified rate of waking. This knowledge can be used to change one or more of the node's operating parameters in order to try and reduce interference; both the interference they experience and this they cause. The nodes operate autonomously so they must try to self-optimise in such a way that deliver the best possible performance for each user. Therefore the creation of an algorithm is required which, after the initial configuration, will allow each node to try and improve its capacity without greatly affecting its neighbours.

The effectiveness of the algorithm may be evaluated by looking at the capacity at the UEs. This involves calculating the signal to noise and interference ratio, SINR.

A basic colouring model can be used to allocate a colour (i.e. channel) to each node from a specified small number of colours, trying to make each node be a different colour from its neighbours. Upon wake-up, a node performs an REM operation and chooses a new colour from a discrete distribution, in which colour i has probability proportional to $\exp(-\alpha \, x[i])$, where $x[i]$ is the SINR in colour i, and α is a constant. This method has been shown by simulation to perform extremely well. There is an advantage over deterministic selection of the quietest channel, in that the system cannot get stuck in a globally bad state. Figure 5 shows an early simulation result using this method for a realistic LTE femtocell scenario, with accurate propagation modelling.
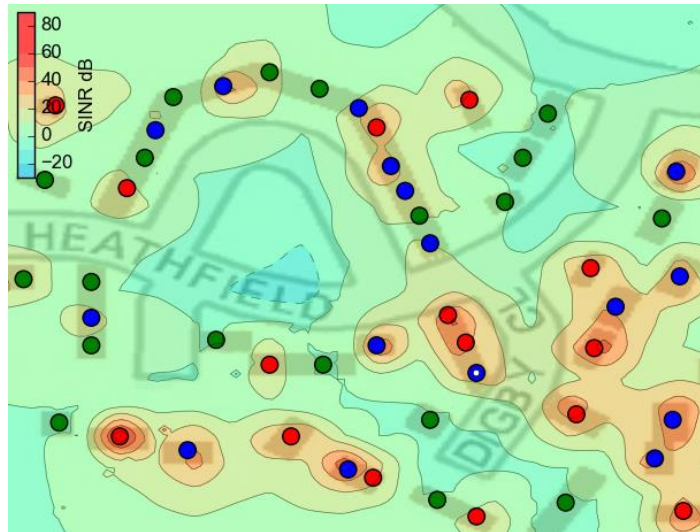

Figure 5. Example small area showing channel allocation algorithm

The algorithm needs further development and testing, e.g. to check if there are any applicable accuracy optimisation methods, and to work out co-working and backwards compatibility with existing LTE standards and with LAA / LTE-U.

3.3.3   Small cell throughput improvement with carrier aggregation

3.3.3.1   Scenario description

Carrier aggregation using components from the same band and using components from different bands are considered. Same-band components will have similar propagation characteristics and will likely have same regulatory constraints, and KPIs and decisions on factors like antenna configuration will the same. All the components will be treated jointly when considering coverage and capacity to a particular UE. In different bands, different regulatory constraints will apply, and different decisions on factors like antenna configuration will depend on UE location, since some of the components will suffer more loss than others.

3.3.3.2   RRM functions

It is necessary to refine the function processes for the case of small cell throughput. Table 18 below is a list of all the RRM functions in Figure 4, with inputs, outputs and processes customised for this case.

Function name	Inputs	Outputs	Processes
Admission and prioritisation	Traffic scenario and QoS requirements. Target KPIs and operator policy.	Priority levels	Terminal discovery, decision on priority given existing load, available capacity and operator

			policy.
Load balancing: offloading control	Traffic scenario and its required QoS, operator policy, deployment topology, e.g. from neighbour lists	Split of traffic flows across aggregated spectrum components	Optimise traffic flows
Load balancing: Traffic steering including base stations used in links	Traffic scenario and its required QoS, operator policy, deployment topology, e.g. from neighbour lists	Base stations involved in the sessions and throughput / delay allowances	Decide which base stations are involved in the sessions
Load balancing: Bearer establishment	Base stations involved in the sessions and throughput / delay allowances	Signalling to establish bearers on uplink and / or downlinks, tagging of packets	Establishes sessions or bearers, and determines their bandwidth / delay requirements
RAT/spectrum and channel selection / aggregation	Bandwidth / delay requirements, spectrum portfolio, spectrum quality, spectrum policy	Spectrum component selection, channel selection within components	Allocates licensed spectrum from available portfolio
Inter-RAT cooperation	Sensitivity of RATs to changes in UE location, error characteristics of spectrum components	Priority of components for traffic of different KPIs	Balancing of spectrum utility across the aggregated components

Table 18. Refinement of RRM functions for the case of Carrier Aggregation

3.4 RRM in support of D2D communications

Recently, wireless local area network (WLAN) technologies based on the IEEE 802.11 standards (e.g. WiFi, WiFi Direct) and wireless personal area network (WPAN) technologies (e.g. Bluetooth, Ultra-Wideband [UWB] technologies) have been increasingly used, due to the proliferation of smartphones and increasing number of people now living in cities. These technologies are designed for short distances between sender and receiver and therefore achieve very high data rates with low energy consumption.

However, communications on a licensed band of a cellular network can be better in terms of interference avoidance under a controlled environment.

Following are the shortcomings of the above mentioned technologies.

a) As WiFi and Bluetooth work in license exempt band, there are no guarantees that they work in every place since there is always the possibility of the presence of an interfering communication system or other sources of interference.

b) WiFi Direct can be used in every public place in the near future as devices become available, but this technology lacks global synchronisation (i.e. synchronisation can be used in wireless systems, generally to enable energy-efficient operations. For devices to discover each other, they must rendezvous in space and time. Only in a synchronised system the discovery periods can be both frequent and of low duty cycle. Thus, in practice, devices operating autonomously without infrastructure support in unlicensed spectrum can synchronise, but only locally).

In order to solve the above mentioned issues, device-to-device communication (D2D) concept has been proposed by LTE-A recently.

D2D allows users to communicate directly (send data directly on licensed/unlicensed band) without access to fixed infrastructure under control of operator (licensed band). The potential advantages of D2D communication are offload traffic, throughput enhancement, coverage expansion and UE energy saving [129].
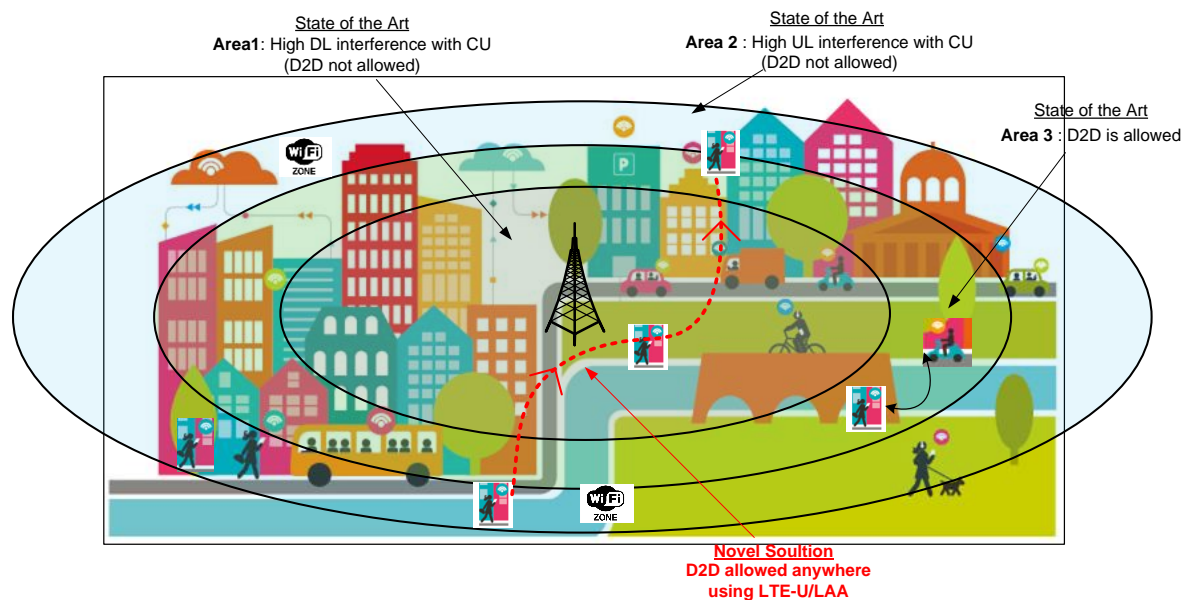

Figure 6. D2D offloading on LTE-U/LAA

In 5G network, D2D coexists as another tier with small cells network which can operate in either licensed or unlicensed band. D2D link will reuse the cellular resource, which creates two types of interference: 1) Intra-cell cross-tier interference between the D2D and cellular users (CU) and 2) Inter-cell interference between the D2D links in coverage area of different BS.

LTE release 12 and state-of-the-art literature on D2D suggest to use LTE uplink band for D2D communication which performs much better than D2D sharing the DL. However, D2D sharing the UL band leads to a higher interference to normal CUs. Therefore, there is a trade-off between D2D and CU performance when considering whether to use UL or DL band. Letting D2D transmission utilise the DL band favours CU reliability over D2D reliability, whereas letting D2D transmission utilise the UL band favours D2D reliability over CU reliability.

Moreover, geometrics area (as shown in Figure 6) is also important in the performance trade-off between D2D and CU.

- Cell centre (area 1) is generally off-limits to D2D transmission using DL band
- Cell edge (area 2) is generally off-limits to D2D using UL band
- If only cellular DL/UL bands can be used, reliable D2D communication would be kept away from cell centre/edge.

Therefore, to solve the above problem in D2D communication, this use case considers D2D communication using lightly licensed/unlicensed/TV white spaces using LTE-U/LAA like mechanism. By using LTE-U/LAA, D2D can operate anywhere in the cell coverage, expect for the region where other unlicensed band RATs are in use. D2D communication in licensed/unlicensed bands under operator control offloads the traffic, possibly using single/different RATs. Traffic with the most stringent QoS requirements is allocated on licensed or lightly licensed spectrum using normal cellular operation. Traffic with non-stringent PER and latency requirements can be moved on D2D using lightly licensed/unlicensed bands. Predetermined association rules (traffic class vs. spectrum band)

could be established: i.e. high data rate (QoS: 2.6 GHz, Non-QoS: 5GHz, 2.3 GHz (ASA), 700MHz (TV white spaces).

The decision as to which channel (i.e. sub-band) within the band is selected is then delegated to the cell, which, for example, can use the exponential weighted randomised algorithm described in section 4.3.4.

4   Mathematical models for optimization

4.1   Introduction

Numerous recent developments in the mathematics of optimization have potential to be applied to the internal algorithms in SPEED-5G, specifically in the RRM. These methods fall into two general classes: (1) strict algorithms, in which the objective (to be minimized or maximized) and constraints can be precisely specified, and for which guaranteed solution methods are available; and (2) heuristics, which cover all other cases and for which solutions methods do not come with guarantees. Both classes come with centralized and distributed variants, and also possibly continuous and discrete variants. In general, strict algorithms should always be preferred to heuristics, but in practice this is rarely possible, and existing and near-future systems use or will use distributed heuristics in most cases. However, it should be recalled that in SPEED-5G, the RRM is centralized.

Some of the most promising recent developments are in the area of convex optimization, in which the objective (to be minimized) is a convex function, and the feasible set (the set of points satisfying all constraints) is a convex set [95]. For this class, which includes examples of direct interest to us such as setting downlink power in a set of small cells in order to maximize the minimal SINR at all UEs, efficient software such as cxvpy (http://cvxpy.org) is available and can be plugged directly into the RRM emulator code. This is currently being implemented, and more detailed will be presented in D4.3.

In the remainder of this chapter, we give detailed optimization models and methods for co-primary spectrum sharing, for D2D systems, and for resource allocation algorithms for coexistence of LTE-U and WiFi. Some of the work is to establish a baseline for Speed-5G and some is new work from Speed-5G.

4.2   Co-primary spectrum sharing in uplink SC-FDMA networks

In this section, we consider a common pool of shared spectrum, for the case of two Mobile Network Operators (MNOs) operating in the uplink direction of LTE. The objective of the spectrum sharing process is to assure exclusive access to the shared spectrum in order to avoid inter-operator interference. This scheme can be applied to both LTE macro and small cells. This is an extension of previous work on uplink LTE resource allocation, adapting the system model and respective formulations in order to be applied to co-primary spectrum sharing scenarios [22].

4.2.1   Uplink resource allocation in LTE systems

In order to appreciate the constraints of the spectrum sharing scheme to be described, we start by briefly summarizing the LTE protocol specification for the transmission of uplink scheduling requests (SRs) and the notification of the eNodeB of a macro cell or the Access Point (AP) of a small cell regarding the buffer status of each User Equipment (UE).

In the uplink direction, LTE systems operate on a Single Carrier Frequency Division Multiple Access (SC-FDMA) physical layer, which achieves considerably improved performance in terms of peak-to-average power ratio (PAPR) compared to Orthogonal Frequency Division Multiple Access (OFDMA). In the case of localized SC-FDMA (LFDMA), employed by the majority of the relevant literature in uplink LTE resource allocation, only groups of contiguous resource blocks can be allocated to each user.

Resources on the LTE uplink are allocated to the users in terms of uplink scheduling grants. A scheduling grant applies to a specific carrier of a UE, and is not limited to a specific application class within the UE. A UE that requires uplink resources in order to transmit one or more of its pending

data packets sends a SR to the uplink scheduler by raising a simple flag, which is transmitted on the Physical Uplink Control Channel (PUCCH). A SR can occur on a periodic manner, and its frequency is a UE-specific parameter provided by the higher layers.

However, in order for the uplink scheduler to be able to determine the required amount of resources to be granted to each user, information on the amount of data available for transmission in the uplink UE buffers is also necessary. Therefore, as part of the uplink transmission through Medium Access Control (MAC) elements, information on the UE buffer situation is provided to the eNodeB/AP in the form of Buffer Status Reports (BSRs). A BSR consists of a buffer size field, which contains information on the amount of data awaiting transmission across all logical channels in a logical channel group. The amount of data is indicated in number of bytes, and refers to all the data that are available for transmission in the Radio Link Control (RLC) and Packet Data Convergence Protocol (PDCP) layers. It has to be noted though that the size of the RLC and MAC headers are not considered in the buffer size computation.

4.2.2 Resource allocation utility function

On the uplink direction of an LTE network, resource allocation is performed on a per subframe basis. In order to perform resource allocation in a fair, QoS and energy efficient manner and evaluate the utility of each scheduling block of the shared spectrum pool to the users of each MNO, we introduce metric $m_{i,j,k}^{UL}(t)$ of user i, $i \in UE$, who is a subscriber of MNO k, for scheduling block j, $j \in \{1, 2, \ldots, N_{SB}\}$, where N_{SB} is the number of scheduling blocks per subframe of the shared spectrum, as follows:

$$m_{i,j,k}^{UL}(t) = p_k \frac{d_i^{UL}(t)}{d_{th,i}} \exp\left(\frac{\overline{D}_i^{UL}(t)}{\overline{R}_i^{UL}(t)}\right) \frac{r_{i,j}^{UL}(t)}{P_{1,i} \cdot N_{RB,SB}} \tag{1}$$

Here p_k is the shared spectrum access priority of MNO k. This can be either a static parameter, or a dynamic value that depends on the performance of spectrum allocations in the past. $d_i^{UL}(t)$ is the time passed since the last uplink grant was allocated to user i or since a SR has been received from this user and $d_{th,i}$ is the delay threshold, beyond which a packet of a real-time application is no longer considered usable and is discarded by the user's buffer. Since the eNodeB/AP does not have accurate information on the exact waiting time of the pending packets of each user, $d_i^{UL}(t)$ is used in order to allow a worst-case estimation of the packet delay, i.e., the case of a new packet entering the user's uplink buffer just after an uplink grant was allocated to the user or a SR was sent. Therefore, with the use of $d_i^{UL}(t)$, the prioritization of users who have waited for a higher amount of time since their last uplink grant or the latest SR, and are in higher risk of packet expiration, is achieved. $\overline{D}_i^{UL}(t)$ and $\overline{R}_i^{UL}(t)$ are the average delay and data rate, respectively, experienced by user i in the past, and are calculated using a weighted moving average formula as follows:

$$\overline{D}_i^{UL}(t) = \beta d_i^{UL}(t) + (1 - \beta)\overline{D}_i^{UL}(t - 1), \tag{2}$$

$$\overline{R}_i^{UL}(t) = \beta r_i^{UL}(t) + (1 - \beta)\overline{R}_i^{UL}(t - 1), \tag{3}$$

where $r_i^{UL}(t)$ is the instantaneous uplink data rate of user i and $0 \leq \beta \leq 1$. The incorporation of $\overline{D}_i^{UL}(t)$ and $\overline{R}_i^{UL}(t)$ in $m_{i,j,k}^{UL}(t)$ allows the prioritization of users that were served with high average delay and low average data rate in the past, thus increasing the fairness of the proposed solution. $P_{1,i}$ is the minimum uplink power per resource block of user i, which, based on the LTE uplink power control specification, is defined as follows:

$$P_{1,i} = \min\left\{P_{CMAX,C}, P_{0,PUSCH} + \alpha PL_i + 10\log_{10}(N_{RB}^{UL})\right\} - 10\log_{10}(N_{RB}^{UL}). \tag{4}$$

$P_{1,i}$ is calculated based on the assumption that all the resource blocks of an uplink slot are allocated to user i. Of course, the actual uplink power per resource block will almost always be higher for the specific user, and will depend on the actual number of its allocated resource blocks, which, in

principle, will be less than N_{RB}^{UL}. $P_{CMAX,C}$ is the configured UE transmit power, $P_{0,PUSCH}$ is the target received power per resource block, while PL_i is the user downlink pathloss estimate calculated in the UE and α, $0 \leq \alpha \leq 1$, is a parameter for pathloss compensation whose value is provided by the higher layers. $r_{i,j}^{UL}(t)$ is the data rate achieved by user i on scheduling block j and is defined as follows:

$$r_{i,j}^{UL}(t) = \frac{\left(L_{SB}^{UL} \log_2 M_{i,j} \right)}{T_{sf}} \tag{5}$$

where L_{SB}^{UL} is the number of data carrying resource elements per uplink scheduling block, which depends on the number of reference signals transmitted in a subframe, $M_{i,j}$ is the Modulation and Coding Scheme (MCS) of user i on scheduling block j and T_{sf} is the subframe length. In a generic SC-FDMA system that allows the selection of different MCS per scheduling block based on the perceived channel conditions, the value of $r_{i,j}^{UL}(t)$ is different for every scheduling block. However, since according to the LTE system specifications all scheduling blocks allocated to the same user have the same MCS, the value of $r_{i,j}^{UL}(t)$ and, consequently, the value of $m_{i,j,k}^{UL}(t)$ will be the same across all scheduling blocks.

4.2.3   Optimal uplink resource allocation (cake-cutting)

The problem of allocating a contiguous collection of scheduling blocks in a subframe of a shared spectrum pool to each user of each MNO is strongly connected to the traditional fair division (or cake-cutting) problem from social choice theory [96], [56], [79], [58]. In the traditional fair division problem there is a cake, represented as the [0, 1] interval, and a set of agents with each one obtaining a given utility for each $[x, y]$ interval, with $0 \leq x \leq y \leq 1$. The cake must be divided among the agents and there are various objectives that one might wish to optimize or adhere to, e.g., some fairness criterion, maximizing social welfare, etc. The version of the fair division problem that is most closely related to our setting is discrete connected cake-cutting. In this case, the cake is a sequence of indivisible items, i.e., non-overlapping (x, y) intervals whose union equals [0,1], and each agent must be allocated a consecutive subsequence of these items. The agent utility functions are assumed to be additive, i.e., an agent's total utility upon receiving a subset of the items is equal to the sum of the individual utilities of each item. There is a straightforward reduction from our setting and the problem of assigning uplink scheduling blocks to users seeking to maximize a total utility function, to the problem of allocating cake pieces to agents in discrete connected cake-cutting, seeking to maximize welfare, i.e., the sum of agents' utilities. More specifically, the users of the LTE system under consideration are mapped to agents in the cake-cutting setting, the uplink scheduling blocks of a subframe are mapped to the sequence of indivisible items that form the cake, and the users' $m_{i,j,k}^{UL}(t)$ metric functions are mapped to the agent utility functions, see Figure 7. Therefore, if we define the set of allocated scheduling blocks to user i, $i \in UE$, as G_i, the total value that this user obtains from this allocation is referred to as $u_i(G_i)$. We define $u_i(G_i)$ as a complex function of the $m_{i,j,k}^{UL}(t)$ values, with the properties that (i) it is non-decreasing in all $m_{i,j,k}^{UL}(t)$'s, and (ii) there is a threshold τ, below which the Signal-to-Noise Ratio (SNR) of the scheduling block is very low, resulting in significantly increased BER, and making the scheduling block, and consequently all the allocated resources to the user, practically unusable. The vector of allocated scheduling blocks G is defined as $G = \{G_i\}_{i \in UE}$. Let $U(G) = \sum_{i \in UE} u_i(G_i)$ be the total utility of allocation G. Therefore, the main objective of the cake-cutting algorithm is to identify the optimal allocation of sets of contiguous scheduling blocks to the different users in a manner that maximizes the total utility, i.e.,

$$G^* = \arg\max_G \{U(G)\}. \tag{6}$$

Results in [126] show that computing the allocation that maximizes welfare in discrete connected cake-cutting is NP-hard. Moreover, it is shown that it is not possible to achieve an arbitrary approximation of the optimal welfare unless P=NP.
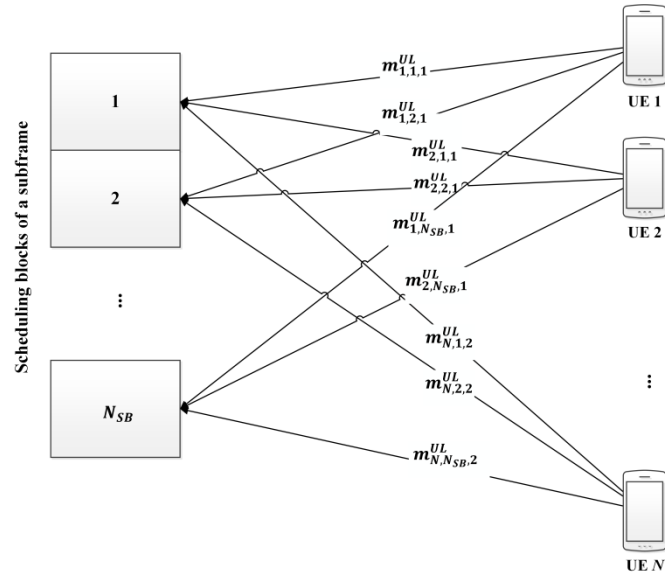
Figure 7. Valuation of the scheduling block utilities by the users.

The best polynomial time approximation algorithm obtained in the same paper achieves an 8-approximation of the optimal welfare, which implies it is hard to obtain an algorithm that offers guarantees of practical importance. Our problem, however, is much more general and positive results (such as this approximation guarantee) do not carry over to our setting. Concluding this section, we introduce the following modification, which is of interest in our setting. Consider the version of discrete connected cake-cutting, which includes an upper bound k_i on the cardinality of the set of contiguous resources G_i allocated to any user i. Each constant parameter k_i models the fact that agent i might be able to utilize at most k_i items. The established version of the problem, which does not consider parameters k_i, is appropriate for systems that assume infinitely backlogged traffic, i.e., users always having data to transmit, and always taking advantage of all their allocated scheduling blocks. However, in a realistic LTE system, the traffic models considered are not infinitely backlogged and a resource allocation algorithm needs to take into consideration the users' buffer status in order to avoid wasting resources by allocating them more scheduling blocks than actually needed. Therefore, we formally define the problem:

Definition 1. *Discrete connected cake-cutting with pieces of bounded size*

Suppose we are given a sequence of items $1, 2, \ldots, N_{SB}$, a set of players UE, and a utility $u_i(S)$, for every player $i \in UE$ and every contiguous subsequence of items S. Let $\mathcal{G}$ be the set of allocations, G, of items to players, such that G_i is a contiguous subsequence of items with $|G_i| \leq k_i$, for all $i \in UE$, and $|G_i \cap G_l| = 0$, for all $i \neq l$, $i, l \in UE$. We wish to find the optimal allocation of items to players that maximizes the total utility, i.e., $G^* = \arg\max_{G \in \mathcal{G}} \sum_{i \in UE} u_i(G_i)$. The literature on maximizing welfare in discrete connected cake-cutting does not explicitly consider the modified version we defined above. However, we note that the reduction in [126] still applies, even if we restrict the number of indivisible items per player to a small constant. This is due to the fact that the authors in [126] use a reduction from 3-dimensional matching to discrete connected cake-cutting, which results in instances such that any welfare maximizing allocation assigns at most 2 items to a player. Therefore, it can be concluded that the modified version of the cake-cutting problem defined in this section is also NP-hard.

4.3   Device to device

This section describes the potential D2D services that will be enabled by SPEED-5G. RRM functionality for D2D is novel to Speed-5G, and the mathematical models in this section are novel the to the project and have been published [30], [35].

4.3.1   Framework for D2D-based C-RAN

The demand for wireless mobile data continues to explode, leading to a surge in the number of smart devices in usage throughout the world. According to an Ericsson report [36], on average a laptop will generate 20 GB, a tablet 9.7 GB and a smartphone 8.5 GB of traffic per month by the year 2021. The main reason for this high traffic volume is the various data hungry applications running on those devices; e.g., high-definition wireless video streaming, health monitoring application and social networking.

To cope with such increasing mobile traffic and with the need for more sophisticated broadband applications and services, current standards and the deployed networks are to be extended towards their next generation, the so-called 5G, which is expected to be deployed around 2020. There is already a general consensus on the requirements of a 5G network[13], which are:

- Capacity: 1000x increase in area capacity
- Latency: 1ms Round Trip Time (RTT) latency
- Energy: 100x improvement in energy efficiency in term of Joules/bit
- Cost: 10-100x reduction in cost of deployment
- Mobility: Seamless indoor/outdoor mobility and always-on connectivity for high throughput users.

To achieve those 5G requirements we propose a novel syndicate architecture of Cloud Radio Access Network (C-RAN) and Device-to-Device (D2D) network (see).

C-RAN is one of the most promising 5G enabling technologies which can optimise network performance and reduce CAPEX/OPEX for the next-generation wireless systems. It has a centralised control ability to support joint resource allocation/processing and an advanced mobility management. The C-RAN access separates the baseband processing units (BBUs) from radio front-ends, also called Remote Radio Heads (RRUs) using a transport link, which is also known as Fronthaul. This new framework, with centralised BBUs, enables centralised and cooperative techniques. Besides the advantage of managing multi-cell and multi-user together, C-RAN is also fit for network deployment with multi-radio access technologies (RATs) and multi-layer networks coexistence, referred to as heterogeneous networks (HetNets). Along with the many advantages, C-RAN also brings some drawbacks, the most relevant of which for our study is the user plane latency, mainly due to fronthaul. With user plane latency here it is meant a RTT from C-RAN to end users (C-RAN↔fronthaul↔users). To alleviate this latency problem, we propose a tighter integration of D2D into C-RAN.

D2D is adopted as an effective and efficient candidate for very low latency in 5G. D2D communication can serve as a candidate paradigm to improve spectrum efficiency as well. In fact, by reusing the spectrum two D2D users can form a direct data link without routing through base stations (BS) and core networks. Further, when the cooperation between cellular users and D2D users is enabled, a win-win situation can be achieved to benefit all. D2D, reducing the distance between a transmitter and a receiver, also makes cellular system more energy-efficient. Moreover, D2D communication is the perfect candidate for centralised location-based services that enable

---

[13] https://5g-ppp.eu/

efficient, flexible and secure applications, including the social network ones. Those social network applications affect the interests of people, their locations and mobility, all important elements for D2D communications.

4.3.2 Objective

The above mentioned platform has led us to develop a new communication architecture called D2D-based C-RAN. The combination between C-RAN and D2D will be a game-changer for future 5G systems, which will provide more answers than questions with respect to the coexistence of different types of communication protocols, services and devices within a single network. This novel architecture solves most of the problems related to emerging 5G systems (capacity, latency, energy efficiency, CAPEX/OPEX and mobility). Moreover, we also propose a novel resource allocation method for D2D and cellular users.

4.3.3 Description of CRAN based D2D system

We consider a smart city scenario, focusing on a use case where most of the users are in close proximity and communicate with each other using their User Equipment (UE) via D2D connection ($UE_1 \leftrightarrow UE_2$). Along with D2D users, we also consider normal cellular users who communicate via RRUs ($UE_1 \leftrightarrow RRU \leftrightarrow UE_2$). RRUs are deployed for radio coverage and can be picocells or microcells depending upon the specific use case under consideration, but for the sake of simplicity we refer to them all as RRU in this paper. RRU is only used to transfer radio signalling, whereas all other signalling processing/baseband processing (PHY, MAC, Transport, and control) are done in the C-RAN.

In C-RAN, baseband processing is centralized and shared among sites in a virtualized BBU Pool. This means that it is very much able to adapt to smart city non-uniform traffic and utilizes the available resources, i.e. the BSs, more efficiently. Not only fewer BBUs are needed in C-RAN compared to the traditional architecture, but also C-RAN has the potential to decrease the cost of the network operation, because power and energy consumption are reduced. New BBUs can be added and upgraded easily, thereby improving scalability and easing network maintenance. Virtualized BBU pools can be shared by different network operators, allowing them to rent RAN as a cloud service. Therefore, mechanisms introduced for LTE-A to increase spectral efficiency, interference management and throughput, such as enhanced Inter-Cell Interference Coordination (eICIC) and CoMP, are greatly facilitated.

In spite of several advantages, one of the main disadvantages of C-RAN is the user-plane delay due to the FH, which should be less than 1ms according to 5G. To alleviate this delay problem, we propose D2D integration into C-RAN. Figure 8 shows the proposed scenario.
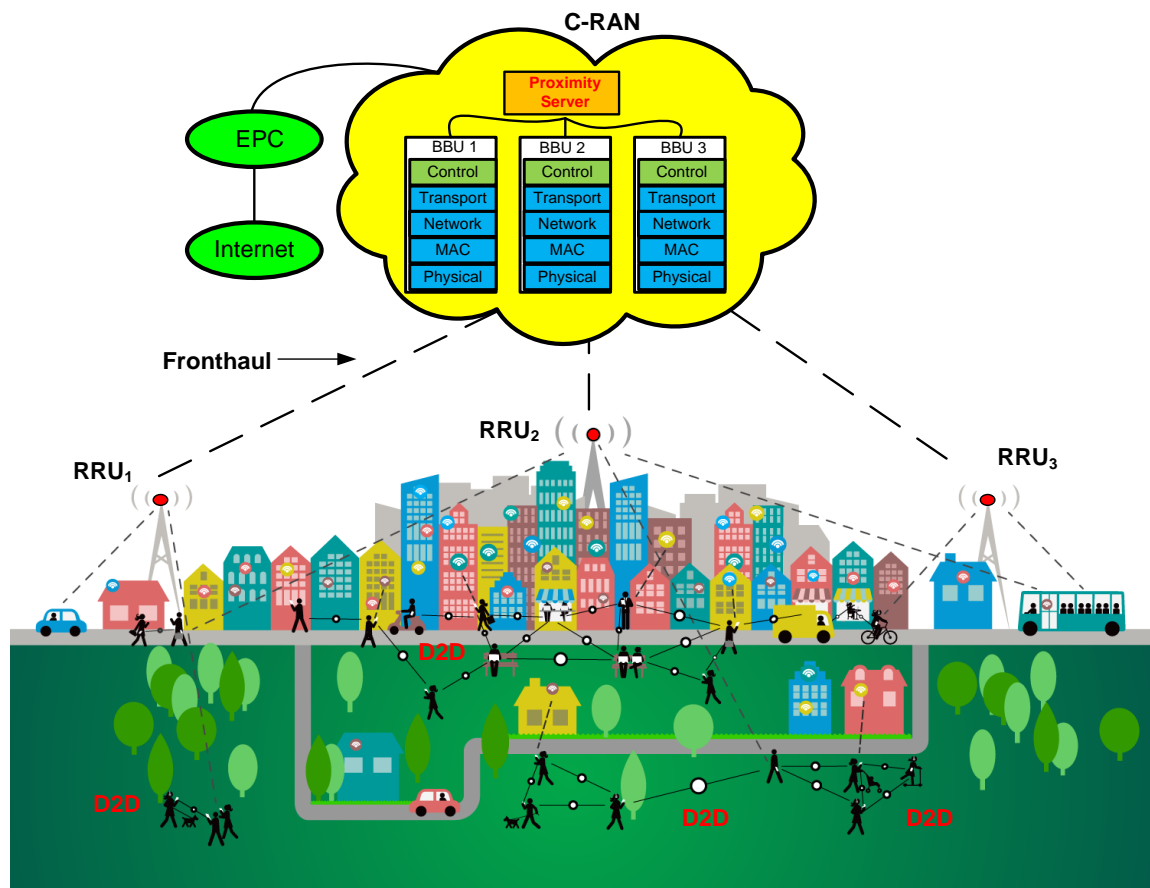
D2D based C-RAN (a)

Figure 8. D2D-based C-RAN

D2D proposes to be a "zero delay" technology for 5G networks and will solve the FH delay problem in C-RAN. To enable D2D communication in a C-RAN architecture, we propose a proximity server which resides inside the C-RAN, with interfaces connected to each BBUs as shown in Figure 8.

The ***proximity server*** functionality can be split up into two portions, proximity discovery and direct communication, as shown in Figure 9(a). The major function of proximity discovery is to discover users which are in proximity. This discovery mechanism can be user-assisted or network assisted, which can also perform as a standalone application to users such as in social networking where a direct communication is not needed to be triggered. Direct communication is initiated when proximity discovery is not needed to transfer data [100].

For the control plane in the access network of C-RAN, the cellular access link (Uu) is utilized by proximity whereas for the data plane a novel Direct Mobile Communication interface (Ud) is required for direct communication between devices, as shown in Figure 9(b).
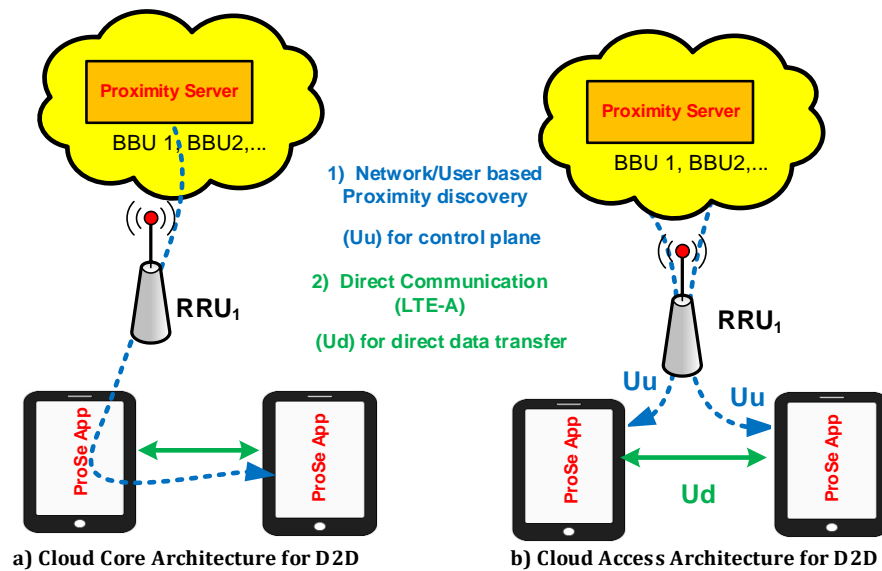
Figure 9. Cloud core architecture for D2D
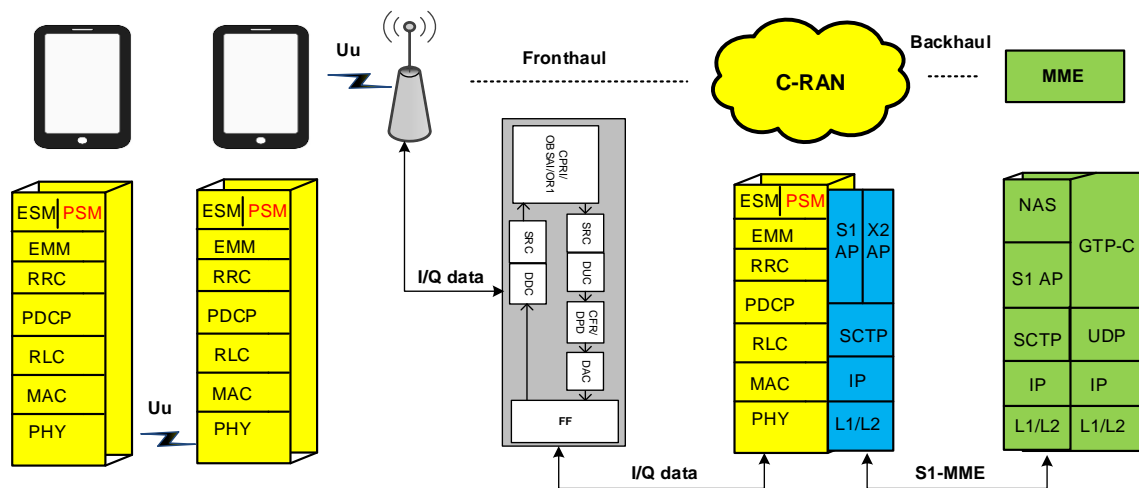
Figure 10. C-RAN-based D2D: control plane and data plane protocols

Figure 10 (b) focuses on the data plane protocols of an LTE-A system and brings in the novel interface named Ud for the data plane, which is enhanced at the PHY layer with D2D functionalities [15]. Radio

bearers—it can be one or more than one— are established for the data plane transmission over direct path transmission. The physical (PHY), medium access control (MAC), radio link control (RLC) and packet data convergence protocol (PDCP) layers of the wireless protocol stack for these radio bearers are all ended at the respective UE side.
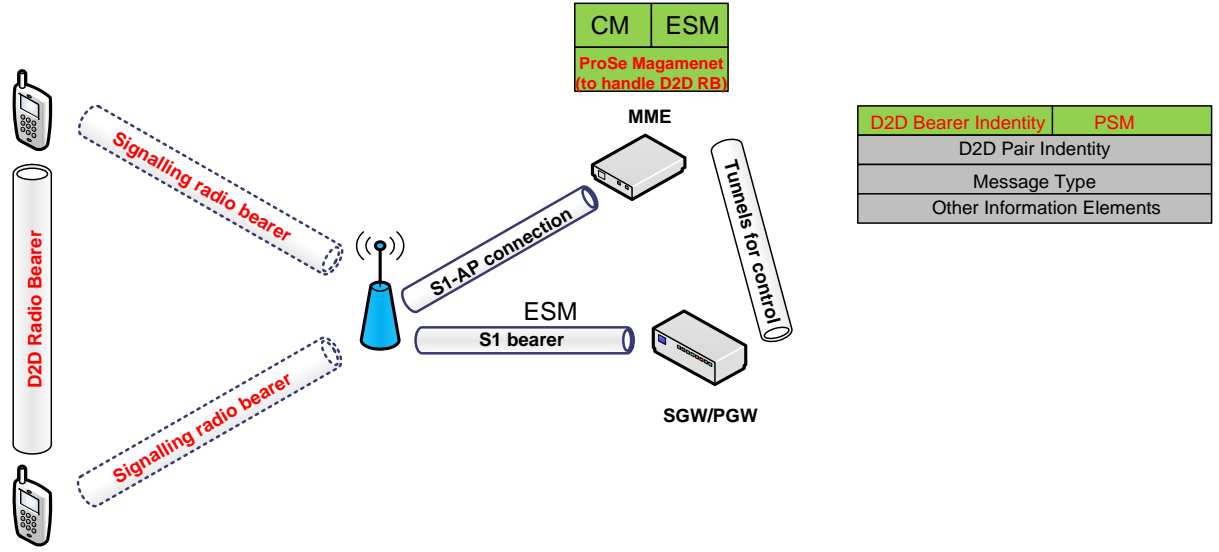

Figure 11. D2D bearer architecture

Figure 11 demonstrates the procedure of D2D Bearer Establishment. The P-GW starts the D2D bearer establishment procedure regarding the QoS control. It sends create D2D bearer request message [pair identity, D2D bearer QoS, D2D traffic flow filter, switchback EPS bearer identity (SBI)] to the MME. The SBI is the EPS bearer identity of the switchback-associated bearer. The MME activates the D2D session and selects a D2D bearer identity, which has not yet been assigned to the UE. The PSM builds ProSe setup request message including the D2D bearer identity. The MME delivers D2D bearer setup request message to the eNB. The eNodeB maps the D2D bearer QoS to the D2D radio bearer QoS. The eNB sends RRC connection reconfiguration message to the UE. The UE stores the D2D bearer identity and links the D2D bearer to the SBI. The ProSe setup request message is included in this message. The UE acknowledges the D2D radio bearer activation to the eNB by sending RRC connection reconfiguration complete message [75].

4.3.4 QoS-aware energy-efficient algorithm

This section presents a novel algorithm to achieve maximum energy efficiency (EE) considering total power consumption for a given data rate of this CRAN based D2D scenario. The SINR, Γ, is used to achieve the optimal rate for each PRB. Using this rate, we optimized the energy efficiency through the rate allocation approach.

The total bandwidth BW is equally divided into B PRBs, each with a bandwidth of $W = BW / B$. Then, the spectral efficiency, SE, obtained by Shannon's theorem, for user u on PRB b is

$$SE_c^{u,b} = \log_2\left(1 + \Gamma_c^{u,b}\right). \tag{7}$$

Then, the maximum achievable data rate, $\overline{r}_c^{u,b}$, for user u on PRB b is

$$\overline{r}_c^{u,b} = W \cdot SE_c^{u,b}. \tag{8}$$

We can also introduce, $r_c^{u,b}$, the data rate for user u on PRB b at any instant. EE is defined as the ratio of the total system throughput over the energy consumption within a given period at the

transmitter side where the unit is bits/Joule [47], [14]. In the subsequent subsection the problem formulation to optimize EE is explained.

4.3.4.1   Problem Formulation

In our scenario we use downlink OFDMA transmission. The overall system energy efficiency, EE, ε_{EE}, is defined as

$$\varepsilon_{EE} = \frac{\textbf{Total System Throughput}}{\textbf{Total Power}} = \frac{R}{P_{Total}}. \tag{9}$$

Accordingly, the optimization problem can be formulated as shown:

$$\max_{\textbf{r}} \varepsilon_{EE} \begin{pmatrix} = \dfrac{\displaystyle\sum_{c=1}^{C}\sum_{u=1}^{U}\sum_{b=1}^{B} r_c^{u,b}}{P_{Total}} = \dfrac{\displaystyle\sum_{c=1}^{C}\sum_{u=1}^{U}\sum_{b=1}^{B} r_c^{u,b}}{\displaystyle\sum_{k=1}^{K} A_k g\left(P^{BS}\right)_k} \\[2em] = \dfrac{\displaystyle\sum_{c=1}^{C}\sum_{u=1}^{U}\sum_{b=1}^{B} r_c^{u,b}}{\displaystyle\sum_{k=1}^{K} A_k g\left(\mu g P + \xi + \rho g R\right)_k} \\[2em] = \dfrac{\displaystyle\sum_{c=1}^{C}\sum_{u=1}^{U}\sum_{b=1}^{B} r_c^{u,b}}{\displaystyle\sum_{k=1}^{K} A_k g\left\{\mu g\left(\displaystyle\sum_{u=1}^{U}\sum_{b=1}^{B} p^{u,b}\right) + \xi + \rho g R\right\}_k} \end{pmatrix}, \tag{10}$$

Subject to:

1) $\hat{r}_c^{u,b} \le r_c^{u,\text{b}} \le \bar{r}_c^{u,\text{b}}$ **or** $r_c^{u,b} = 0$,

2) $r_c^{u,b} \ge 0$,

where $\hat{r}_c^{u,b}$ denotes the minimum rate requirement of this system for user u on PRB b and our optimization variable is the rate vector **r** and $R = \displaystyle\sum_{c=1}^{C}\sum_{u=1}^{U}\sum_{b=1}^{B} r_c^{u,b}$. Since the corresponding optimization problem is an optimization approach that exhibits non-linear and non-convex features and is NP-hard (non-deterministic polynomial-time hard), we aim at developing efficient sub-optimal algorithms that show a good trade/off between performance and complexity." To ensure the convexity of the proposed optimization problem, constraint 1 needs to be re-defined to meet the convexity requirement. Based on the non-negativity constraint 2, constraint 1 further simplifies to a cubic inequality [26] which helps us to solve our problem as a convex optimization problem. We reformulated our constraint based on non-negativity, as given by the following:

$$\left(r_c^{u,b}\right) \cdot \left(r_c^{u,b} - \hat{r}_c^{u,b}\right) \cdot \left(\bar{r}_c^{u,\text{b}} - r_c^{u,b}\right) \ge 0. \tag{11}$$

From here, we then derive four conditions with the 1st and 3rd conditions satisfying our constraint's non-negativity and convexity:

$$\textbf{1)}\, r_c^{u,b} = 0; \; \textbf{2)}\, r_c^{u,b} \in \left(0, \hat{r}_c^{u,b}\right);$$
$$\textbf{3)}\, r_c^{u,b} \in \left[\,\hat{r}_c^{u,b}, \bar{r}_c^{u,b}\,\right]; \; \textbf{4)}\, r_c^{u,b} \in \left(\bar{r}_c^{u,b}, +\infty\right). \tag{12}$$

In the following, we provide the analysis regarding those conditions that define our solution space where we assume some arbitrary value (it can be any real positive value) as a starting point.

Let's say $\hat{r}_c^{u,b} = 2$ and $\overline{r}_c^{u,b} = 5$.

Condition 1: $r_c^{u,b} = 0$;

We put values of $r_c^{u,b}, \hat{r}_c^{u,b}$ and $\overline{r}_u^{-b}$ in Equation (11) above, which leads to the following:

$$
\left(r_c^{u,b}\right) \cdot \left(r_c^{u,b} - \hat{r}_c^{u,b}\right) \cdot \left(\overline{r}_c^{u,b} - r_c^{u,b}\right)
$$
$$
= (0) \cdot (0 - 2) \cdot (5 - 0)
$$
$$
= 0
$$

That means this condition satisfies the non-negativity.

Condition 2: $r_c^{u,b} \in \left(0, \hat{r}_c^{u,b}\right)$;

In this condition we can take any value for $r_c^{u,b}$ between 0 to $\hat{r}_c^{u,b}$ excluding endpoints (since open interval does not include endpoints) i.e., $0 < r_c^{u,b} < \hat{r}_c^{u,b}$. We assume $r_c^{u,b} = 1$. We put values of $r_c^{u,b}, \hat{r}_c^{u,b}$ and $\overline{r}_c^{u,b}$ in Equation (11) to attain the following

$$
\left(r_c^{u,b}\right) \cdot \left(r_c^{u,b} - \hat{r}_c^{u,b}\right) \cdot \left(\overline{r}_c^{u,b} - r_c^{u,b}\right)
$$
$$
= (1) \cdot (1 - 2) \cdot (5 - 1)
$$
$$
= (1) \cdot (-1) \cdot (4)
$$
$$
= -4
$$

That means this condition does not satisfy the non-negativity.

Condition 3: $r_c^{u,b} \in \left[\hat{r}_c^{u,b}, \overline{r}_c^{u,b}\right]$;

In this condition we can take any value for $r_c^{u,b}$ between $\hat{r}_c^{u,b}$ to $\overline{r}_c^{u,b}$ including endpoints (since closed interval includes endpoints) i.e., $\hat{r}_c^{u,b} \leq r_c^{u,b} \leq \overline{r}_c^{u,b}$. We assume $r_c^{u,b} = 3$. We put values of $r_c^{u,b}, \hat{r}_c^{u,b}$ and $\overline{r}_c^{u,b}$ in Equation (11). Afterwards we get given in the following

$$
\left(r_c^{u,b}\right) \cdot \left(r_c^{u,b} - \hat{r}_c^{u,b}\right) \cdot \left(\overline{r}_c^{u,b} - r_c^{u,b}\right)
$$
$$
= (3) \cdot (3 - 2) \cdot (5 - 3)
$$
$$
= (3) \cdot (1) \cdot (2)
$$
$$
= 6
$$

This means that this condition satisfies the non-negativity.

When $r_c^{u,b} = \hat{r}_c^{u,b}$, we put values of $r_c^{u,b}, \hat{r}_c^{u,b}$ and $\overline{r}_c^{u,b}$ in Equation (11) to attain:

$$
\left(r_c^{u,b}\right) \cdot \left(r_c^{u,b} - \hat{r}_c^{u,b}\right) \cdot \left(\overline{r}_c^{u,b} - r_c^{u,b}\right)
$$
$$
= (2) \cdot (2 - 2) \cdot (5 - 2)
$$
$$
= (3) \cdot (0) \cdot (2)
$$
$$
= 0
$$

This means that this condition satisfies the non-negativity.

When $r_c^{u,b} = \overline{r}_c^{u,b}$, we place values of $r_c^{u,b}, \hat{r}_c^{u,b}$ and $\overline{r}_c^{u,b}$ in Equation (11) to attain:

$$\left(r_c^{u,b}\right)\cdot\left(r_c^{u,b}-\hat{r}_c^{u,b}\right)\cdot\left(\overline{r}_c^{u,b}-r_c^{u,b}\right)$$
$$=(5)\cdot(5-2)\cdot(5-5)$$
$$=(5)\cdot(3)\cdot(0)$$
$$=0$$

That means this condition satisfies the non-negativity in all 3 circumstances.

Condition 4: $r_c^{u,b}\in\left(\overline{r}_c^{u,b},+\infty\right)$.

In this condition we can take any value for $r_c^{u,b}$ between $\overline{r}_c^{u,b}$ to $+\infty$ excluding endpoints (since open interval does not include endpoints) i.e., $\overline{r}_c^{u,b}<d_u^m<+\infty$. We assume $r_c^{u,b}=6$. We place values of $r_c^{u,b},\hat{r}_c^{u,b}$ and $\overline{r}_c^{u,b}$ in Equation (11). Afterwards we get given in the following

$$\left(r_c^{u,b}\right)\cdot\left(r_c^{u,b}-\hat{r}_c^{u,b}\right)\cdot\left(\overline{r}_c^{u,b}-r_c^{u,b}\right)$$
$$=(6)\cdot(6-2)\cdot(5-9)$$
$$=(6)\cdot(4)\cdot(-4)$$
$$=-96$$

That means this condition does not satisfy the non-negativity.

From the above analysis we can easily deduce that only the 1st and 3rd conditions satisfy the non-negativity. Therefore, applying those two conditions we can optimize our objective function that maximizes energy efficiency.

4.3.4.2   Optimization

Our problem has an optimal solution since its objective function is quasiconcave and the solution space defined by the constraints is convex. In other words, this is a convex optimization problem [95].

In this section we develop an algorithm using the gradient method by constructing Lagrangian function and checking the solution with KKT to meet the global optimal solution. Let us formulate the Lagrangian of our problem with the Lagrange multiplier λ :

$$L(\mathrm{r},\lambda)=\varepsilon_{EE}+\sum_{c=1}^{C}\sum_{u=1}^{U}\sum_{b=1}^{B}\lambda_c^{u,b}\left\{\left(r_c^{u,b}\right)\cdot\left(r_c^{u,b}-\hat{r}_c^{u,b}\right)\cdot\left(\overline{r}_c^{u,b}-r_c^{u,b}\right)\right\},\qquad(13)$$

with the corresponding Lagrange dual function

$$g(\lambda)=\max_{\mathbf{r}}L(\mathbf{r},\lambda).\qquad(14)$$

Hence the dual problem is formulated as follows:

$$\min_{\lambda}g(\lambda);\lambda\geq0.\qquad(15)$$

Since the objective functions of problem in equations (13) and (15) are differentiable, with respect to the primal variable **r** and dual variable λ , both problems can be solved by the projected gradient method [95]:

$$r_c^{u,b}(\tau+1)=\left[r_c^{u,b}(\tau)+\eta\cdot\frac{\partial L(\mathbf{r},\lambda)}{\partial r_c^{u,b}}\right]^+,\qquad(16)$$

$$\lambda_c^{u,b}\left(\tau+1\right)=\left[\lambda_c^{u,b}(\tau)-\delta\cdot\frac{\partial L\left(\mathbf{r},\lambda\right)}{\partial\lambda_c^{u,b}}\right]^+ , \tag{17}$$

where τ denotes the iteration index and η, δ are positive step sizes [122], and $[\cdot]^+$ is a projection onto the set of $\mathfrak{i}^+$. By setting $\dfrac{\partial L\left(\mathbf{r},\lambda\right)}{\partial\mathbf{r}}$ to zero (i.e., the KKT condition), one can obtain a solution.

The derivative of $\dfrac{\partial L\left(\mathbf{r},\lambda\right)}{\partial\mathbf{r}}$ is given below

$$\frac{\partial L\left(\mathbf{r},\lambda\right)}{\partial\mathbf{r}}=\frac{\sum_{k=1}^{K}A_k\,\mathrm{g}\left\{\mu\mathrm{g}\left(\sum_{u=1}^{U}\sum_{b=1}^{B}p^{u,b}\right)+\xi\right\}_k}{\left[\sum_{k=1}^{K}A_k\,\mathrm{g}\left\{\mu\mathrm{g}\left(\sum_{u=1}^{U}\sum_{b=1}^{B}p^{u,b}\right)+\xi+\rho\mathrm{g}R\right\}\right]_k^2}$$
$$+\lambda_c^{u,b}\left(\left(r_c^{u,b}-\hat{r}_c^{u,b}\right)\left(\overline{r}_c^{u,b}-r_c^{u,b}\right)+r_c^{u,b}\left(\overline{r}_c^{u,b}-r_c^{u,b}\right)-r_c^{u,b}\left(r_c^{u,b}-\hat{r}_c^{u,b}\right)\right) \tag{18}$$

The derivative of $\dfrac{\partial L\left(\mathbf{r},\lambda\right)}{\partial\lambda}$ is given below

$$\frac{\partial L\left(\mathbf{r},\lambda\right)}{\partial\lambda}=\left\{\left(r_c^{u,b}\right)\cdot\left(r_c^{u,b}-\hat{r}_c^{u,b}\right)\cdot\left(\overline{r}_c^{u,\mathrm{b}}-r_c^{u,b}\right)\right\} \tag{19}$$

According to the analysis of the above subsection, we propose the optimal resource allocation algorithm for achieving maximum EE in the downlink CoMP as summarized by Algorithm 1:

Algorithm 1: The novel optimizing rate allocation algorithm to maximize EE
1: Step 1→ Initialization
2: Set $r_c^{u,b}\left(0\right)$ and $\lambda_c^{u,b}\left(0\right)$ to some non-negative value for all users u and resource blocks b.
3: Step 2→ Optimization
4: Apply Gradient method.
5: Update $r_c^{u,b}\left(\tau+1\right)$ according to equation (16).
6: Update $\lambda_c^{u,b}\left(\tau+1\right)$ according to equation (17).
7: Step 3→
8: Iterate until the implementation converges to the optimality (or the total number of iterations are achieved), the algorithm stops, or else return to step 2.

4.3.4.3 Complexity analysis

The complexity of the proposed algorithm can be approximated by using the number of operations required by the sorting operation. The complexity of the sorting algorithm is bounded by the square of the number of elements used in the list. In this case the total number of elements is *UB* and thus the complexity is (*UB*) ^{2}, which can be quite high. A method to reduce the complexity of the sorting operation is to sort the elements resource per resource, obtaining the best user combination per resource and then sort again these values to select the best values for each user. This technique

provides exactly the same results as the original method, but with a complexity of $U^2 + B^2$ as compared to the original method with $(UB)^2$. In other words, the second method shows polynomial complexity which enhances its practical implementation.

For example, in a network with U=10 UEs and B=100 PRBs, the complexity of original algorithm gives $(10 \times 100)^2$=1e6, while the second algorithm gives $10^2 + 100^2 = 1.01$e4, which represents a considerable reduction.

Simulation modelling and results will be provided in D4.2.

4.4 Dynamic resource allocation algorithms for coexistence of LTE-U and WiFi

4.4.1 Introduction

This is baseline modelling of LTE-U and WiFi, to give a reference point for Speed-5G performance measurement. In recent years, the mobile data usage has grown by 70-2000 percent per annum. There is novelty in adaptation to the Speed-5G project and a paper submitted to EUROSIP [39].

This exponential increase in mobile data is driven by the fact that internet applications of all kinds are rapidly migrating from wired PCs to mobile smartphones, tablets, MiFis and other portable devices [108]. The applications, such as high definition video streaming, real-time interactive games, wearable mobile devices, ubiquitous health care, mobile cloud running on these devices, do not only demand higher data rates but also require an improved Quality of Experience (QoE). To meet such high data rate and QoE demands, three main solutions are proposed [108]: a) addition of more radio spectrum for mobile services (increase in MHz); b) deployment of smallcells (increase in bit/Hz/km^2); c) efficient spectrum utilisation (increase in bit/second/Hz/km^2).

For mobile operators, efficient spectrum utilisation is the most essential resource in this pursuit. Therefore, mobile operators have been offloading more and more data traffic from their overloaded networks to a large number of WiFi hotspots over the past years [40]. The approach of utilizing both licensed band and unlicensed band has helped mobile operators to narrow the gap between the limited capacity of cellular network and the fast growing demand of mobile broadband.

With a significant amount of unlicensed spectrum globally available in the 5GHz band [51], 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 utilise unlicensed spectrum to carry data traffic for mobile services with initial focus on the 5725-5850 MHz band for this use [51].

The FCC opened up unused TV spectrum for "Super WiFi" hotspots with superior characteristics, many of which are due to the relatively low carrier frequencies of TV bands [93]. LTE-U extends LTE to the unlicensed spectrum and aggregates the unlicensed spectrum with the licensed spectrum leveraging the existing carrier aggregation [69] technology. It can provide better coverage and larger capacity than cellular/WiFi interworking while allowing seamless data flow between licensed and unlicensed spectrum through a single Evolved Packet Core (EPC) network. For operators, LTE-U means synchronised integrated network management, same authentication procedures, more efficient resource utilisation and thus lower operational costs. For wireless users, LTE-U means enhanced user experience, i.e. higher data rates, seamless service continuity 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 meaningful degradation on the system performance. Currently, WiFi systems adopt a contention-based MAC protocol with random back-off mechanism [69]. 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 not only degrade the WiFi devices throughput but also overall throughput of the system.

4.4.2   Related work

This section summarises coexistence studies and techniques proposed in the literature between different wireless technologies. These studies mainly focus on the simultaneous operation of IEEE 802.15.4 and WiFi, LTE and Wide band - code division multiple access (W-CDMA), and WiFi and LTE in TV white space, ISM bands and also licensed spectrum [82].

Co-existence models between different wireless technologies based on transmit power and receive sensitivity is presented in [121], and on different implementation aspects of Bluetooth and WiFi in [103]. Adaptive frequency hopping mechanism and interference aware scheduling strategy for Bluetooth was proposed in [81]and experimental study was conducted focusing on physical and network layers aspects in [63]. A measurement based co-existence study was conducted in [101] and an interference mediation scheme in [20]. Experiments are carried out to identify Zigbee under co-existence with WiFi in [87], and co-existence of UWB and WiMAX is studied using spectrum sensing techniques in [99]. WiFi and WiMAX co-existence in adjacent frequency bands is analysed in [86], and in [76] authors investigated the deployment of heterogeneous vehicular wireless networks consisting of IEEE 802.11 b/g and 802.11e inside a tunnel for surveillance applications.

An approach to facilitate the coexistence between LTE and WCDMA in the same frequency band is presented in [73], and an approach to facilitate simultaneous operation of LTE, WCDMA and WiFi in the same band is presented in [71]. A coexistence study about downlink high-speed railway communication system with TDD-LTE is presented in [21] and a co-existence study between LTE and DVB-T2-Lite system is presented in [65]. Coexistence performance between LTE and mobile satellite services (MSS) systems is evaluated in [133].

Co-existence between LTE and WiFi has generated a lot of papers. For LTE in LAA, some proposed operating modes and different deployment scenarios are given in [52] and [24]. Possible sub-bands within 5 GHz unlicensed spectrum is investigated and case study about feasibility of introducing DFS for LAA is in [118]. Proposed Listen-Before-Talk (LBT) for some regions is presented in [3], and motivating factors for deploying LAA in the unlicensed spectrum are presented in [83]. 3GPP organised a workshop in June 2014 on LAA in [4]. Experiment and simulation-based studies of WiFi and LAA coexistence are carried out in [49], Coexistence between WiFi and LTE (900 MHz) is investigated in [17]. Qualcomm provides some important studies about facilitating WiFi-LAA coexistence in the unlicensed spectrum in [88]and in [89]. In [55] it is proposed that the unlicensed spectrum access by LTE-A small cells should be done on-demand basis. LBT-based coexistence performance of LAA and WiFi-based on different deployment scenarios is shown in [72]. Simultaneous operation of WiFi and LTE in the TV white space; two techniques are proposed to facilitate interference management in [94].

LBT-based LTE transmission along with RTS/CTS message exchange prior to starting the actual LTE transmission is proposed in [34]. An LBT-based approach proposed for LTE transmission considers handling of both inter-radio access technology (RAT) interference and intra-RAT interference is shown in [27]. Blank sub frame allocation technique by LTE is introduced in to facilitate simultaneous WiFi and LTE operation in the unlicensed spectrum in [38]. Silent gaps with a predefined duty cycle to facilitate better coexistence with WiFi are proposed in [111]. An UL power control-based mechanism is evaluated for LTE systems to allow simultaneous operation of WiFi and LTE in the unlicensed spectrum in [32]. Exchanging spectrum allocation information between WiFi and LTE via a common database for enabling simultaneous access to the unlicensed spectrum by LTE and WiFi in [84]. In [105], a game theoretic approach to share unlicensed spectrum between several operators is proposed. UE side necessary enhancements for the LTE unlicensed band operation is discussed in [77]. In [104], the authors propose a QoS-based strategy to split the unlicensed spectrum between WiFi and femtocell network.

Dynamic channel selection mechanisms are proposed in [[1], [34], [19], [91], [25]] for WiFi-LTE coexistence. In [34], authors propose a centralized algorithm for channel selection in LAA. The algorithm maximizes the area spectral efficiency in the network, subject to interference power and

spatial distribution of desired signal. Simulated annealing is used to get sub-optimal solution. In [1], authors discuss how least congested channel selection in WiFi along with subcarrier allocation in OFDM can be exploited for WiFi-LTE coexistence. In [91] and [19], authors propose a WiFi-LTE unlicensed system architecture for channel selection. In [91], dynamic channel selection based on time division and optimal power control is discussed. In [25], interference aware channel selection techniques are proposed for outdoor LTE-U pico and indoor WiFi cells.

4.4.3 Objective

The objective of this work is to maximize network throughput in multi-operator scenario for 5G mobile systems by jointly considering licensed & unlicensed band, user association and power allocation subject to minimum rate guaranty and co-channel interference threshold. We benchmark our results with results obtained from outer approximation algorithm (OAA).

The existing works focus on individual aspects, as discussed in the previous section. The main contributions of proposed work are summarized below

- We formulate a constrained optimization problem, Joint User Association and Power Allocation for Licensed and Unlicensed Spectrum (JUAPALUS) that maximizes sum rate of 5G mobile network in multi-operator scenario, subject to minimum rate guarantee and co-channel interference threshold.

- The JUAPALUS is mixed integer nonlinear programming (MINLP) problem. MINLP problems are generally considered as NP-hard, where optimal solution requires exponential increase in computational complexity when number of variables are increased. In this paper, we have used Nonlinear Optimization by Mesh Adaptive Direct Search (NOMAD), for resource allocation in 5G Heterogeneous Networks, to reach sub-optimal solution.

- Extensive simulation work has been carried out to verify ε-optimal solution.

- We benchmark our results with results obtained from outer approximation algorithm (OAA).

4.4.4 System Model and Problem Formulation

We consider multi-operator LTE-U HetNet, where two operators (A&B) have non-collocated macro cell, small cell, and WiFi Access point (AP) as shown in Figure 12. We assume that ideal backhaul exits between small cell and macro cell. We also assume that UE in an operator will be served by one base station (BS) among macro eNB on licensed band, small cell on licensed band/unlicensed band (5Ghz) and WiFi AP on unlicensed 5Ghz band for downlink transmissions. Let $x_i^{(k,o)}$ be binary indicator to show UE k is connected to which operator, operator A or operator B, i.e $o \in O$ and $O = \{A, B\}$. It also indicates UE k is connected by whom among the macro eNB on licensed band (m_l), small cell on licensed band (s_l), small cell on unlicensed band (s_u) and WiFi on unlicensed band (w_u), i.e $i \in I$ and $I = \{m_l, s_l, s_u, w_u\}$. This indicator can be described as

$$x_i^{(k,o)} = \begin{cases} 1, \text{if UE k is connected to BS i inoperator o} \\ 0, otherwise \end{cases} \tag{20}$$

Initially UEs connected to small cell will be served on licensed band, when traffic load is increased, LTE-U interface is turned-on in small cell and UEs in small cell are served with 5Ghz unlicensed band, same as in WiFi AP. For simplicity we assume that while sharing licensed band between small cell and macro cell, different orthogonal channels are used to avoid strong interference. Let K be the number of UEs that want to communicate among each other. The channel gain between k^{th} UE and BS i is h_i^k, G_o be the antenna gain and $\varsigma 10^{\varsigma/10}$ be the lognormal shadowing, where ς is the zero mean

gaussian random variable with standard deviation σ [89], and the channel gain h_i^k is modeled as

$$h_i^k = \bar{h}_i^k \varsigma G_o \left(\frac{d_o}{d} \right)^{\alpha}$$ where d is the distance between transmitter and receiver, d_o denotes antenna far field reference distance, α is path loss exponent, $\bar{h}_i^k$ is Rayleigh random variable.
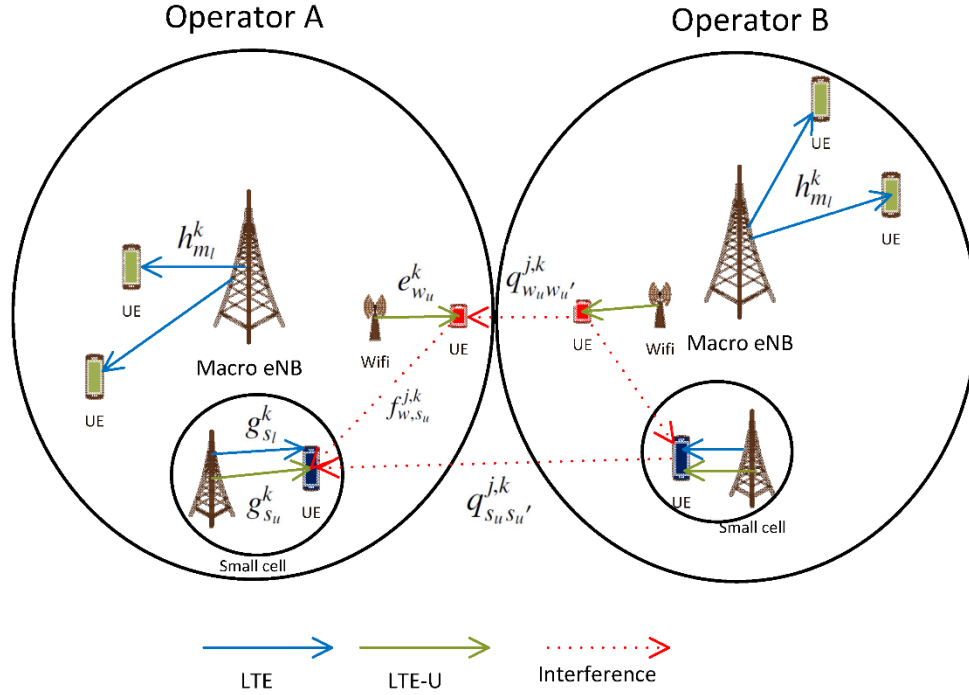

Figure 12 Scenario

Symbols	Definition
K	Number of Users (UE)
t_u	Portion of time unlicensed band is used by small cell UE
t_l	Portion of time licensed band is used by small cell UE
P_{s_u}	Max. power of small cell on unlicensed band
P_{w_u}	Max. power of WiFi A.P on unlicensed band
P_{s_l}	Max. power of small cell on licensed band
P_{m_l}	Max. power of macro cell on licensed band
$p_{s_u}^k$	power of user K connected to small cell on unlicensed band
$p_{w_u}^k$	power of user K connected to WiFi A.P. on unlicensed band
$p_{s_l}^k$	power of user K connected to small cell on licensed band
$p_{m_l}^k$	power of user K connected to macro cell on licensed band
R_{s_u}	Sum rate of UEs on small cell on unlicensed band
R_{s_l}	Sum rate of UEs on small cell on licensed band

R_{w_u}	Sum rate of UEs on WiFi AP on unlicensed band
R_{m_l}	Sum rate of UEs on macro cell on licensed band
R^k	Min required rate
$g^k_{s_u}$	Channel gain between UE and small cell on unlicensed band
$g^k_{s_l}$	Channel gain between UE and small cell licensed band
h^k_m	Channel gain between UE and macro cell on licensed band
e^k_w	Channel gain between UE and WiFi A.P on unlicensed band
$f^{j,k}_{w,s_u}$	Channel gain between UE k on small cell s_u and UE j on WiFi A.P
$q^{j,k}_{s_u s_u}$	Channel gain between UE k on small cell s_u and UE j on another small cell s_u
$q^{j,k}_{w_u w_u}$	Channel gain between UE k on WiFi w_u and UE j on another WiFi w_u
B_w	Total bandwidth on unlicensed band
B_l	Total bandwidth on licensed band
N_o	Additive white Gaussian noise
G_o	Antenna gain
ς	Zero mean Gaussian random variable for shadowing
C1-C11	Constraint 1 to Constraint 11

Table 19: Notations

4.4.5   Achievable data rate in unlicensed spectrum

The basic technique for unlicensed channel access in WiFi is distributed coordination function (DCF). DCF uses binary exponential back-off with carrier sense multiple access/collision avoidance (CSMA/CA). When a UE has a data packet to transmit, it monitors the channel. UE transmits if the channel is found idle for time equal to distributed interframe space (DIFS) time. Otherwise, if the channel is busy, UE waits for random backoff time, which is necessary to minimize chances of collision with data from other UEs. Moreover UE have to wait for random backoff time between two consecutive data transmissions, even if channel is found idle for DIFS time, which is necessary to avoid channel seizure by single user. We assume that small cell while on unlicensed band s_u, also maintains same random back-off mechanism, to ensure the coexistence with the WiFi A.P on unlicensed band w_u. According to author in [42], the probability to transmit for WiFi UE on unlicensed channel ρ_{w_u} is given by

$$\rho_{w_u} = \frac{2(1-\rho_{wc})}{(1-2\rho_{wc})(\omega_w+1) + \rho_{wc}\omega_w(1-(2\rho_{wc})^\ell)} \qquad (21)$$

where ρ_{wc} is collision probability, ℓ is maximum back-off stage and ω_w is back-off window size. The collision probability for WiFi UE on unlicensed channel ρ_{wc} is given by

$$\rho_{wc} = 1-(1-\rho_{s_u})(1-\rho_{w_u})^{n_w-1} \qquad (22)$$

where ρ_{s_u} is the probability to transmit for small cell UE on unlicensed channel, n_w is number of users on WiFi. The probability of successful transmission by a WiFi UE is given by

$$S\rho_{w_u} = n_w \rho_{w_u} ((1-\rho_{w_u})^{n_w-1})(1-\rho_{s_u}) \tag{23}$$

The probability to transmit for small cell UE on unlicensed band, ρ_{s_u}, is given by

$$\rho_{s_u} = \frac{2(1-\rho_{sc})}{(1-2\rho_{sc})(\omega_s+1)+\rho_{sc}\omega_s(1-(2\rho_{sc})^\ell)} \tag{24}$$

where ρ_{sc} is collision probability, ℓ is maximum back-off stage and ω_s is back-off window size. The collision probability for small cell UE on unlicensed channel, ρ_{sc}, is given by

$$\rho_{sc} = 1-(1-\rho_{w_u})^{n_w} \tag{25}$$

The probability of successful transmission by a small cell UE on unlicensed band is given by

$$S\rho_{s_u} = \rho_{s_u}((1-\rho_{w_u})^{n_w}) \tag{26}$$

We assume that the unlicensed band is used for the same duration by the small cell and WiFi system and the small cell changes its minimum backoff window size adaptively. The time fraction occupied by the small cell on unlicensed band, can be defined as t_u, which will be equal to $S\rho_{s_u}$, i.e $t_u = S\rho_{s_u}$ [91]. Based on the value of t_u, small cell adaptively changes the minimum size of back-off window. Small cell and WiFi will share unlicensed band, so we can write achievable sum rate for small cell on unlicensed band (R_{s_u}), as follows

$$R_{s_u} = \sum_{k \in \mathrm{K}} t_u log\left(1+\frac{p_{s_u}^k g_{s_u}^k}{N_o + \sum_{\substack{j=1 \\ j \neq k}}^{K} p_{w_u}^j f_{w,s_u}^{j,k} + p_{s_u}^j q_{s_u s_u}^{j,k}}\right) \tag{27}$$

where $f_{w,s_u}^{j,k}$ is channel gain between UE k on small cell s_u and UE j on WiFi A.P, $g_{s_l}^k$ is channel gain between UE and small cell licensed band and $q_{s_u s_u}^{j,k}$ is channel gain between UE k on small cell s_u and UE j on another small cell s_u. Achievable sum rate for WiFi on unlicensed band(R_w)

$$R_{w_u} = \sum_{k \in \mathrm{K}} log\left(1+\frac{p_{w_u}^k e_{w_u}^k}{N_o + \sum_{\substack{j=1 \\ j \neq k}}^{K} p_{s_u}^j f_{w,s_u}^{j,k} + p_{w_u}^j q_{w_u w_u}^{j,k}}\right) \tag{28}$$ where e_w^k is channel gain betwe

4.4.6　　Achievable data rate in licensed spectrum

For simplicity we assume that macro cell and small cell use orthogonal channels in licensed band in TDM fashion with no interference, achievable sum rate for small cell on licensed band (R_{s_l}) is

$$R_{s_l} = \sum_{k \in K} t_l \, log\left(1 + \frac{p_{s_l}^k g_{s_l}^k}{N_o}\right) \tag{29}$$

where t_l is time sharing factor for licensed band. Achievable sum rate for macro cell on licensed band R_m is given below

$$R_{m_l} = \sum_{k \in K} (1-t_l) log\left(1 + \frac{p_{m_l}^k h_{m_l}^k}{N_o}\right) \tag{30}$$

The optimization problem here is How to maximize sum rate of all UEs in network including s small cell UEs, m macro cell UEs and w WiFi UEs , by using power allocation (using parameter: $p_{s_u}^k, p_{w_u}^k, p_{m_l}^k, p_{s_l}^k$) and time allocation(using parameters: t_u, t_l), subject to minimum rate guaranty and co-channel interference threshold. Summary of notations is given in Table 19.

4.4.7   Problem formulation

We formulate joint power transmission and user association in licensed & unlicensed band such that sum rate is maximized. Mathematically, we have

$$\max_{t_u, t_l, p_{s_u}^k, p_{w_u}^k, p_{m_l}^k, p_{s_l}^k} \sum_{o \in O} \sum_{k \in K} \sum_{i \in I} x_i^{(k,o)} R_i^k$$

$$\text{subject to}$$

$$C1: \sum_{o \in O} \sum_{i \in I} x_i^{(k,o)} \leq 1, \ \forall \, k \in K,$$

$$C2: r^k \geq \sum_{o \in O} \sum_{i \in I} x_i^{(k,o)} R^k, \ \forall \, k \in K,$$

$$C3: \sum_{k \in K} p_{s_u}^k \leq P_{s_u}, \ \forall \, o \in O,$$

$$C4: \sum_{k \in K} p_{w_u}^k \leq P_{w_u}, \ \forall \, o \in O,$$

$$C5: \sum_{k \in K} p_{s_l}^k \leq P_{s_l}, \ \forall \, o \in O,$$

$$C6: \sum_{k \in K} p_{m_l}^k \leq P_{m_l}, \ \forall \, o \in O,$$

$$C7: p_i^k \leq x_i^k P_i \ \forall \, i \in I \, , k \in K, \forall \, o \in O,$$

$$C8: p_{w_u}^j f_{w,s_u}^{j,k} + p_{s_u}^j q_{s_u s_u}^{j,k} \leq \gamma$$

$$C9: p_{s_u}^j f_{w,s_u}^{j,k} + p_{w_u}^j q_{w_u w_u}^{j,k} \leq \gamma$$

$$C10: 0 \leq t_u \leq 1, 0 \leq t_l \leq 1,$$

$$C11: p_{s_l}^k \geq 0, p^k s_u \geq 0, p_{w_u}^k \geq 0, p_{m_l}^k \geq 0. \tag{31}$$

Constraint $C1$ is operator selection and BS is the selection constraint, which ensures that the UE is connected to one of the BSs (macro eNB on licensed band (m_l), small eNB on licensed band (s_l), small eNB on unlicensed band (s_u) and WiFi on unlicensed band(w_u)) in one of operators (operator A or operator B). $C2$ ensures that minimum rate requirement of each user is guaranteed. Constraint $C3$ to $C6$ are maximum power constraints for small eNB on unlicensed band (s_u), WiFi on unlicensed band(w_u), small eNB on licensed band (s_l) and macro eNB on licensed band (m_l), respectively. $C7$ ensure that power experienced by any UE must be zero if it not connected to

concerned BS. Constraints $C8$ and $C9$ guarantees the interference threshold. $C10$ is time sharing constraint for licensed and unlicensed band. $C11$ is minimum power constraint for each user.

4.4.8 Proposed Approach

The combination of integer and continuous variables along with their non-linear behaviour makes the problem in equation (31) very complex and challenging. However by exploiting special structure of the problem, we can use mesh adaptive direct search algorithm (MADS) [28] to reach sub-optimal solution. The MADS algorithm is an extension of generalized pattern search (GPS) algorithm [120].

4.4.9 Description of Algorithm

The MADS is an iterative pattern search algorithm, it evaluates objective function f on mesh of points. The mesh M_i at iteration i, given by

$$M_i = \bigcup_{y \in T_i} \{ y + \Delta_i^l \, \mathsf{D}_w : w \in N^n \},$$

(32)

where $\Delta_i^l \in R^+$ is the size of mesh, T_i is set of points where objective function is calculated at iteration i and $\mathsf{D} \in R^n$ is set of directions having maximum of n_{Max} directions. D can be considered as $n \times n_{Max}$ matrix containing n_{Max} directions. D must be positive spanning set [8], equal to product of G and W ($\mathsf{D} = GW$), where G is $n \times n$ nonsingular matrix and W is $n \times n_{Max}$ matrix.

There are three main steps of MADS algorithm: search, poll and update. In search step, objective function f is evaluated at any finite set of points T_i on mesh in the feasible region. If point y is not in the feasible region then value of function is set to $+\infty$. The search step allows creation of point anywhere on mesh, this flexibility restricts the search step to take part in convergence analysis.

If an improved mesh point (y_{i+1}) is found, iteration may continue with search step or it may stop according to user's choice. If improved mesh point is not generated in search step, the poll step is invoked. The poll step explores space of optimization variable near current solution with following set of poll trial points

$$P_i = \{ y_i + \Delta_i^l d : d \in \mathsf{D}_i \} \subset M_i$$

(33)

where D_i is positive spanning set, depicting poll directions. Points of P_i are generated so that their distance to the current solution y_i is limited by a parameter, called poll size $\Delta_i^p \in R^+$. Δ_i^p is always greater than $\Delta_i^l d$ (i.e $\Delta_i^p \geq \Delta_i^l d$). Moreover $\lim_{i \in I} \Delta_i^l d = 0$ if and only if $\lim_{i \in I} \Delta_i^p = 0$ for infinite subset of iteration l. The update step determines whether iteration i was successful or not. This step updates parameters, Δ_i^l, Δ_i^p, and T_i at end of each iteration. The pseudo-code for MADS is given in following Algorithm:

Algorithm:

Mesh Adaptive Direct Search

//Initialization

1) $i \leftarrow 0$

2) $\Delta_o^l \in R^+$, $\Delta_o^p \in R^+$, $y_o \in T_o$

//Search and Poll

3) *The Search Step*: Evaluate objective and constraint functions on finite number of points of $M(i, \Delta_i^l)$, to find y_{i+1}

4) *The Poll Step*: If y_{i+1} is not found, compute p MADS directions $D_i \in R^n$ .Construct set of points $P_i \subset M(i, \Delta_i^l)$ with y_i , D_i , Δ_i^p . Evaluate the objective and constraint functions on p points of P_i

//Update

5) Determine succes/failure of iteration i

6) Update solution(y_{i+1})

7) Update mesh(Δ_{i+1}^l)

8) Update poll size (Δ_{i+1}^p)

9) $i \leftarrow i+1$ Check stopping conditions and go to Search and Poll step.

4.4.10   Experimental Results

Simulation[14] results obtained from our simulation setup, depict the performance of proposed approach to solve MINLP problem in equation (31). Performance is portrayed in terms of network throughput and user distribution among base stations with licensed and unlicensed band. To implement MADS algorithm, non-linear optimization with MADS (NOMAD) software [15] is used. Simulation parameters are given in Table 20.

Parameter	Value
P_{s_u}	4 Watt
P_{w_u}	4 Watt
P_{s_l}	6 Watt
P_{m_l}	6 Watt
d_o	10 m
α	2
Small cell and WiFi radius	30 m
Macro cell radius	1000 m
G_o	50
ς	10dB

---

[14] Detail of simulation modelling and design will be provided in D4.2

R^k	1 Mbps
Unlicensed bandwidth	20 MHz
Licensed bandwidth	20 MHz
Min UEs	20
Max UEs	120
UE Increment	20

Table 20 : System Parameters

Figure 13 depicts the throughput of network when LTE-Licensed (LTE-L i.e m_l and s_l) is combined with LTE-Unlicensed (LTE-U i.e s_u) and WiFi (W_u) . It is clear that LTE-U gives better results in terms of throughput compared to WiFi when combined with LTE-L . Figure 13 also depicts the throughput of HetNet which combines LTE-L, LTE-U and WiFi, throughput is better than LTE-L only, however throughput slightly less compared to LTE-L+LTE-U case due to interference between WiFi and LTE-U.
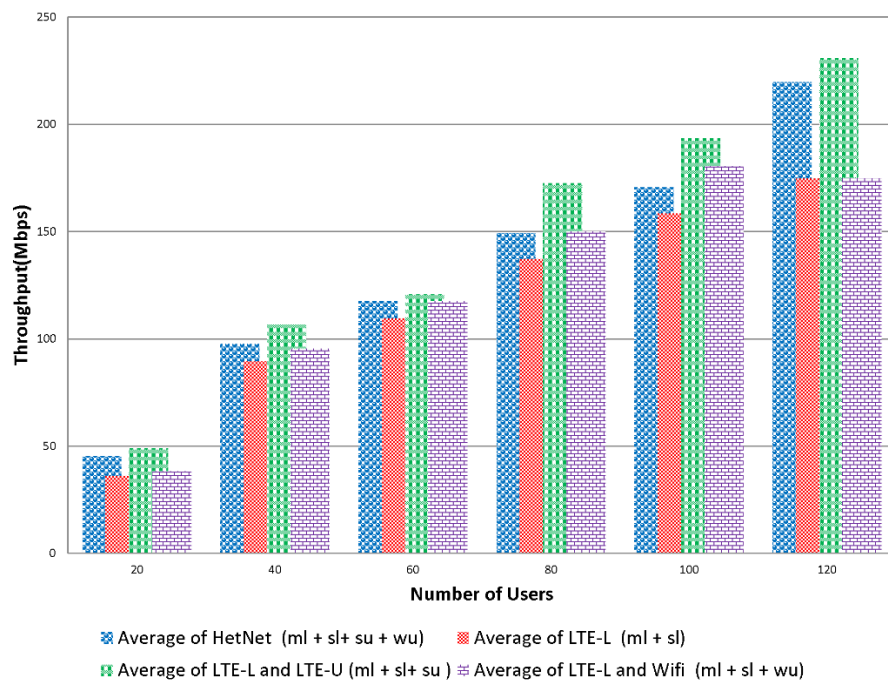

Figure 13: Network throughput comparison for LTE-U and WiFi when combined with LTE-licensed

Figure 14 shows UEs selected by each technology in case of HetNet. For low number of total UEs, LTE-L (m_l and s_l) gets more UEs, as we increase total number of UEs, more UEs camp on LTE-U and WiFi. In Figure 15, we benchmark the throughput results for LTE-U obtained from MADS against OAA, which shows MADS outperforms OAA.
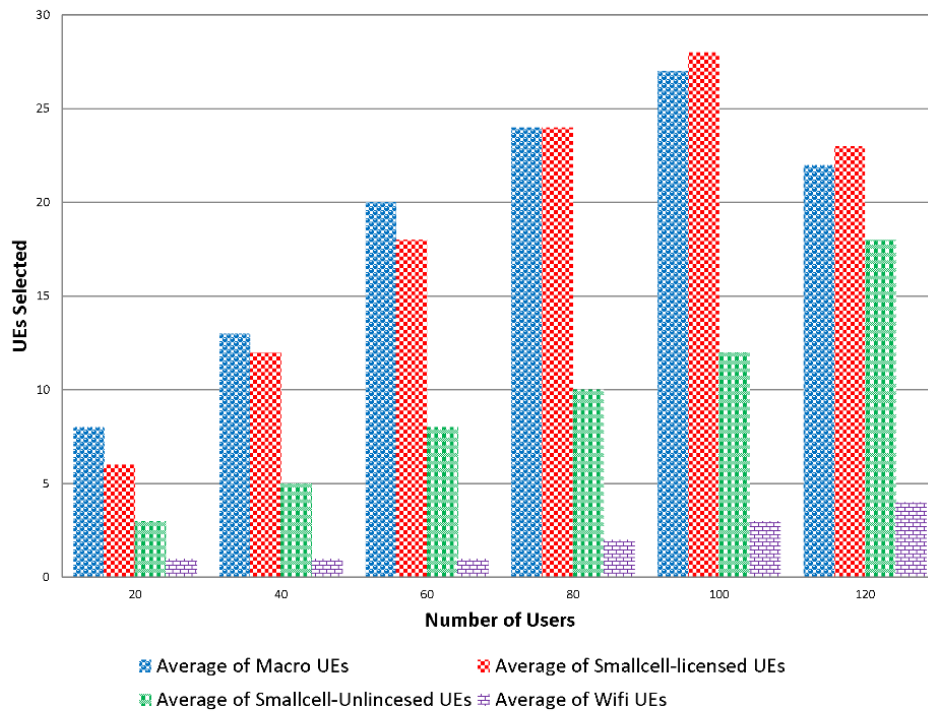
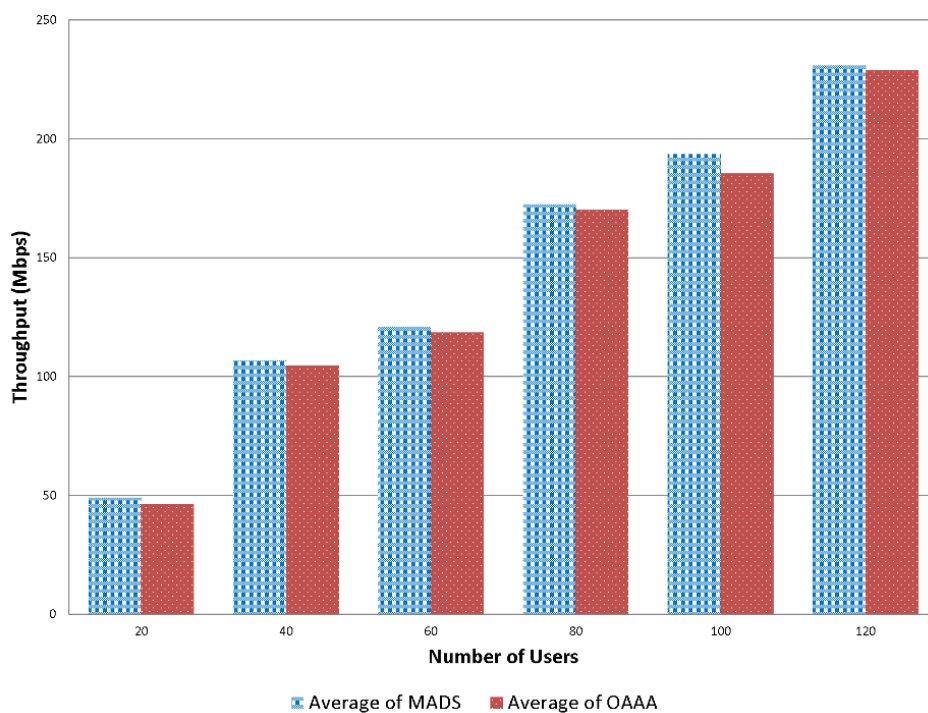*Figure 14: UE distribution in HetNet*

Figure 15: Throughput comparison of OAA vs MADS for LTE-U network

4.5 Interference and QoS aware channel segregation for heterogeneous networks

4.5.1 Objective

One of the main scenarios addressed in SPEED-5G is the case of heterogeneous networks [119], where a massive deployment of small cells is put in place to provide a uniform broadband experience to the users considering the provisioning of applications with different QoS requirements (such as high resolution multimedia streaming, gaming, video calling, and cloud services). A significant challenge in these networks is the efficient management of co-channel interference (CCI) that occurs due to proximity among the base stations of small cells (SBSs). Particularly, given that the same channels are reused among SBSs due to the scarce spectral resources, CCI constitutes an important factor that restricts the network performance.

4.5.2 Related work

To confront this challenge, Dynamic Channel Assignment (DCA) techniques have been proposed in the literature, either considering a centralised approach [107] or a distributed one [31]. In particular, in [107], a centralised DCA technique considering a heterogeneous network that consists of femtocells and macro cells is investigated based on the graph approach. It should be noted that the centralised approaches have several advantages in terms of performance. Nevertheless, the high computational complexity renders them inappropriate for the case of a heterogeneous network with a massive number of small cells. Therefore, the distributed DCA have gained the interest of many researchers as a solution that can be applied in future wireless networks. Towards this direction, one promising category of DCA mechanisms, which has recently been studied in the literature, refers to the channel segregation-based DCA (CS-DCA) mechanisms [98], [128]. According to this approach, each cell creates a priority table with the available channels and tries to use the channels with the highest priority to improve the spectrum efficiency. However, in the majority of the DCA schemes in the literature, the SBSs do not differentiate between traffic requests from UE applications, even if the applications do not have the same priority from the user point of view. Considering that in 5G networks, the traffic will range from high data rate to machine type traffic, covering a variety of different applications, there is an emerging need for DCA schemes that provides differentiated QoS to each user, coping with the changing network conditions and the time-varying CCI.

Based on this remark and the work presented in [90], we will study a modified distributed channel segregation mechanism that takes into account the CCI and the QoS characteristics of the users. The proposed Interference and QoS aware Channel Segregation-based DCA (IQ-CS-DCA) can be employed in order to use the spectral resources efficiently and at the same time prioritise the users with delay-constrained applications (such as video streaming).

4.5.3 Interference and QoS aware Channel Segregation-based DCA

A first version of the proposed solution is provided in the text that follows [102]-[54]. Specifically, two algorithms have been investigated. The first one is a random channel assignment algorithm (). This algorithm does not have a certain logic for assignment of channels that's why it's called 'random'. As input we have a set of UEs U that want to transmit, a set of macro BSs M, a set of channels C, and a set of available channels $C_a \subseteq C$. Then, a certain procedure is followed in order to allocate certain channels to UEs as the flowchart illustrates.
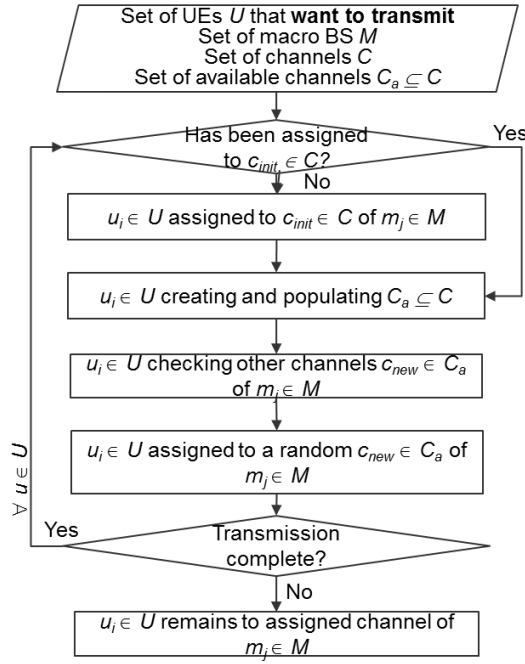
Figure 16. Random channel assignment algorithm

The random channel assignment algorithm is used as a baseline in order to evaluate the effectiveness of the next algorithm that we propose and is called "Best SINR-based channel assignment algorithm". As input we have again a set of UEs U that want to transmit, a set of macro BS M, a set of channels C, a set of available channels $C_a \subseteq C$. The selection procedure differs from the random channel assignment algorithm since, here, we introduce a control point for checking the best available channels in order to select these (if available). The best channel is identified according to the SINR and if the SINR of a new channel is better than the currently utilised one, then the UE will switch to the better channel. In general, it is expected that through this algorithm, it will be possible to achieve better quality (e.g. higher throughput, less latency, shorter session duration).

Also, it should be mentioned that the algorithm tries to tackle the aspect of scheduling starvation by creating a prioritisation list of served UEs (as a result, some priority will be given to UEs which have been served fewer times, compared to others).
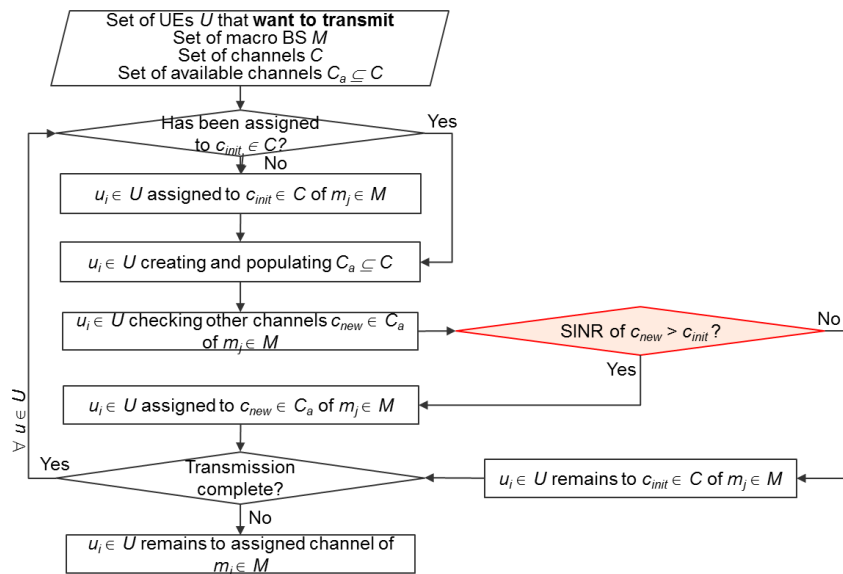

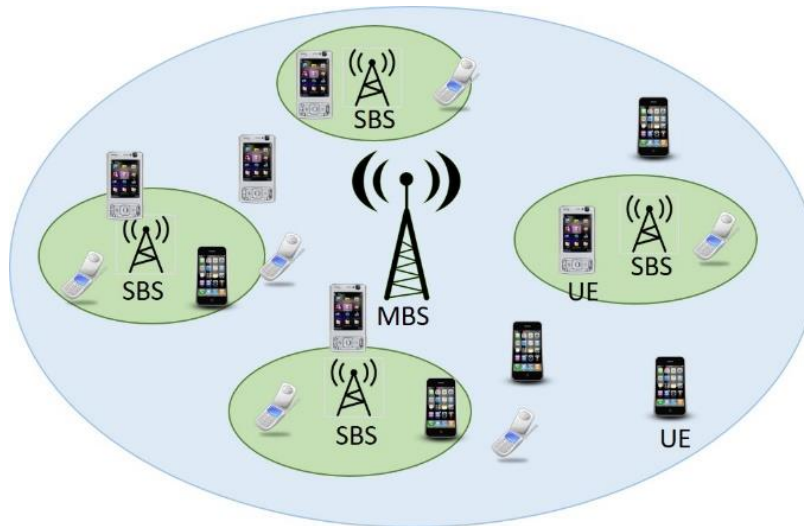*Figure 17. Best SINR-based channel assignment algorithm*

Figure 18. Heterogeneous network scenario

4.6 Context-aware User-driven framework

This section constructs a context-aware user-driven framework that can be used as a "fall-back" option when the new Speed-5G architecture (i.e., cRRM and reconfigurable MACs) is not fully supported. Based on the proposed framework, a generic fuzzy multiple attribute decision making (MADM) strategy is developed to select the best combination of spectrum bands, available RATs and licensing regimes in a given context. This work is new to the project, and has one paper published and another submitted [37], [43].

4.6.1 Functional architecture

A set of K available RATs ($\{RAT_k\}_{1\leq k\leq K}$) are considered by user equipments (UEs) to establish a set of L applications ($\{A_l\}_{1\leq l\leq L}$) that are characterised in terms a set of heterogeneous QoS/QoE requirements ($\{Req_l\}_{1\leq l\leq L}$). At the time of establishing each of these applications, the various contextual information that may be available about the UE (e.g., velocity, remaining power and remaining balance) and the network (e.g., operator strategy and regulation rules) should be also taken into account as they may have a strong impact on the suitability of each of these RATs.

Therefore, the problem considered here is whenever an application A_l needs to be established, how to:

- make the **UE** select the best RAT,
- to meet the set of **application requirements**,
- in the considered **context**?

To enable such context-aware user-driven mode of operation, a functional split between the UE and network domains should be clearly made to identify the logical entities and scope of each side. In this respect, the functional architecture described in Figure 19 is proposed [37]. Specifically, a connection manager (CM) is introduced at the UE to collect the relevant components of the context from both the terminal and network sides to implement a given decision-making policy (e.g., the proposed fuzzy MADM). The collected contextual information is combined with a radio characterisation of each available RAT in terms of a set of short-term attributes (e.g., signal strength, SNR and load) obtained e.g., 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 parameters (e.g., algorithm, fuzzy logic membership functions and MADM weights). The content of the policy repository may be

retrieved in practice from a local instance following a pull or push mode using e.g., the Open Mobile Alliance-Device Management (OMA-DM) protocol [2]. To offer a higher degree of flexibility to the network manager, a policy designer entity enables to build and update the policy repository content based on a set of measurement reports collected from UEs and the various network-level strategies and constraints (e.g., operator strategy and regulation rules). For instance, the policy designer may dynamically adjust some of the policy-related parameters (e.g., MADM weights) or the RAT attributes (e.g., cost) to optimise some network-level metrics (e.g., spectrum efficiency and energy efficiency) or implement a form of traffic steering (i.e., push UEs to use specific RATs during some periods of time). Finally, the CMs of different UEs may collaborate to further improve their individual performances.
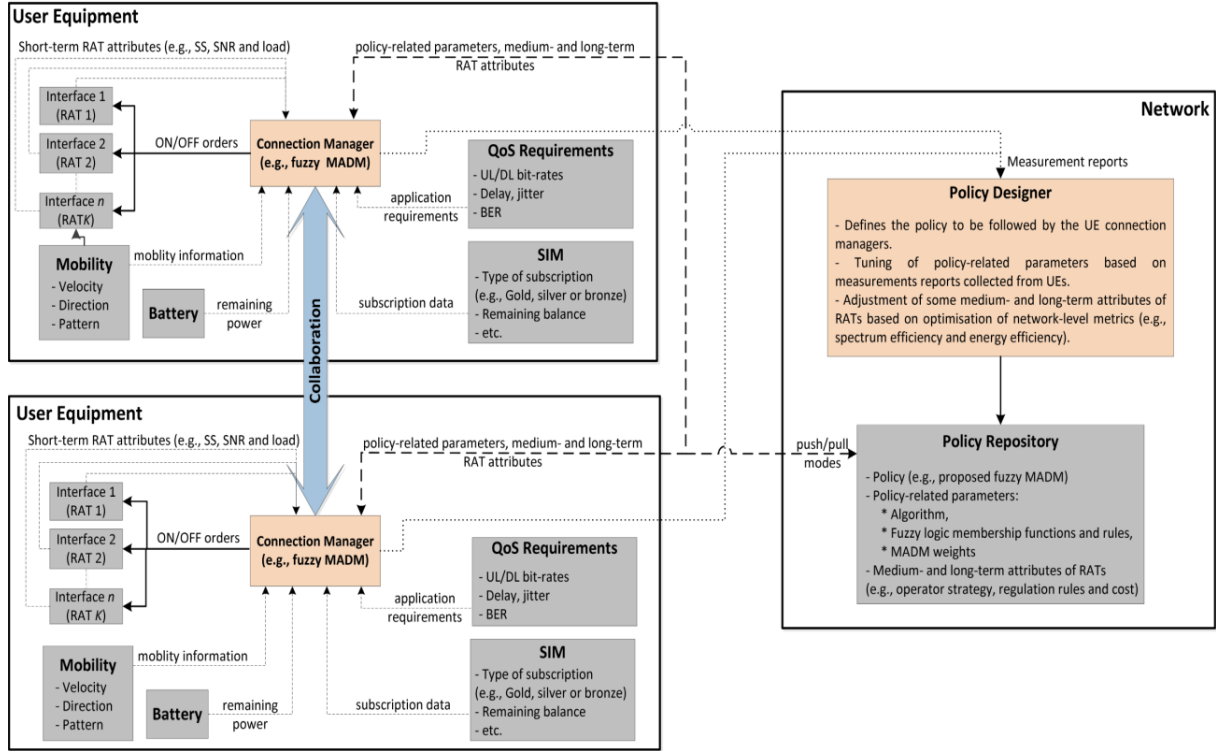

Figure 19. Functional architecture of the proposed context-aware user-driven framework

4.6.2 Connection manager: Fuzzy MADM decision-making

This section proposes to select, for a given application A_l, the RAT_k that would best meet each of the associated QoS/QoE requirements in a given context. To this end, a three-step approach is proposed [37]:

1) Design a fuzzy logic calculator to estimate the "out-of-context" suitability level to meet the application requirements ($s_{k,l}^{oc}$) out of the available radio parameters.

2) Develop a fuzzy MADM methodology to combine $s_{k,l}^{oc}$ with the various components of the context to derive the so-named "in-context" suitability level ($s_{k,l}^{ic}$).

3) Select the RAT that maximizes the in-context suitability level $s_{k,l}^{ic}$.

4.6.2.1 Out-of-context suitability levels

Given the uncertainty and lack of information associated with UEs, this section develops a fuzzy logic calculator to estimate the suitability levels of each RAT to meet the various requirements of the considered applications.

The key building block of fuzzy logic reasoning is the fuzzy logic controller (FLC) whose block diagram is described in Figure 20. It is composed of three main modules, namely the fuzzifier, inference engine and defuzzifier. During fuzzification, crisp (i.e., real) input data are assigned a value between 0 and 1 corresponding to the degree of membership in a given fuzzy set. Then, the inference engine executes a set of *if-then* rules on the input fuzzy sets. These rules, referred to as inference rules, are maintained in a rule base that is typically built based on previous expert knowledge. Finally, the aggregated output fuzzy sets are converted into crisp outputs using a given defuzzification method.
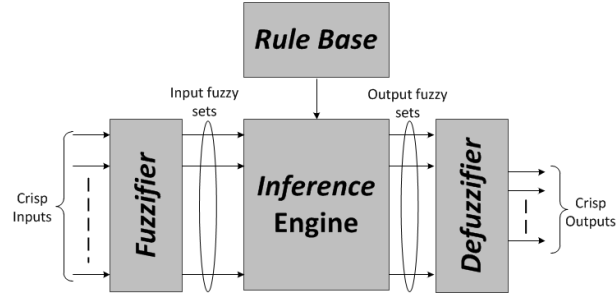

Figure 20. Block diagram of fuzzy logic controller

In our case, the main challenge is how to design a FLC that would reliably estimate the QoS/QoE suitability level in the most scalable way. To this end, the main design requirements can be highlighted as follows:

1) How to minimize the number of FLC inputs?
2) How to adjust all the relevant parameters (e.g., membership functions and inference rules)?

To tackle these issues, it is proposed to develop a single FLC for each pair of RAT_k and application A_l. This enables to directly incorporate the set of application requirements (i.e., Req_l) when designing the set of rules. To further improve scalability, the minimum set of input radio parameters that are relevant for the considered application should be designed. Finally, all membership functions and inference rules need to be adjusted by the policy designer of Figure 19 based on the actual measurements reported by the various UEs to capture any dependency on the proprietary features and settings of the network (e.g., scheduler in LTE).

4.6.2.2   In-context suitability levels

In this section, the previously determined estimates are combined with the various components of the context to derive the in-context suitability levels.

To particularly cope with the heterogeneity of the various components of the context, the following MADM methodology is proposed:

- **MADM formulation**: The decision-maker is in our case a UE who wants to establish an application A_l and has to select among a set of alternatives (i.e., RATs). 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, 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:

$$
\mathbf{D_l} = \begin{array}{c} \\ \\ RAT_1 : \\ \\ \vdots \\ \\ RAT_k : \\ \\ \vdots \\ \\ RAT_K : \end{array} \begin{array}{ccccc} QoS / QoE & cost & power & range \\ s_{1,l}^{oc} & cost_1 & power_1 & range_1 \\ \vdots & \vdots & \vdots & \vdots \\ s_{k,l}^{oc} & cost_k & power_k & range_k \\ \vdots & \vdots & \vdots & \vdots \\ s_{K,l}^{oc} & cost_K & power_K & range_K \end{array}
\tag{34}
$$

To adjust the relative importance of the various attributes for each of the considered applications, a vector $\mathbf{w_l}$ of M weights $\left(\left\{w_{l,m}\right\}_{1\leq m \leq M}\right)$ is introduced:

$$
\mathbf{w}_l = \begin{bmatrix} w_{l,QoS} \\ w_{l,cost} \\ w_{l,power} \\ w_{l,range} \end{bmatrix}
\tag{35}
$$

When both performance metrics $d_{k,m}^{l}$ and weights $w_{l,m}$ are crisp numbers, the traditional MADM ranking methods, such as simple additive weighting (SAW) and technique for order preference by similarity to ideal solution (TOPSIS), can be efficiently used to first rank the various alternatives and then select the best one. Without loss of generality, SAW will be considered in the remainder of this paper for its simplicity.

- **Cope with fuzziness**: In our case, most of the attributes (e.g., cost and power) and their corresponding weights are difficult to quantify and would be much better expressed in terms of linguistic terms (e.g., LOW, MEDIUM and HIGH). The considered problem becomes fuzzy MADM and classical MADM methods cannot be directly applied to solve it. To handle such imprecision, over a dozen fuzzy MADM methods have been developed [106]. However, these approaches are too cumbersome to be implemented in UEs because fuzzy data are operationally difficult to manipulate. Therefore, it is proposed to solve the considered fuzzy MADM problem using the simplified approach proposed in [106]. It consists in two steps, first converting fuzzy data to crisp numbers, and second applying classical MADM methods to rank the alternatives. In what follows, it is assumed that all fuzzy data in $\mathbf{D_l}$ and $\mathbf{w_l}$ have been converted to crisp numbers. Therefore, classical MADM ranking methods can be applied.

- **MADM ranking**: According to the SAW ranking, the vector $\mathbf{s_l^{ic}}$ of in-context suitability levels ($\left\{s_{k,l}^{ic}\right\}_{1\leq k \leq K}$) can be obtained by combining the various MADM attributes and weights as follows:

$$
\mathbf{s_l^{ic}} = \begin{bmatrix} s_{1,l}^{ic} \\ \vdots \\ s_{k,l}^{ic} \\ \vdots \\ s_{K,l}^{ic} \end{bmatrix} = \mathbf{D_l} \cdot \mathbf{w_l}
\tag{36}
$$

4.6.2.3   Decision-making

Based on the previous sub-section, the best RAT that maximises the in-context suitability level is selected for application A_l:

$$k*(l) = \arg \max_{k \in \{1,\ldots,K\}} (s_{k,l}^{ic}) \qquad (37)$$

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 handover (HO) to the best RAT is performed during sessions. This may be triggered on an event-basis (e.g., emergency situation due to QoS degradation) or periodically (i.e., comfort HO).

4.7   Mapping of maths models to RRM high level functions

Table 21 shows the relevance of each maths model to the RRM high level functions. The models can be complementary or alternative, and further developments of these will be presented in report D4.2.

	Discrete exhaustive search (exact)	Convex optimization (exact)	Co-primary sharing of uplink SC-FDMA (heuristic)	Dynamic resource allocation LTE-U and WiFi co-existence (heuristic)	Interference and QoS-aware channel segregation (heuristic)	Context-aware user-driven framework (heuristic)
Admission, prioritization, and steering	X		X			X
Load-balancing and offloading control				X		X
RAT and spectrum selection and aggregation				X		
Channel selection and downlink power control	X	X			X	X
Inter-RAT cooperation			X	X		

Table 21. Mapping of maths models to RRM high level functions

5 Conclusions and next steps

The project uses cases are Broadband Wireless, Massive IoT, Ultra-Reliable Communications and High Speed Mobility. For these use-cases, the Key Performance Indicators (KPIs) are derived from several white papers from industry associations (FMCF, NGMN), standards (3GPP), research projects and vertical sectors work groups in the 5G PPP framework. The KPIs have been analysed and performance metrics have been derived.

The state of the art has been reviewed from relevant studies on other EU projects, 3GPP and the wider industry, and gaps identified for next steps. Architectures are discussed that focus on multi-RAT (Radio Access Technologies) and joint usage of unlicensed and licensed bands, such as LTE-U/MuLTEFire, LAA, LTE-WiFi link aggregation/LTE-H and related technologies (carrier aggregation, dual connectivity, C/U plane split). Further state-of-the-art is presented for the potential available and emerging technologies that will be used to implement the RRM functions. These technologies include (i) Spectrum sharing particularly on uplink, (ii) Resource allocation in D2D-based C-RAN, (iii) Dynamic Resource Allocation Algorithms for coexistence of LTE-U and WiFi, (iv) Interference and QoS Aware Channel Segregation for Heterogeneous Networks and (v) context aware user-centric RAT selection.

It should be noted that SPEED-5G is considering a reallocation of certain functions as detailed in the SPEED-5G deliverable D5.1 "MAC approaches with FBMC and simulation results" that may depart from the 3GPP allocation of these functions to the RRM layer. In particular, (i) the Dynamic Resource Allocation (DRA) - Packet Scheduling (PS) function is clearly allocated to the MAC layer, (ii) the Load Balancing (LB) function may be allocated to the SON layer although currently located in the cRRM for the traffic offloading use case and finally, (iii) Inter-Cell Interference Coordination (ICIC), and enhanced Inter-Cell Interference Coordination – (eICIC) that may be allocated to the SON layer, but have not been explicitly studied in SPEED-5G yet.

The next steps are to perform detailed design of RRM functions and interfaces, and to test these by simulation. The focus will be on centralised RRM and interfaces to the MAC layer and also northbound interfaces to the OAM, spectrum manager and KPI manager. Then the RRM will be integrated with the MAC layer to simulate the eDSA, and some aspects of these will be carried forward to hardware in the loop demonstrations later in the project.

References

[1]   "Cisco visual networking index: Global mobile data traffic forecast update, 2015to-2020 white paper," Feb 2016.

[2]   "OMA Device Management Protocol," Open Mobile Alliance, Tech. Rep. version 1.2.1, June 2008.

[3]   "Review of Regulatory Requirements for Unlicensed Spectrum," 3GPP RP-140054, Fukuoka, Japan, Sep. 2014.

[4]   "Summary of a workshop on LTE in Unlicensed Spectrum," Tech. Rep. 3GPP TSG RAN Meeting 63, Fukuoka, Japan, Mar. 2014.

[5]   3GPP TR 22.803 v12.1.0, "Feasibility study for Proximity Services (ProSe)", 2013.

[6]   3GPP TR 36.889 V13.0.0, "Feasibility Study on Licensed-Assisted Access to Unlicensed Spectrum", June 2015.

[7]   3GPP TR 37.870 V13.0.0, "Study on Multiple Radio Access Technology (Multi-RAT) joint coordination", June 2015.

[8]   3GPP TS 36.300 V13.2.0, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2", January 2016.

[9]   3GPP TSG RAN3, R3–071438, Load balancing SON Use case. Athens: Greece, Alcatel-Lucent, August, 2007.

[10]   Nikos Dimitriou et al, "Context-aware Radio Resource Management in Hetnets.". Workshop on Interference and Design Issues for Heterogeneous Networks, IEEE WCNC 2014.

[11]   Wang, Hossain, Bhagava, 'Backhauling 5G small cells, a Radio Resource Management perspective. IEEE Wireless Communications, Volume 22, Issue 5.

[12]   HP Technical White paper 'RRM technology', October 2014.

[13]   3GPP, Radio Access Network Technical Specification Group, "TR 36.927, Potential solutions for energy saving for E-UTRAN (Release 10)," 2010.

[14]   3GPP, Radio Access Network Technical Specification Group, "TS 36.300 V.13.3.0; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description", Mar. 2016.

[15]   3GPP, TS 36.321, "Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification."

[16]   Abolfazl Mehbodniya and Fumiyuki Adachi, "Radio Resource Management for Next Generation Wireless Networks", 2014.

[17]   A. Cavalcante, E. Almeida, R. Vieira, F. Chaves, R. Paiva, F. Abinader, S. Choudhury, E. Tuomaala, and K. Doppler, "Performance Evaluation of LTE and Wi-Fi Coexistence in Unlicensed Bands," in Proc. IEEE Vehi. Technol. Conf. (VTC), June 2013, pp. 1–6.

[18]   A. De Domenico, E. Calvanese Strinati, A. Capone, "Enabling Green cellular networks: A survey and outlook," Computer Communications, Volume 37, 1 January 2014, Pages 5-24.

[19]   A. Mukherjee, J. F. Cheng, S. Falahati, L. Falconetti, A. Furuskr, B. Godana, D. H. Kang, H. Koorapaty, D. Larsson, and Y. Yang, "System architecture and coexistence evaluation of licensed-assisted access lte with ieee 802.11," in 2015 IEEE International Conference on Communication Workshop (ICCW), June 2015, pp. 2350–2355.

[20]   B. H. Jung, J. W. Chong, C. Jung, S. M. Kim, and D. K. Sung, "Interference Mediation for Coexistence of WLAN and Zigbee Networks," in Proc. IEEE Int. Symp. On Personal, Indoor and

Mobile Radio Communi., Sept 2008, pp. 1–5.

[21] B. Han, Y. Liang, L. Huo, X. Zhang, and D. Yang, "Coexistence of Downlink High-Speed Railway Communication System with TDD-LTE Cellular Communication System," in Proc IEEE Vehicular Technol. Conf. (VTC Spring), May 2012, pp.1–6.

[22] D. Triantafyllopoulou, K. Kollias, and K. Moessner, "QoS and Energy Efficient Resource Allocation in Uplink SC-FDMA Systems", IEEE Transactions on Wireless Communications, vol. 14, no. 6, pp. 3033-3045, June 2015.

[23] Boggis, J., Castro, J., Liu, H., Minokuchi, A., Salam, D., Schwinghammer, P., & Shahbaz, M. (2006). Next generation mobile networks beyond HSPA & EVDO. A white paper by the NGMN alliance, 5 December 2006.

[24] Broadcom Corporation, "Discussion on Licensed Assited LTE," 3GPP RAN1 standard contribution - RWS-140022, Sophia Antipolis, France, Jun. 2014.

[25] C. Davis, "Theory of positive linear dependence," American Journal of Mathematics, vol. 76, no. 4, pp. 733–746, 1954.

[26] C. Li, H. Xiong, J. Zou, and Z. He, "Joint Source and Flow Optimization for Scalable Video Multirate Multicast Over Hybrid Wired/Wireless Coded Networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 5, pp. 550 –564, May 2011.

[27] CableLabs, "Cable Labs perspective on LTE-U Coexistence with Wi-Fi and Operational Modes for LTE-U," 3GPP RAN1 standard contribution - RWS-140004, Sophia Antipolis, France, Jun. 2014.

[28] Charles Audet and John E Dennis Jr. Mesh adaptive direct search algorithms for constrained optimization. *SIAM Journal on optimization*, 17(1):188–217, 2006.

[29] D. Feng, L. Lu, Y. Yuan-Wu, G. Ye Li, S. Li, and G. Feng, "Device-to-device communications in cellular networks," IEEE Communications Magazine, vol. 52, no. 4, pp. 49–55, Apr. 2014.

[30] D Kazi, M.S Huq, Shahid Mumtaz, Jonathan Rodriguez, Paulo Marques, Bismark Okyere, Valerio Frascolla, " Enhanced C-RAN using D2D Network", IEEE Communication Magazine , Dec 2016.

[31] D. Goodman, S.A. Grandhi, and R. Vijayan, "Distributed dynamic channel assignment schemes," Proc. IEEE 43rd Vehicular Technology Conference (VTC1993-Spring), May 1993.

[32] Dongning Guo, "LTE in Unlicensed Spectrum," Communication Theory Workshop, Curacao, May. 2014.

[33] Doppler, K.; Rinne, M.; Wijting, C.; Ribeiro, C.B.; Hugl, K., "Device-to-device communication as an underlay to LTE-advanced networks," in Communications Magazine, IEEE , vol.47, no.12, pp.42-49, Dec. 2009.

[34] E. Almeida, A. M. Cavalcante, R. C. D. Paiva, F. S. Chaves, F. M. Abinader, R. D. Vieira, S. Choudhury, E. Tuomaala, and K. Doppler, "Enabling lte/wifi coexistence by lte blank subframe allocation," in 2013 IEEE International Conference on Communications (ICC), June 2013, pp. 5083–5088.

[35] Kazi Mohammed Saidul Huq, Shahid Mumtaz, Jonathan Rodriguez and Rui Aguiar, "Green HetNet CoMP: Energy Efficiency Analysis and Optimization" in IEEE Transactions on Vehicular Technology, 2014

[36] Ericsson Mobility Report, November 2015, http://www.ericsson.com/res/docs/2015/mobility-report/ericsson-mobility-report-nov-2015.pdf

[37] F. Bouali, K. Moessner, and M. Fitch, "A Context-aware User-driven Framework for Network Selection in 5G multi-RAT Environments," in IEEE 84th Vehicular Technology Conference (VTC Fall), September 2016.

[38] F. Chaves, E. Almeida, R. Vieira, A. Cavalcante, F. Abinader, S. Choudhury, and K. Doppler, "LTE UL Power Control for the Improvement of LTE/Wi-Fi Coexistence," in Proc. IEEE Vehic. Technol.

Conf. (VTC), Sep. 2013, pp. 1–6.

[39] Mudassar Ali, Shahid Mumtaz: "Joint User-Association and Power Allocation for Licensed & Unlicensed Spectrum in 5G Heterogeneous Networks" *EURASIP Journal on Wireless Communications and Networking* special issue on 'Dynamic Spectrum Access and Cognitive techniques for 5G' (Submitted),

[40] Feilu Liu; at el, "A framework for femtocells to access both licensed and unlicensed bands," in Modelling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2011 International Symposium on, vol., no., pp.407-411, 9-13 May 2011

[41] Fodor, G.; Dahlman, E.; Mildh, G.; Parkvall, S.; Reider, N.; Miklós, G.; Turányi, Z., "Design aspects of network assisted device-to-device communications," in IEEE Communications Magazine, vol.50, no.3, pp.170-177, March 2012.

[42] G. Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, March 2000.

[43] F. Bouali, K. Moessner, and M. Fitch, "*A Context-aware User-driven Strategy to Exploit Offloading and Sharing in Ultra-dense Deployments*," Submitted to ICC 2017.

[44] G. Fodor et al. "Intercell Interference Coordination in OFDMA Networks and in 3GPP Long Term Evolution Systems", IEEE JSAC, vol.4, no.7, pp.445-453, August 2009.

[45] G. Lim et al., "Energy-efficient resource allocation for OFDMA-based multi-RAT networks", IEEE Transactions on Wireless Communications, vol.13, no.5, pp 2696–2705, May 2014.

[46] G. Yu, "Multi-Objective Energy-Efficient Resource Allocation for Multi-RAT Heterogeneous Networks", IEEE Journal on Selected Areas in Communications, vol.33, n.10, pp.2118-2127, October 2015.

[47] H. M. Kwon and T. G. Birdsall, "Channel capacity in bits per joule," IEEE J. Ocean. Eng., vol. 11, no. 1, pp. 97–99, Jan. 1986.

[48] H. Min, J. Lee, S. Park, and D. Hong, "Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks," IEEE Trans. Wireless Communications., vol. 10, no. 12, pp. 3995–4000, Dec. 2011.

[49] HiSilicon, "Scenarios, spectrum considerations and preliminary assessment results of U-LTE," 3GPP RAN1 standard contribution - RWS-140026, Sophia Antipolis, France, Jun. 2014.

[50] Hongguang Sun; Wildemeersch, M.; Min Sheng; Quek, T.Q.S., "D2D Enhanced Heterogeneous Cellular Networks With Dynamic TDD," in Wireless Communications, IEEE Transactions on , vol.14, no.8, pp.4204-4218, Aug. 2015.

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

[52] Huawei, "U-LTE: Unlicensed Spectrum Utilization of LTE," 2014. [Online]. Available: www.huawei.com/ilink/en/download/HW 327803

[53] Hyunkee Min; Jemin Lee; Sungsoo Park; Daesik Hong, "Capacity Enhancement Using an Interference Limited Area for Device-to-Device Uplink Underlaying Cellular Networks," in Wireless Communications, IEEE Transactions on , vol.10, no.12, pp.3995-4000, December 2011.

[54] I.P. Belikaidis et al., "Context-aware Radio Resource Management Below 6 GHz for Enabling Dynamic Channel Assignment in the 5G era", EURASIP Journal on Wireless Communications and Networking (EURASIP JWCN), submitted.

[55] Intel, "LTE System Deployment and Performance in Unlicensed Bands," Tech Talk at IEEE ComSoc Event on LTE-U, Nov. 2014.

[56] J. B. Barbanel, The geometry of efficient fair division. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[57] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-Device Communication in LTE-Advanced Networks: A Survey," IEEE Communications Surveys Tutorials, vol. 17, no. 4, pp. 1923–1940, Fourth quarter 2015.

[58] J. M. Robertson, and W. A. Webb, Cake-cutting algorithms - be fair if you can. Natick, MA: A K Peters, 1998.

[59] J. Oh and Y. Han, "Cell selection for range expansion with almost blank subframe in heterogeneous networks," in IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications, pp.653-657, 9-12 Sept. 2012

[60] Jiang, Y.; Liu, Q.; Zheng, F.; Gao, X.; You, X., "Energy Efficient Joint Resource Allocation and Power Control for D2D Communications," in Vehicular Technology, IEEE Transactions on , vol. PP, no.99, pp.1-1, 2016

[61] K. Doppler, M. P. Rinne, P. Janis, C. Ribeiro, and K. Hugl, "Device-to-Device Communications; Functional Prospects for LTE-Advanced Networks". IEEE International Conference on Communications Workshops, 2009, Jun. 2009, pp.1– 6.

[62] Kae Won Choi; Zhu Han, "Device-to-Device Discovery for Proximity-Based Service in LTE-Advanced System," in Selected Areas in Communications, IEEE Journal on , vol.33, no.1, pp.55-66, Jan. 2015.

[63] L. Angrisani, M. Bertocco, D. Fortin, and A. Sona, "Experimental Study of Coexistence Issues Between IEEE 802.11b and IEEE 802.15.4 Wireless Networks," IEEE Trans. Instrumentation and Measurement, vol. 57, no. 8, pp. 1514–1523, Aug 2008.

[64] L. Juho, K. Younsun, L. Hyojin, B. Loong Ng, D. Mazzarese, L. Jianghua, X. Weimin, and Z. Yongxing. (2012). Coordinated multipoint transmission and reception in LTE-advanced systems. IEEE Communications Magazine, 50 (11), 44-50.

[65] L. Polak, O. Kaller, L. Klozar, and J. Prokopec, "Exploring and Measuring the Co-Existence between LTE and DVB-T2-Lite Services," in Proc. Int. Conf. on Telecommuni.and Signal Processing (TSP), Jul. 2013, pp. 316–320.

[66] Laya, A.; Kun Wang; Widaa, A.A.; Alonso-Zarate, J.; Markendahl, J.; Alonso, L., "Device-to-device communications and small cells: enabling spectrum reuse for dense networks," in Wireless Communications, IEEE , vol.21, no.4, pp.98-105, August 2014.

[67] Li Wang; Huaqing Wu, "Fast Pairing of Device-to-Device Link Underlay for Spectrum Sharing With Cellular Users," in Communications Letters, IEEE , vol.18, no.10, pp.1803-1806, Oct. 2014.

[68] LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (3GPP TS 36.300 version 13.3.0 Release 13)

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

[70] Lu Yang; Wei Zhang; Shi Jin, "Interference Alignment in Device-to-Device LAN Underlaying Cellular Networks," in Wireless Communications, IEEE Transactions on , vol.14, no.7, pp.3715-3723, July 2015.

[71] M. B. Çelebi, İsmail Güvenc¸, H. Arslan, and K. A. Qaraqe, "Interference suppression for the LTE uplink," Physical Commun., vol. 9, no. 0, pp. 23 – 44, 2013.

[72] M. Beluri, E. Bala, Y. Dai, R. Di Girolamo, M. Freda, J. Gauvreau, S. Laughlin, D. Purkayastha, and A. Touag, "Mechanisms for LTE Coexistence in TV White Space," in Proc. IEEE Int. Symp. on Dynamic Spectrum Access Networks (DYSPAN), Oct. 2012, pp. 317–326.

[73] M. E. Şahin, İsmail Güvenc, and H. Arslan, "An iterative interference cancellation method for co-channel multicarrier and narrowband systems," Physical Commun., vol. 4, no. 1, pp. 13 – 25, 2011.

[74] M. Ismail et al., "A Distributed Multi-Service Resource Allocation Algorithm in Heterogeneous Wireless Access Medium", IEEE Journal on Selected Areas in Communications, vol.30, n.2, pp.425-432, February 2012.

[75] M. J. Yang, S. Y. Lim, H. J. Park, and N. H. Park, "Solving the data overload: Device-to-device bearer control architecture for cellular data offloading," IEEE Vehicular Technology Magazine, vol. 8, no. 1, pp. 31–39, Mar. 2013.

[76] M.Bennis, at el "When cellular meets WiFi in wireless small cell networks" IEEE Commun. Mag, 2013

[77] M.Charitos and G.Kalives "Heterogeneous hybrid vehicular WiMAX-WiFi network for in-tunnel surveillance implementations", IEEEs ICC June 2013

[78] Maghsudi, S.; Stanczak, S., "Hybrid Centralized-Distributed Resource Allocation for Device-to-Device Communication Underlaying Cellular Networks," in Vehicular Technology, IEEE Transactions on, vol. PP, no.99, pp.1-1, 2016.

[79] Moulin, Fair division and collective welfare. The MIT Press, 2003.

[80] Mustafa, H.A.; Shakir, M.Z.; Imran, M.A.; Imran, A.; Tafazolli, R., "Coverage Gain and Device-to-Device User Density: Stochastic Geometry Modeling and Analysis," in Communications Letters, IEEE , vol.19, no.10, pp.1742-1745, Oct. 2015.

[81] N. Golmie, N. Chevrollier, and O. Rebala, "Bluetooth and WLAN coexistence: challenges and solutions," IEEE Wireless Communi., vol. 10, no. 6, pp. 22–29, Dec 2003.

[82] Nadisandla Rupasinghe "WiFi and LTE Coexistence in Unlicensed Specturm"

[83] NOKIA, "LTE in Unlicensed Spectrum: European Regulation and Co-Existence Considerations," 3GPP RAN1 standard contribution - RWS-14002, Sophia Antipolis, France, Jun. 2014.

[84] NOKIA, "Nokia LTE for Unlicensed Spectrum," 2014, White Paper.

[85] P. Mach, Z. Becvar, and T. Vanek, "In-Band Device-to-Device Communication in OFDMA Cellular Networks: A Survey and Challenges," IEEE Communications Surveys Tutorials, vol. 17, no. 4, pp. 1885–1922, Fourthquarter 2015.

[86] P. Rathod, A. Karandikar, and A. Sahoo, "Facilitating Non-collocated coexistence for WiFi and 4G wireless networks," in Proc. IEEE Conf. Local Comp. Netw. (LCN), Oct. 2012, pp. 1–9.

[87] Q. Liu, X. Li, W. Xu, and D. Zhang, "Empirical Analysis of Zigbee and WiFi Coexistence," in Proc., Int. Conf. Innovative Design and Manufacturing (ICIDM), Aug 2014, pp. 117–122.

[88] Qualcomm presentation, "Extending the benefits of LTE Advanced to Unlicensed Spectrum," https://www.qualcomm.com/media/documents/files/extending-the-benefits-of-lte-advanced-to-unlicensed-spectrum.pdf, Apr. 2014.

[89] Qualcomm white paper, "LTE in Unlicensed Spectrum: Harmonious Coexistence with WiFi," https://www.qualcomm.com/media/documents/files/extending-the-benefits-of-lte-advanced-to-unlicensed-spectrum.pdf, Jun. 2014.

[90] R. Matsukawa, T. Obara, and F. Adachi. "A dynamic channel assignment scheme for distributed antenna networks." Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th. IEEE, 2012.

[91] R. Yin, G. Yu, A. Maaref, and G. Li, "A framework for co-channel interference and collision probability tradeoff in lte licensed-assisted access (laa) networks," IEEE Transactions on Wireless Communications, vol. PP, no. 99, pp. 1–1, 2016.

[92] R1-132861. Final Report of 3GPP TSG RAN WG1 #73 v1.0.0, August 2013

[93] Ratasuk, at el,"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

[94] Ratasuk, M. Uusitalo, N. Mangalvedhe, A. Sorri, S. Iraji, C.Wijting, and A. Ghosh, "License-exempt LTE deployment in heterogeneous network," in Proc. Int. Symp. Wireless Commun. Sys. (ISWCS), Aug. 2012.

[95] S. Boyd, L. Vandenberghe, Stephen Boyd, and Lieven Vandenberghe, Convex Optimization. Cambridge University Press, 2004, p. 730.

[96] S. Brams, and A. Taylor, Fair division: from cake-cutting to dispute resolution. Cambridge, U.K.: Cambridge Univ. Press, 1996.

[97] S. Deb et al., "Algorithms for Enhanced Inter-Cell Interference Coordination (eICIC) in LTE HetNets", IEEE/ACM Transactions on Networking, vol.22, n.1, pp.137-150, March 2013.

[98] S. Glisic, and B. Lorenzo. Advanced wireless networks: cognitive, cooperative & opportunistic 4G technology. John Wiley & Sons, 2009.

[99] S. Mishra, R. Brodersen, S. Brink, and R. Mahadevappa, "Detect and Avoid: An Ultra-Wideband/WiMAX Coexistence Mechanism," IEEE Commun. Mag., vol. 45, no. 6, pp. 68–75, Jun. 2007.

[100] S. Mumtaz, K. M. S. Huq, and J. Rodriguez, "Direct mobile-to-mobile communication: Paradigm for 5G," IEEE Wireless Communication, vol. 21, no. 5, pp. 14–23, Oct. 2014.

[101] S. Pollin, I. Tan, B. Hodge, C. Chun, and A. Bahai, "Harmful Coexistence Between 802.15.4 and 802.11: A Measurement-based Study," in Proc Int. Conf. on Cognitive Radio Oriented Wireless Networks and Communi., May 2008, pp. 1–6.

[102] S. Vassaki, A. Georgakopoulos, F. Miatton, K. Tsagkaris and P. Demestichas, "Interference and QoS aware channel segregation for heterogeneous networks: A preliminary study," 2016 European Conference on Networks and Communications (EuCNC), Athens, June 2016, pp. 195-199.

[103] S. Zacharias, T. Newe, S. O'Keeffe, and E. Lewis, "Coexistence measurements and analysis of IEEE 802.15.4 with Wi-Fi and bluetooth for vehicle networks," in Proc. IEEE Conf. Telecommun. (ITST), Nov. 2012.

[104] S.Dimatteo, "Celluar traffic offloading through WiFi Network", in IEEE MASS. Oct 2011

[105] S.Hajomhammad at el, "Unlicensed spectrum splitting between Femtocell and WiFi", in IEEE ICC, June 2013

[106] S.-J. J. Chen and C. L. Hwang, Fuzzy Multiple Attribute Decision Making: Methods and Applications, M. J. Beckmann and W. Krelle, Eds. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1992.

[107] S.-J. Kim, I. Cho, Y.-K. Kim, and C.-H. Cho, "A two-stage dynamic channel assignment scheme with graph approach for dense femtocell networks," IEICE Trans. Commun., vol.E97-B, no.10, pp.2222– 2229, Oct. 2014.

[108] Sagari,at el., "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

[109] Salem, M., Adinoyi, A., Rahman, M., Yanikomeroglu, H., Falconer,D.,& Kim, Young-Doo (2009). Fairness-aware joint routing and scheduling in ofdma-based cellular fixed relay networks. IEEE International Conference on Communication 2009 (pp. 1–6), June 2009.

[110] Seyedmehdi, S.H.; Boudreau, G., "An Efficient Clustering Algorithm for Device-to-Device Assisted Virtual MIMO," in Wireless Communications, IEEE Transactions on , vol.13, no.3, pp.1334-1343, March 2014.

[111] SONY, "Requirements and Coexistence Topics for LTE-U," 3GPP RAN1 standard contribution - RWS-140010, Sophia Antipolis, France, Jun. 2014.

[112] SPEED-5G Public Deliverable, "D5.1: MAC approaches with FBMC and simulation results", ICT-671705, H2020-ICT-2014-2, June 2016.

[113] Sunghyun, C., Jang, E.W.,&Cioffi, J. M. (2009). Handover in multihop cellular networks. IEEE Communication Magazine, 47(7), 64–73.

[114] T. Elkourdi et al., "Cross-Layer Optimization for Opportunistic Multi-MAC Aggregation", IEEE Vehicular Technology Conference (VTC Spring), pp.1-5, June 2013.

[115] Tang, R.; Zhao, J.; Qu, H., "Joint optimization of channel allocation, link assignment and power control for device-to-device communication underlaying cellular network," in Communications, China , vol.12, no.12, pp.92-100, December 2015.

[116] Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description; Stage 2 (Release 11), 3GPP TS 36.300 V11.7.0, Sep. 2013.

[117] Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description; Stage 2 (Release 11), 3GPP TS 36.300 V11.7.0, Sep. 2013.

[118] Texas Instruments, "LTE Operation in Unlicensed Spectrum," 3GPP RAN1 standard contribution - RWS-140012, Sophia Antipolis, France, Jun. 2014.

[119] U. Herzog et al., "Quality of service provision and capacity expansion through extended-DSA for 5G", 2016 European Conference on Networks and Communications (EuCNC), Athens, June 2016, pp. 200-204.

[120] Virginia Torczon. On the convergence of pattern search algorithms. *SIAM Journal on* optimization, 7(1):1–25, 1997.

[121] W. Yuan, X. Wang, and J. P. M. G. Linnartz, "A Coexistence Model of IEEE 802.15.4 and IEEE 802.11b/g," in Proc. IEEE Symp. on Commun. Vehic. Techno. in the Benelux, Nov. 2007, pp. 1–5.

[122] Wei Yu and R. Lui, "Dual Methods for Nonconvex Spectrum Optimization of Multicarrier Systems," IEEE Transactions on Communications, vol. 54, no. 7, pp. 1310-1322, Jul. 2006.

[123] Wei, L.; Hu, R.Q.; Qian, Y.; Wu, G., "Energy Efficiency and Spectrum Efficiency of Multihop Device-to-Device Communications Underlaying Cellular Networks," in Vehicular Technology, IEEE Transactions on, vol.65, no.1, pp.367-380, Jan. 2016.

[124] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," IEEE Communications Magazine, vol. 52, no. 4, pp. 40–48, Apr. 2014

[125] Xingqin Lin; Heath, R.W.; Andrews, J.G., "The Interplay Between Massive MIMO and Underlaid D2D Networking," in Wireless Communications, IEEE Transactions on , vol.14, no.6, pp.3337-3351, June 2015.

[126] Y. Aumann, Y. Dombb, and A. Hassidim, "Computing socially efficient cake divisions," in Proc. Int. Conf. AAMAS, Saint Paul, MN, USA, 2013, pp. 343-350.

[127] Y. Choi et al., "Joint Resource Allocation for Parallel Multi-Radio Access in Heterogeneous Wireless Networks", IEEE Transactions on Wireless Communications, vol.9, n.11, pp.3324-3329, November 2010.

[128] Y. Furuya, and Y. Akaiwa. "Channel segregation, a distributed adaptive channel allocation scheme for mobile communication systems." IEICE Transactions on Communications 74.6 (1991): 1531-1537.

[129] Y. Wu et al., "Device-to-device meets LTE-unlicensed," in IEEE Communications Magazine, vol. 54, no. 5, pp. 154-159, May 2016

[130] Y. Zhou et al., "An overview on intercell interference management in mobile cellular networks: From 2G to 5G", IEEE International Conference on Communication Systems (ICCS), pp.217-221,

November 2014.

[131] Yan Chen, Shunqing Zhang, Shugong Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," IEEE Commun. Mag., vol. 49, no. 6, pp. 30–37, Jun. 2011.

[132] Yong Niu; Chuhan Gao; Yong Li; Li Su; Depeng Jin; Vasilakos, A.V., "Exploiting Device-to-Device Communications in Joint Scheduling of Access and Backhaul for mmWave Small Cells," in Selected Areas in Communications, IEEE Journal on, vol.33, no.10, pp.2052-2069, Oct. 2015.

[133] Z. Meng, Y. Hai, F. Yulin, and C. Yongyu, "Coexistence Studies on interference Performance between Mobile Satellite System and LTE Network," China Communications, Vol. 10, no. 7, pp. 1-11, July, 2013.

[end of document]