Coordinated Control and Spectrum Management for 5G Heterogeneous Radio Access Networks

Grant Agreement No. : 671639
Call: H2020-ICT-2014-2

Deliverable D3.1

First report on physical and MAC layer modelling and abstraction

Version:	1.0
Due date:	30.06.2016
Delivered date:	02.07.2016
Dissemination level:	PU

The project is co-funded by

Authors

Antti Anttonen (Editor-in-Chief), Tao Chen, Tapio Suihko, Aarne Mämmelä (VTT); Sundar Daniel Peethala, Nidal Zarifeh (UDE); Furqan Ahmed, Junquan Deng, Ragnar Frej-Hollanti, Sergio Lembo, Olav Tirkkonen (AALTO); Antonio Cipriano, Dorin Panaitopol (TCS); Per Kreguer, Akhila Rao, Rebecca Steinert (SICS); Chia-Yu Chang (EUR); Roberto Riggio, Shah Nawaz Khan (CNET); Mariana Goldhamer (4GC); Pawel Kryszkiewicz (PUT); Fang-Chun Kuo (EICT); George Agapiou (OTE); Dimitri Marandin, Yi Yu (CMA);

Coordinator

Dr. Tao Chen

VTT Technical Research Centre of Finland Ltd

Tietotie 3

02150, Espoo

Finland

Email: tao.chen@vtt.fi

Disclaimer

Acknowledgement

This report is funded under the EC H2020 5G-PPP project COHERENT, Grant Agreement No. 671639.

Version history

Version	Date	Remarks
0.1	5.4.2016	Draft ToC created and content from the intermediate milestone report M3.1 added.
0.2	13.6.2016	Content added to Sections 1-6.
0.3	29.6.2016	Revised contributions and Annexes added based on review comments.
1.0	30.6.2016	Final version

Executive summary

This deliverable presents the first results on physical and medium access control layer abstraction methods to enhance coordination and control capabilities of 5G heterogeneous networks. Efficient control and coordination of upcoming 5G heterogeneous mobile networks will be critical to achieve the target values of key performance indicators. One of the main problems is to maximize the use of available global information in making optimal resource control decisions while minimizing the control overhead and effects of having outdated information in a scalable network. A key concept in solving this fundamental problem is to use a sophisticated network abstraction framework. Deriving and developing such abstractions for coordinated network control is the main focus point in COHERENT project.

This deliverable contains the first representable results from the initial studies on physical and medium access control layer modelling and abstraction which will be completed in a follow-up deliverable D3.2. The underlying idea behind the work is to find new ways to provide and utilize diversified information about the global network state in particular for resource management tasks which leverage the centralized control and coordination concept.

An overview of network graphs, control paradigms, abstraction principles, standardization activities, and some key challenges is first provided to which the developed network graph concept should respond. Next, the COHERENT system model is outlined and the target abstraction framework is introduced. The work of WP2 is closely taken into account, namely the key performance indicators of deliverable D2.1 and the network architecture developed in deliverable D2.2.

The relevant metrics and measurements that are used to generate the COHERENT network graph are then investigated. Subsequently, some new ideas to the existing standards are introduced. Moreover, selected interesting control application examples are evaluated where the network graph concept can be applied. Specifically, we investigate LTE downlink performance, channel quality indicators for WiFi system, distributed and coordinated multi-antenna systems, inter-node interference modelling, load balancing, cognitive radios, device-to-device communications, mobility management, and coverage extension. Finally, some important remarks and conclusions are made on the investigated network graphs along with indications on how the proposed network graph concept can be used to address the defined problems.

Table of contents

Executive summary 4
List of abbreviations 8
List of figures 14
List of tables 15
1. Introduction 17
1.1 Objectives and motivation 17
1.2 Definition of abstraction and related terms 17
1.3 Structure of the deliverable 18
2. Background and challenges 19
2.1 Network graphs 19
2.2 Control paradigms in heterogeneous networks 19
2.2.1 Introduction 19
2.2.2 Control functionality 20
2.2.3 Control utility 21
2.2.4 Control centralization degree 21
2.2.5 Control time scales 22
2.2.6 Control protocol structure 22
2.2.7 Summary of control challenges 22
2.3 Abstracting wireless systems 23
2.3.1 Background 23
2.3.2 Promoting analytic understanding of performance of stochastic wireless systems 24
2.3.3 Simplification of system level simulations 24
2.3.4 Simplification of control procedures for RRM 24
2.3.5 Remaining challenges 24
2.4 D2D channel modelling 25
2.4.1 3GPP activities on D2D channel modelling 26
2.4.2 D2D channel measurements in EU/FP7 ABSOLUTE project 30
2.4.3 Conclusion of D2D channel modelling review 31
2.5 Relevant standardization activities for supporting a central coordinator in LTE standards 31
2.6 Open problems 32
3. COHERENT system model and abstraction framework 34
3.1 COHERENT network architecture 34
3.2 COHERENT network graph and abstraction framework 35
3.2.1 Nodes 36
3.2.2 Edges 36
3.2.3 Abstracted network graph 36
3.2.4 Overview of abstraction procedure 37
4. Measurements and metrics to generate COHERENT network graph 40
4.1 Available measurements and metrics 40
4.1.1 PHY layer – LTE 40
4.1.2 PHY layer – WiFi 48
4.1.3 MAC layer – LTE 49
4.1.4 MAC layer – WiFi 49

4.2 Possible new metrics and approaches for constructing network graphs 55
4.2.1 Probabilistic input to network graphs 55
4.2.2 Coupling loss 56
4.3 Management of measurements and metrics 56
5. Evaluation of abstraction opportunities in selected COHERENT control applications 58
5.1 LTE downlink PHY transmission model and link-level performance 58
5.1.1 LTE downlink PHY transmitter 58
5.1.2 LTE downlink PHY receiver 60
5.1.3 LTE PHY link-level simulations 63
5.2 Modelling of channel quality indicators for WiFi 65
5.2.1 Problem description 65
5.2.2 Proposed solution 65
5.2.3 Evaluation 69
5.2.4 Final discussion and summary 72
5.3 Distributed antenna system/massive MIMO 72
5.3.1 Problem description 72
5.3.2 Modelling 72
5.3.3 Measurement setup 75
5.3.4 Results 76
5.3.5 Usage of abstraction 77
5.4 Coordinated multipoint transmission 77
5.4.1 Joint processing 78
5.4.2 Coordinated scheduling and beamforming 78
5.5 Dynamic inter-node OOB interference modelling 78
5.5.1 Inter-system interference characterization 78
5.5.2 Influence of PHY modulation techniques on inter-node interference 80
5.5.3 Proposed solution and results 81
5.6 Load balancing 84
5.6.1 Enhanced distributed load balancing 84
5.6.2 Load balancing in energy-limited HetNets 93
5.7 Cognitive radio in HetNets 99
5.8 Device to device communications 99
5.8.1 Problem description 99
5.8.2 COHERENT system model for D2D network 100
5.8.3 Abstraction approach for D2D management 101
5.8.4 Proposed graph-based solution for D2D network management 102
5.8.5 Evaluation of graph-based interference coordination in D2D network 104
5.9 D2D relaying for coverage extension 105
5.9.1 Problem description 105
5.9.2 State-of-the-art of D2D communications in 3GPP 107
5.9.3 Proposed solution 120
5.10 Mobility management in HetNets 121
5.10.1 Problem description 121
5.10.2 Solution approach 122
5.11 Mobility study for high speed platforms 124

5.11.1 Problem description 124
5.11.2 Proposed system 125
5.11.3 Proposed solution 127
6. Conclusions 130
6.1 Summary of main outcome 130
6.1.1 Network graphs 130
6.1.2 Control applications 131
6.2 Future work needed in D3.2 133
References 135
Annex: Objectives mapping of WP3 146

List of abbreviations

3GPP	Third Generation Partnership Project
5G	Fifth Generation
ABS	Almost Blank Subframe
ACIR	Adjacent Channel Interference Ratio
ACK	Acknowledge
ACLR	Adjacent Channel Leakage Ratio
ACS	Adjacent Channel Selectivity
AIF	Antenna Interface
ANDSF	Access Network Discovery and Selection Function
AoA	Angle of Arrival
AoD	Angle of Departure
AP	Access Point
API	Application Programming Interface
ARQ	Automatic Repeat request
AWGN	Additive White Gaussian Noise
BBU	Baseband Unit
BCH	Broadcast Channel
BER	Bit Error Rate
BLER	Block Error Rate
BPSK	Binary Phase Shift Keying
BS	Base Station
BSR	Buffer Status Report
C3	Central Controller and Coordinator
CA	Carrier Aggregation
CAP	Consistency Availability Partition
CB	Coordinated Beamforming
CCTV	Closed Circuit Television
CDF	Cumulative Distribution Function
CL	Coupling Loss
CN	Cognitive Networks
CoMP	Coordinated Multipoint
CPRI	Common Public Radio Interface
CRE	Cell Range Expansion
CRS	Cell-Specific Reference Signal
CS	Coordinated Scheduling
CSG	Closed Subscriber Group

CSI	Channel State Information
CSI-RS	Channel State Information Reference Signal
CSMA/CA	Carrier Sense Multiple Access/ Collision Avoidance
CQI	Channel Quality Indicator
CQM	Channel Quality Map
CW	Continuous Wave
D2D	Device to Device
DA2GC	Broadband Direct Air to Ground Communications
DAS	Distributed Antenna System
DCI	Downlink Control Information
DCS	Dynamic Cell Selection
DFN	Direct Frame Number
DFT	Discrete Fourier Transform
DL	DownLink
DSSS	Direct Sequence Spread Spectrum
EDP	Energy Depletion Probability
eMBB	enhanced Mobile Broadband
EMA	Exponential Moving Average
eNB or eNodeB	Evolved Node B
EPC	Evolved Packet Core
ETSI	European Telecommunications Standards Institute
E-UTRA	Evolved UMTS Terrestrial Radio Access
eV2X	enhanced Vehicle-to-X communication
EVM	Error Vector Magnitude
EWMA	Exponentially Weighted Moving Average
FBMC	Filter Bank Multicarrier
FDD	Frequency Division Duplexing
FFT	Fast Fourier Transform
GPS	Global Positioning System
GSM	Global System for Mobile Communication
H2H	Human to Human
HARQ	Hybrid Automatic Repeat reQuest
HETNET, HetNet	Heterogeneous Network
HMN	Heterogeneous Mobile Network
HO	HandOver
ICIC	Intercell Interference Coordination
IDFT	Inverse Discrete Fourier Transform
IoT	Internet of Things

ISD	Inter Site Distance
ISO	International Organization for Standardization
JT	Joint Transmission
LAA	Licenced Assisted Access
LAN	Local Area Network
LBT	Listen Before Talk
LC	Logical Channel
LCG	Logical Channel Group
LE	Licence Exempt
LNV	Local Network View
LOS	Line Of Sight
LSB	Least Significant Bits
LTE	Long Term Evolution
LTE-A	LTE-Advanced
LVAP	Light Virtual Access Point
M2M	Machine to Machine
MAC	Medium Access Control
MBS	Macro Base Station
MCPTT	Mission Critical Push to Talk
MCS	Modulation and Coding Scheme
MF	Monitoring Function
MIB	Master Information Block
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity
MoM	Method of Moments
NACK	Negative Acknowledge
NGMN	Next Generation Mobile Networks
NLOS	Non Line Of Sight
NOS	Network Operating System
NZP	Non-Zero Power
OBSAI	Open Base Station Architecture Initiative
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OOB	Out-of-Band
PDR	Packet delivery Ratio
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PDSCH	Physical Downlink Shared Channel

PDU	Protocol Data Unit
PER	Packet Error Rate
PHY	Physical Layer
PLCP	Physical Layer Convergence Protocol
PLMU	Portable Land Mobile Unit
PMF	Probability Mass Function
PMR	Professional Mobile Radio
PRACH	Physical Random Access Channel
PRB	Physical Resource Block
ProSe	Proximity Services
PS	Public Safety
PSBCH	Physical Sidelink Broadcast CHannel
PSCCH	Physical Sidelink Control Channel
PSD	Power Spectral Density
PSDCH	Physical Sidelink Discovery CHannel
PSS	Primary Synchronization Signal
PSSCH	Physical Sidelink Shared Channel
PSSS	Primary Sidelink Synchronization Signal
PU	Primary User
PUSCH	Physical Uplink Shared Channel
QAM	Quadrature Amplitude Modulation
QoE	Quality of Experience
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RAN	Radio Access Network
RAT	Radio Access Technology
RB	Resource Block
RCPI	Received Channel Power Indicator
RI	Rank Indicator
RLC	Radio Link Control
RNTP	Relative Narrowband Transmit Power
RNTI	Radio Network Temporary Identifier
RP	Resource Pool
RRC	Radio Resource Control
RRH	Remote Radio Head
RRM	Radio Resource Management
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality

RSSI	Received Signal Strength Indicator
RT	Radio Transceiver
RTC	Real-Time Controller
RTP	Real Transmission Point
RX	Receiver
SBCH	Sidelink Broadcast Channel
SBS	Small Base Station
SC-FDMA	Single Carrier- Frequency Division Multiple Access
SCI	Sidelink Control Information
SDK	Software Development Kit
SDM	Space Division Multiplexing
SDN	Software Defined Networking
SD-RSRP	Sidelink Discovery Reference Signal Received Power
SDU	Service Data Unit
SFN	System Frame Number
SIB	System Information Block
SIMO	Single Input Multiple Output
SINR	Signal to Interference plus Noise Ratio
SIR	Signal to Interference Ratio
SISO	Single Input Single Output
SNR	Signal to Noise Ratio
SL	SideLink
SL-BCH	Sidelink Broadcast CHannel
SL-DCH	Sidelink Discovery CHannel
SL-SCH	Sidelink Shared Channel
SLSS	Sidelink Synchronization Signals
SNV	Slice-specific Network View
SON	Self Organizing Network
SRS	Sounding Reference Signal
S-RSRP	Sidelink Reference Signal Received Power
SSM	Signal Strength Map
SSRS	Sidelink Sounding Reference Signal
SSS	Secondary Synchronization Signal
SSSS	Secondary Sidelink Synchronization Signal
STA	Station
STCH	Sidelink Traffic Channel
SU	Secondary User
TAS	Transmit Antenna Selection

TBS	Transport Block Size
TDD	Time Division Duplexing
TFT	Traffic Flow Template
TN	Transport Node
T-RPT	Time Resource Pattern of Transmission
TTT	Time to Trigger
TX	Transmission
UE	User Equipment
UL	UpLink
UMTS	Universal Mobile Telecommunications System
USIM	Universal Subscriber Identity Module
V2V	Vehicle to Vehicle
VDSL	Very high rate Digital Subscriber Line
WCDMA	Wideband Code Division Multiple Access
WiFi	Wireless Fidelity
WLAN	Wireless Local Area Network
WP	Work Package
WTP	Wireless Termination Point
ZP	Zero Power

List of figures

Figure 2-1 Overview of existing control paradigms in HetNets (user-centric coverage view) 20
Figure 2-2 Examples of abstraction levels and procedures in different kind of applications adopted from [52, 54, 55, 57] 23
Figure 2-3 Geometry for d_1 and d_2 path-loss model for B4 NLoS 28
Figure 2-4 Results on WINNER II channel model B4 29
Figure 2-5 Results on WINNER II channel model C2 30
Figure 3-1 COHERENT Architecture 34
Figure 3-2 A downlink network graph 35
Figure 3-3 Abstracted cellular network graph in a homogeneous network, with resource assignments represented by colours 37
Figure 3-4 Network abstraction in a heterogeneous network, with two macro-cells and a small cell used for offloading and range extension 37
Figure 3-5 Abstraction flow for the network graph in C3 instance 38
Figure 3-6 Abstraction flow for the network graph in C3 instance 39
Figure 4-1 An example of RSSI Map 52
Figure 4-2 An example of link statistics map 54
Figure 4-3 An example of traffic matrix map 54
Figure 4-4 Measurement entities triplet 57
Figure 5-1 LTE downlink PHY transmitter 58
Figure 5-2 LTE downlink PHY receiver 61
Figure 5-3 Control block diagram of CRS-based channel estimation 61
Figure 5-4 MIMO receivers at UE PHY receiver 62
Figure 5-5 Test bench for PHY link-level modelling 63
Figure 5-6 Time impulse responses for different channel models 64
Figure 5-7 A conceptual overview of monitoring functions (MFs) implementing local aggregation and probabilistic modelling inside APs 69
Figure 5-8 Plot of ξ_i as a time series over a sample link, with window size of 10 s. The black overlay curve on each plot is the smoothed Bayesian estimation of the metric in the background. 70
Figure 5-9 Plot of quantile RSSI as a time series over a sample link, with window size of 10 s. The black overlay curve on each plot is the smoothed Bayesian estimation of the metric in the background. 71
Figure 5-10 Plot of F_{PDR} as a time series over a sample link, with window size of 10 s. The black overlay curve on each plot is the smoothed Bayesian estimation of the metric in the background. 71
Figure 5-11 Possible DAS setup in a stadium 72
Figure 5-12 Collocated antenna cell and distributed antenna cell 73
Figure 5-13 DLSCH transmitter processing chain adaptation for DAS [107] 74
Figure 5-14 Hardware implementation blocks [107] 75
Figure 5-15 Measurement setup 76
Figure 5-16 Constellation and EVM - Single transmit antenna case 76
Figure 5-17 Constellation and EVM – Two antennas separated by 300 cm 76
Figure 5-18 Constellation and EVM – Two antennas separated by 200 cm 77
Figure 5-19 CoMP downlink transmission schemes [108] 77
Figure 5-20 TX and RX characteristic influence on the effective interference 80
Figure 5-21 Plots of coexistence between single occupied RB and GSM carrier 83
Figure 5-22 Normalized ACIR values for interference cause by/to GSM carrier to/by single RB 83

Figure 5-23 Load distributions for unbalanced and balanced cases for a macro cell and an adjacent small cell 86
Figure 5-24 Load, risk and CRE value time series for comparable high risk scenarios. Left column balances load, right balances risk. 90
Figure 5-25 Spatial configuration used in qualitative examples 91
Figure 5-26 Spatial configuration of scenario 3 in Table 5-3..... 92
Figure 5-27 Abstracted network graph concept in a HetNet with diversified design constraints 94
Figure 5-28 Illustration of probabilistic energy graph of a centrally coordinated energy-limited SBS access tier. Energy consumption is heavily affected by the operation of the PHY layer. 95
Figure 5-29 Control procedure for centralized coordination of new user request leading to optimized load balance with respect to applied probabilistic energy graph principle..... 96
Figure 5-30 Comparison of accuracy of energy depletion abstraction methods with integral-based approach and upper bound..... 97
Figure 5-31 Comparison of energy-limited SBS connection dropping and blocking probabilities of centralized and distributed approaches for different average energy rate difference $E(\beta_j)$ of random energy arrival and departure processes..... 98
Figure 5-32 Access approach of a secondary user to available channels of an unused spectrum 99
Figure 5-33 Interference graph for D2D-assisted HMN 101
Figure 5-34 D2D management application and underlying D2D network management 102
Figure 5-35 A simple example of graph-based resource allocation for D2D..... 103
Figure 5-36 HMN with D2D-relaying feature..... 104
Figure 5-37 Simulation result of graph-based interference coordination for D2D-relaying 105
Figure 5-38 Coverage extension with UE-relay 106
Figure 5-39 Simplified 3GPP architecture for ProSe (from 3GPP) 107
Figure 5-40 MAC structure for SL communications (Fig. 4.2.1-3 in TS 36.321 [167]])..... 109
Figure 5-41 SL subframe structure with resources filled with synchronization and broadcast signals 112
Figure 5-42 Resource pool definition for SL discovery. In the figure, only subframes for UL (eligible to SL) are shown..... 114
Figure 5-43 Resource pool definition for SL communication, temporal aspect. In the figure, only subframes for UL (eligible to SL) are shown 117
Figure 5-44 Mobility management application interaction with SDN controller..... 123
Figure 5-45 Comparison between broadband direct air to ground communications satellite communications..... 124
Figure 5-46 Estimation of needed eNodeB for European coverage for DA2GC 125
Figure 5-47 LTE measurement issues related to high Speed (part I) 128
Figure 5-48 LTE measurement issues related to high Speed (part II)..... 129
Figure A-1 D2D Network graph.....**Error! Bookmark not defined.**

List of tables

Table 2-1 A1, B4 and C2 WINNER II Channel Models [193] 27
Table 4-1 RSRP measurement report mapping 41
Table 4-2 RSRQ measurement report mapping 41
Table 4-3 Table 7.2.3-1 in TS 36.213 v12.8.0, CQI index mapping for QPSK, 16-QAM, 64-QAM.. 43
Table 4-4 Table 7.2.3-2 in TS 36.213 v12.8.0, CQI index mapping for QPSK, 16-QAM, 64-QAM and 256-QAM 43
Table 4-5 Size and number of sub-bands for aperiodic CQI reporting in LTE..... 44

Table 4-6 Size and number of sub-bands for periodic CQI reporting in LTE 45
Table 4-7 UL-DL configuration for Type 2 frame (TDD), as in Table 4.2-2 in TS 36.211 v12.8.0 47
Table 4-8 Subcarrier spacing and corresponding CP as in Table 6.12-1 in TS 36.211 v12.8.0. CP length is given in sampling times 47
Table 4-9 Resource block tuples 51
Table 4-10 Southbound interface: RSSI map request 52
Table 4-11 Southbound interface: Transmission summary request 53
Table 4-12 Southbound interface (RSSI trigger request) 53
Table 4-13 Southbound interface (link statistics map request) 54
Table 4-14 Southbound interface (link statistics) 54
Table 5-1 Downlink reference signals 59
Table 5-2 Summary of results 64
Table 5-3 Proportion of requested throughput served for 3 random placements of one macro and two smaller cells, and ~1250 active/passive and mobile UEs and each of three balancing methods 92
Table 5-4 Simulation parameters for D2D network 105
Table 5-5 Mapping of channels for Sidelink in 3GPP LTE 108
Table 5-6 Modulations and coding parameters of SL PHY channels 110
Table 5-7 Traffic model for business and non-business needs for DA2GC 125
Table 5-8 Throughput requirements for DA2GC 125
Table 5-9 Capacity [Mbps] for FDD, 1 RF chain 126
Table 5-10 Capacity [Mbps] for FDD, 2 RF chain 126
Table 5-11 Capacity [Mbps] for FDD, 4 RF chain 126
Table 5-12 QoS requirements per service for DA2GC 127
Table A-1 Example of CSI-RS NZP and ZP allocation between several eNBs **Error! Bookmark not defined.**
Table A-2 Example of SSRS allocation ..... **Error! Bookmark not defined.**
Table B-1 DoW objectives to Sections in D3.1 mapping, Task T3.1 view 146
Table B-2 DoW objectives to Sections in D3.1 mapping, Task T3.2 view 146
Table B-3 DoW objectives to Sections in D3.1 mapping, Task T3.3 view 147

1. Introduction

1.1 Objectives and motivation

Efficient control and coordination of upcoming 5G heterogeneous mobile networks (HMN), interchangeably abbreviated as HetNet, will be critical to achieve the targeted key performance indicators described extensively in the COHERENT deliverable D2.1 [15]. One of the main problems is to maximize the use of available global information in making optimal resource control decisions while minimizing the control overhead and effects of having outdated information in a scalable network. A key concept in solving this fundamental problem is to use a sophisticated network abstraction framework.

This deliverable D3.1 (First report on physical and MAC layer modelling and abstraction) presents the first results achieved in WP3 of COHERENT project. The underlying idea behind the work is to find new ways to provide and utilize diversified information about the global network state. Our contribution is two-fold:

1) abstraction of relevant information from physical (PHY) layer and medium access control (MAC) layer, including that of LTE-A and WiFi technologies,

2) application of abstracted information in selected control problems of HMNs, leveraging the centralized control and coordination concept.

More specifically, the aim of WP3 is to construct abstracted network graphs used for inter-cell radio resource allocation, resource allocation for distributed multiple-input-multiple output (MIMO), device-to-device (D2D) communications, and advanced mobility management. Based on the proposed abstraction framework, the goal is to develop methods and algorithms for flexible inter-cell resource allocation and inter-cell interference coordination in HMNs. New control methods are considered for physical layer cooperation between BSs, which take advantage of the abstraction based control layer for efficient implementation of distributed and massive MIMO. User cooperation is considered to create new control methods for D2D communications, which utilize the abstraction based control layer for efficient resource allocation among multiple base stations and user devices. The proposed network graph is used to efficiently manage the user and network mobility as well as designing a control protocol suitable for current and next generation wireless standards.

The purpose of this deliverable is to present the initial results on modelling and abstraction of HMNs from the studies undertaken thus far. The complete results from all considered use-cases will be presented in a follow-up deliverable D3.2 (Final report on physical and MAC layer modelling and abstraction).

1.2 Definition of abstraction and related terms

There are many interpretations of the term abstraction in different disciplines. For instance, abstraction can mean a nonspecific concept that is not concrete or it can be a concept that connects or classifies related matters as a group. In the ISO standard [139], an abstraction is defined as follows: 1) a view of an object that focuses on the information relevant to a particular purpose and ignores the remainder of the information and 2) the process of formulating such a view. Following this definition, abstraction is understood as a technique for managing complexity of a system by carefully establishing the model description level and interactions between main system subparts and system parameters. This is accomplished by focusing only on the necessary details while hiding unnecessary information of the system from the particular subpart to represent only the essential features (e.g. classes) to a desired context.

In other words, abstraction is more about describing the transition between different abstraction hierarchy levels of a system whereas interrelated modelling is needed to describe the behaviour at each level, possible errors caused by abstraction, and achieved performance gain from using abstraction in a particular application. The transition process is often defined as deriving higher level context data from

lower context data (inductive bottom-up transition) or hiding some unnecessary high level information to simplify the development of a system subpart (reductive top-down transition). Therefore, the former transition involves induction or synthesis from particular facts whereas the latter transition involves specification or analysis of a high level system by breaking it down into subparts.

The abstractions are essentially used by coordination and control tasks for making sophisticated resource management decisions. We use the term central controller and coordinator (C3) and adopt the definition from COHERENT deliverable D2.2 [140] as follows: C3 is as a logically centralized entity which provides network-wide control. We use terms control and coordination interchangeably in this deliverable. To address the scalability problem, a logically centralized C3 can also be distributed so that the regional network graphs are shared across several physical control instances.

1.3 Structure of the deliverable

The rest of the deliverable is organized as follows. In Section 2, some relevant background information is provided on network graphs, HetNet control paradigms, abstraction principles, channel modelling, standardization activities, and some key challenges. Next, in Section 3, the COHERENT system model is outlined and the target abstraction framework is introduced. The purpose of Section 4 is to provide metrics and measurements that can be used to generate the COHERENT network graph. Subsequently, Section 5 presents some interesting case examples where the network graph concept can be applied. Finally, some important remarks and conclusions are made in Section 6.

2. Background and challenges

In this section, some relevant background information is provided on network graphs, HetNet control paradigms, abstraction principles, standardization activities, and 5G system requirements. Furthermore, some key remaining challenges that are addressed in the work are outlined.

2.1 Network graphs

Programming wireless networks requires identifying how network resources are exposed (and represented) to software modules written by developers and how software modules can affect the network state. Current state-of-the-art cellular analytics systems continuously collect per connection information such as radio resource usage, associated base stations and handoffs at network elements such as the mobility management entity (MME) and probes deployed in strategic locations. The collected information is then backhauled to centralized servers in batches and fed to the analysis engine [1, 2]. Moreover, recently cellular network analytics systems have adopted the Hadoop based framework [3, 4].

Graphs play a key role in modelling and abstracting the properties of wireless networks. A classic example of this is the frequency assignment operation modelled as a graph colouring problem. A broad survey of graph modelling techniques for wireless networks can be found in [5]. Graphs can be used to model both connectivity and channel quality. We use the generic term channel quality rather than, for example, interference because we want to encompass measurements of the channel quality different from pure interference.

Passive and/or active measurements are leveraged by several authors to derive either the conflict graph or the interference graph of a wireless network. In [6, 7] active measurements are exploited in order to study how mutual link interference affects packet delivery ratio and throughput. In [8] micro–probing (i.e. active measurements lasting few milliseconds) is used in order to detect conflicts between links. Passive interference graph construction techniques are presented in [9, 10]. WIT [11] and Jigsaw [12] are two other examples of passive interference monitoring techniques aimed at modelling cross–link interference. A framework capable of performing root cause analysis in WiFi networks is presented in [13].

Conflict graphs are leveraged in [14] to manage the effect of interference when multiple transmitters employ variable channel widths. An architecture using micro–probing to jointly address channel assignment and transmission power control is presented in [14]. Centralized scheduling is exploited in [14], in order to mitigate hidden and exposed terminals issues in WiFi–based networks.

2.2 Control paradigms in heterogeneous networks

2.2.1 Introduction

Moving from homogeneous macrocell-based networks into multi-tier heterogeneous networks (HetNets) with high and low power nodes involves several control paradigm changes. In essence, the coordination of intra- and inter-tier interference with spectrum reuse, traffic load, and energy consumption is of paramount importance, and is handled via different kind of radio resource management (RRM) tasks. This complicated control framework of HetNets has several different dimensions which are handled separately in the literature. These dimensions have different levels of significance for the various usage scenarios presented [15]. The purpose is also to collect most relevant surveys dealing with state-of-the-art LTE-A systems and main control paradigms which are fragmented in different existing surveys.

In the following, we shortly gather some interesting viewpoints by classifying control paradigms into (see also Figure 2-1):

- **Control functionality (what):** Inter/intra-tier spectrum allocation, traffic and mobility management, packet scheduling, and link adaptation.

- **Control utility (why):** Network- and user-centric utilities; cost- and profit-based optimization.
- **Control centralization degree (where):** Centralized, semi-centralized, and distributed.
- **Control time scales (when):** Short, long, and mixed time scales.
- **Control protocol structure (how):** Control-data separation and abstracted layer coupling.

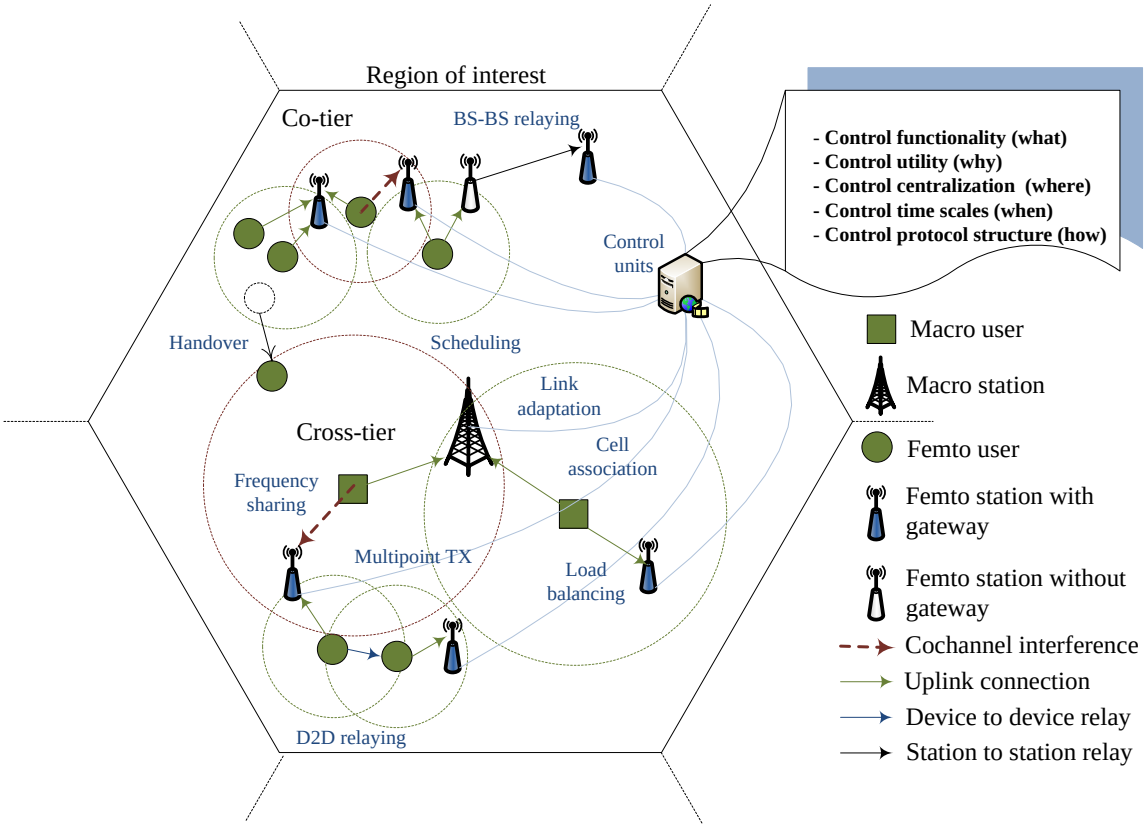


Figure 2-1 Overview of existing control paradigms in HetNets (user-centric coverage view)

2.2.2 Control functionality

The current LTE-A cellular system must perform a number RRM tasks summarized in [16]. The objective of RRM techniques is to ensure sufficient quality of service (QoS) for individual users while maximizing the resource utilization of the network without causing significant cross- or co-tier interference to other collocated users. Moreover, in modern multiservice networks with heterogeneous QoS user demands require some fairness and prioritization capabilities from the RRM. Some RRM functions are executed during the setup of new connections while some functions are conducted for every packet arrival.

The first significant control decision is the resource allocation between different tiers of the HetNet. The approaches vary from static to dynamic spectrum allocation with or without frequency reuse with an objective to balance between intercell interference avoidance and resource utilization level. Detailed spectrum sharing architectures are surveyed in [17]. The second control decision level is related to admission control, traffic steering, and mobility management. Extensive surveys on these RRM tasks are given in [18, 19]. The various phases include cell identification, admission control for QoS support, cell association, enforced handover for load balancing, possible multihop relaying, etc. The use of

femtocells complicates all these steps. The third level of RRM involves packet scheduling. An advanced scheduling decision typically takes into account the QoS requirements and current channel state as surveyed in [20]. Finally, the fourth level of RRM deals with link adaptation. Link adaptation is responsible of adaptive multi-antenna, modulation, and transmission power control whose details in the LTE system can be found from [21]. Cooperation is an important feature of advanced RRM and can take, e.g., the forms of coordinated multipoint or multi-hop relay transmission.

2.2.3 Control utility

A number of control objectives for RRM in cellular HetNets can be found depending on the target scenario. Most RRM schemes can be classified as a user- or network-centric method [22], where the former tries to maximize the incentives of subscribers and the latter one optimizes the interests of the network operator. Minimization of individuals' QoS outage probability is an example of user-centric RRM objective while maximization of sum rates of a network or resulting revenue is a natural choice for the C3. Since the two approaches are based on different optimization objectives, they result in different resource allocations. While both approaches are widely used, a recent trend is to address a user-centric approach more aggressively. Also users may have different utilities, namely best-effort users map utilities based on throughput while delay-sensitive applications emphasise delay [23, 24]. Moving from the thinking that users are just dots in the network architecture is changing to a paradigm where users' needs are understood in a better way. The challenge of future service providers is to profile heterogeneous users more accurately not only by the physical demand attributes (e.g., speed, spatial distribution, device capabilities), but also by the physiological attributes (e.g., sensory response time, language difference) and psychological attributes (e.g. content preference, service consumption habits, quality tolerance) [25, 26].

Furthermore, massive machine-to-machine communication will bring some specific features in the selection process of the controlling objectives [27]. The overall efficiency of the system is not only defined via the throughput or delay parameter but also via the QoS-aware balance of the resource provision (fairness) and quality of resource allocation (cost vs. profit) [28]. Minimization of power consumption of network nodes is important as shown in [29]. Therefore, a typical cost-based optimization function minimizes the power consumption demand or interference while the profit-based optimization function maximizes a profit such as data rate. A number of multiobjective optimization functions are also possible and receiving more attention in 5G systems [30]. Typically, the optimization functions are too complicated to be directly solved online. As a result, more practical approximate frameworks, such as game theory, graph theory, and iterative machine learning approaches, can be used to simplify the original optimization problems [16].

2.2.4 Control centralization degree

One significant problem affecting control efficiency is the selection of centralization degree for the RRM. The RRM can be divided into centralized, decentralized (also distributed), and hybrid (also semi-centralized, partially decentralized) schemes [16, 31]. Centralized schemes collect all necessary information to make all relevant RRM decisions while fully decentralized schemes use only local information at the base stations. Typically, centralized schemes allow optimal decision making but in practice only in small-sized networks whereas decentralized schemes scale better to larger network sizes, due to use of only local signalling, but lead to suboptimal decisions with a need for iterative techniques. In hybrid schemes, many features of centralized schemes are used but in this case the role of central unit is secondary or auxiliary [32]. For instance, coordination can be done at two levels: central node level controls a bulk of base station resources and base station level controls the user level resources.

In addition to control decisions, the centralization may concern also data processing as in cloud or centralized radio access networks where simple distributed radio remote heads are used and most of the data processing is centralized [33]. The centralization degree of the processing can be flexible based on current network needs and backhaul or fronthaul network limitations. The decentralized approach is further divided into coordinated distributed and autonomous distributed schemes in [31]. In coordinated distributed approach, base stations perform RRM operations locally but are allowed to exchange

channel information to enhance intercell interference coordination. In autonomous distributed approach, the local RRM decisions are based on only the information provided by the local users. Finding the optimal centralization degree is not a trivial task and many variants have been implemented in the past [34] while GSM and UMTS systems use a centralized RRM approach, the LTE system employs a more distributed RRM approach to reduce the access latencies at the cost of reducing coordination efficiency.

2.2.5 Control time scales

RRM decision execution is related to implementation complexity in terms of control overhead and needed information exchange between different cells and the decision making entities [16, 35]. Typically, control is performed at the frame level in the physical or medium access layer whereas controlling operates over longer time scales in the upper layers. Therefore, the proper separation of time scales and dimensioning of the subproblems is important. The selection of the time scale has an effect also on the necessary processing time to execute the RRM decisions and sets a constraint on the allowable computation time. While the different time-scale problems can be studied separately, mixed multi-time scale problems aim to combine different timescales into a common optimization framework [36]. When selecting the time scale for an RRM task, it is important to include the limited backhaul capacity into the design process, especially with very high network densification [37, 38, 39].

2.2.6 Control protocol structure

In addition to some general frameworks such as self-organizing networks [40] and cognitive radio networks [37], two important trends in the general control protocol structure of the HetNet concept are the control-data separation and cross-layer design using layer abstraction. In the control-data separation, the objective is to separate low rate control signals requiring full coverage from macro cells and more local high rate user data provided by small cells to improve the resource usage and interference resilience [41, 42]. This means that idle users are connected with the control macro base stations only and the link to data base station can be switched off. The control base station can help in selecting the best data base station. Closely related concept terms with some variations on how control and data are separated and how the control problem is abstracted include software-defined networks [42], phantom cells [43], information-bridlhed control [44], hyper-cellular networks [45], SoftCell [46], SoftRAN [47], MobileFlow [48], and SoftMobile [34]. The actual separation of the underlying control plane and data plane is a nontrivial task. It is essential to be able to use well defined interfaces from lower layers in a flexible way using easily programmable control as well as to support virtual network functions to improve resource utilization [34].

Cross-layer design paradigm aims to utilize information on channel variations, traffic requirements, and service demands in a common framework [28, 35, 49, 50]. This is accomplished by carefully coupling the information hidden inside different layers. A central part of this cross-layer design is the ability to abstract layer-specific parameters so that they are representable to other layers with a minimum number of parameters but without losing any significant information. Similarly, the abstractions that transform only relevant information to the C3 are important in the programmable control approach having separated control and data planes [51]. One key abstraction mapping is between signal to interference and noise ratio and packet error rate or throughput which is not an easy task to be done without semi-analytical measurements on a particular network and protocol setup.

2.2.7 Summary of control challenges

Based on previous subsections, we summarize some main control challenges in HetNets, relevant to the design of network graphs, as follows:

- RRM in the presence of unpredicted cross- and co-tier interference.
- Scalability and complexity of RRM with multiobjective network optimization and small cells.
- Optimal centralization degree with small cells.
- The effect of backhaul architectures with different time scales and ultra-densified HetNets.
- Efficient designs for control-data separation, layer abstraction, virtualization, and programmable control.

2.3 Abstracting wireless systems

2.3.1 Background

Abstraction can treat both the control actions and data structures. Examples of different abstraction levels and procedures are illustrated in Figure 2-2. The interpretation and details of the necessary hierarchical levels or abstraction steps depend on the application, including design of integrated circuits [52], software systems [53], system-level simulators [54], resource abstractions for virtualized networks [55], programming abstractions for software-defined networks [56], or information extraction from raw data [57].

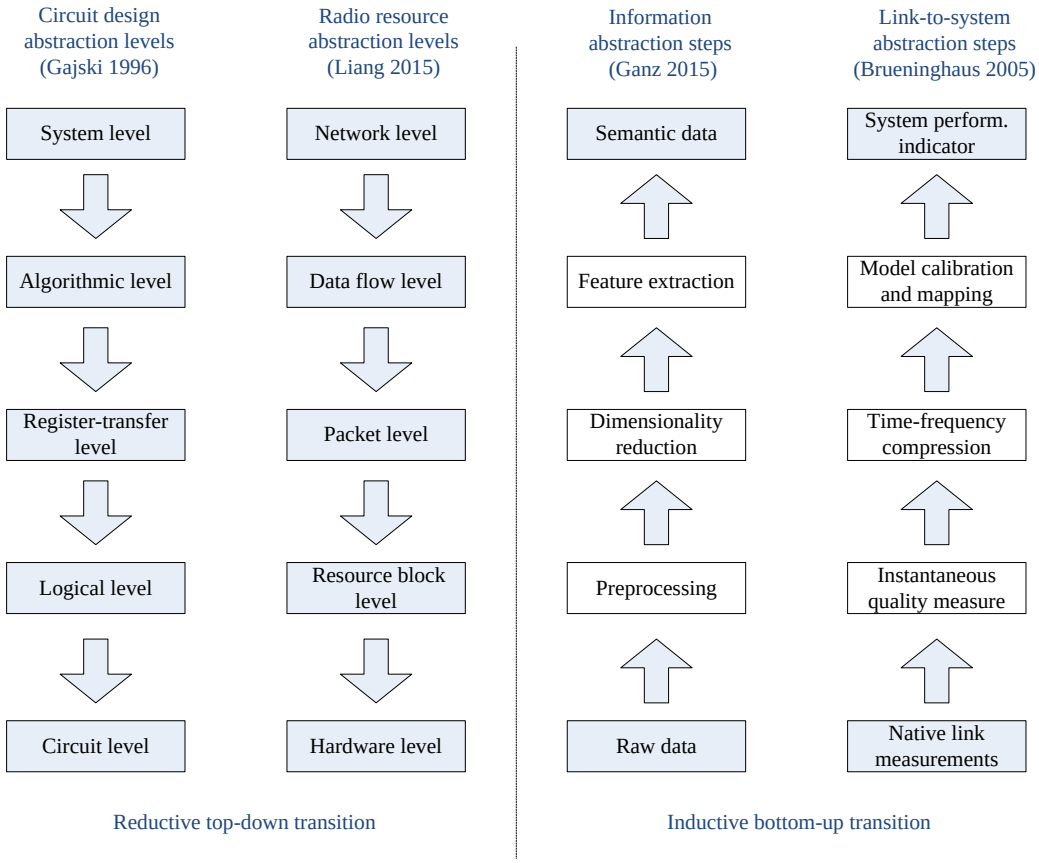


Figure 2-2 Examples of abstraction levels and procedures in different kind of applications adopted from [52, 54, 55, 57]

Abstractions of various phenomena have been extensively used for different purposes in the design of wireless networks. In a broader view, the most frequently used study cases include:

a) Promoting analytic understanding of performance of stochastic wireless systems

b) Simplification of system level simulations via link-to-system interfaces

c) Simplification of control procedures for RRM.

These study cases are highly interrelated, providing different viewpoints on abstraction procedures and their needs in wireless networks, and are treated next in more detail.

2.3.2 Promoting analytic understanding of performance of stochastic wireless systems

Regarding the case a), a vast amount of literature exists with an ultimate aim of finding closed form relationships between various system parameters and system performance. At the link level, a primary goal has been in finding unified error rate expressions in different fading channels and modulation methods [58, 59]. On the other hand, on the network level the main aim has been the analytical modelling of aggregate network interference as a function of wireless propagation effects, PHY and MAC transmission parameters, spatial density, and related performance indicators [60, 61, 62, 63]. Sophisticated abstractions of stochastic phenomena for fading and spatial distributions are needed in finding analytic expressions between system parameters (e.g. fading distribution, interference power, noise variance) are performance indicators (e.g. outage probability, data rate). Different abstractions of identifying if the packet transmission is successful are given and compared in [64]. It is concluded that the cumulative interference has a major impact and needs to be included in the abstraction model.

2.3.3 Simplification of system level simulations

Regarding the case b), the objective of abstraction is to reduce the simulation time of complicated system level evaluation by abstracting the lower layer phenomena into desired higher layer performance indicator without using complicated physical layer simulations in the actual system level simulator. A typical case is to map the measured signal to noise ratio (SINR) with resulting packet error rate (PER) for given modulation and coding method which is not typically available in an analytic form as function of desired system parameters. Different mapping functions between SINR and PER are compared in [54, 65, 66, 67, 68, 69], in the single antenna and multiantenna cases. Effect of automatic repeat request is included in [70]. Another important case is to identify the mapping between SINR and throughput, which can be achieved via modified Shannon capacity or PER-dependent throughput expression [68, 71, 72, 73]. In these works, the main problem is to identify case-sensitive post-detection SINR values that include the effects of spatio-temporal processing of the transceiver. Typically the main approach is to find an effective SINR value that is an abstract model of the transceiver and stochastic phenomena in the channel, and is then mapped to given mathematical model using model calibration techniques such as least squares curve-fitting method. Such a popular link-to-system abstraction approach used for LTE systems is described in [74].

2.3.4 Simplification of control procedures for RRM

Regarding case c), the derived abstractions are primarily used to simplify the controlling approach with the aim to reduce the control overhead and complexity of required control methods for RRM. To this end, more advanced abstraction models suitable for link adaptation control are proposed in [73, 75, 76]. In the LTE system, the control decisions for link adaptation are based on quantized channel quality indicator, which is well summarized in [21]. Abstraction models including queueing analysis for advanced scheduling control approaches have been proposed in [75, 77, 78]. Abstracted SINR models for different intercell interference coordination techniques are analysed in [79]. In [80], the packet error rate is approximated via outage probability to assist multihop routing control.

In case of control concepts where a (semi-)centralized control paradigm is used, it becomes relevant to construct a (semi-)global view of the network status. Such a view is important in many modern network control frameworks, including cognitive radio networks with spectrum awareness of primary and secondary users [81], heterogeneous networks with co- and cross-tier interference awareness [82], network virtualization [55], and software-defined networks with separated control and data planes [42, 51]. In these works, the view of the network status is often called as a map or graph that allows the C3 to perform the RRM tasks more efficiently. There are a number of ways to create and utilize such a graph. Typically, the main aim is to reduce the decision space so that the RRM decision problem becomes tractable [83].

2.3.5 Remaining challenges

Based on previous subsections, we summarize some challenges in abstracting wireless systems as follows with an emphasis on control procedures:

- Addressing optimal trade-offs between overheads and accuracy of used abstractions to enable scalable RRM decision frameworks
- Identification of possible new phenomena arisen from massive MIMO and massive machine-type access
- Smart aggregation of PHY (channel, interference, and power) and MAC (data buffer) based graphs
- Deriving probabilistic abstracted metrics to predict consequences of control decisions
- Extending the abstraction to higher levels and use it to enable a new control paradigm for complex densified wireless networks
- Identifying what kind of abstractions are missing or could be improved in current systems; in essence, what parameters that are currently handled in a distributed way in LTE base stations can be used by a C3 in a reasonable way.

2.4 D2D channel modelling

Accurate channel models are fundamental to the build-up of a wireless communication system and play a vital role in the process of algorithm design, protocol definition, performance evaluation, and network optimization. From a practical point of view, we need a model that combines all the aspects of radio channel including the path loss, shadowing and small-scale fading, rather than the theoretical one such as the Rayleigh distribution that only captures the small-scale fading.

The propagation characteristics of device-to-device (D2D) communication radio channels are distinct from that of the conventional channels in terrestrial cellular systems due to the differences on antenna height, propagation distance, moving speed, surrounding environment, and so on. That is to say, the D2D links, including human-to-human (H2H), vehicle-to-vehicle (V2V), and machine-to-machine (M2M), are required to be modelled for facilitating the system design and performance evaluation of the upcoming 5G systems.

Special requirements for D2D channel modelling
There exists tens of, if not hundreds of, models to describe the propagation characteristics of radio channels. However, most existing channel models are not as such suitable for modelling D2D links. Compared to the conventional cellular models, which capture the radio links between the base stations and the mobile terminals, D2D links have special features:

- Both ends of the link are typically at low heights.
- Both ends of the link can be moving.
- Both ends of the link are subject to shadowing by the user and surrounding obstructions.

Because of this dual mobility, both ends of the link can be at arbitrary location -usually close to each other- which has its impact on channel modelling requirements.

Basically, the most important factors to affect the channel characteristics are the carrier frequency, the propagation distance, the height of antenna, the propagation environment, and the movement of transmitter and receiver. As is well known, plenty of models such as COST, SCME and WINNER II, have been generated during the research and development of cellular communication systems in the past decades. However, these models are specifically developed for the radio channels between the terrestrial base stations and the terminals. The following factors contribute to the specific propagation characteristics of D2D channels:

(1) The antenna heights of terrestrial base station are normally ranged from 25m to 50m for urban macro-cell, 10m for urban micro-cell, and 3-6m for indoor hot spots. In D2D communications, the antenna heights of both transmitter and receiver should be around 1.5m (in comparison with the floor in case of indoor).
(2) In general, the scatters are close to the terminal while locating far away from base station. Both the transmitter and receiver in D2D channels are surrounded by the scatters, leading to different channel characteristics especially in multi-antenna scenarios.

(3) The movement of the transmitter or receiver cause Doppler spread. Although the terrestrial base stations are always static, the terminals at both ends of D2D link are highly possible to be in motion.
(4) As given in a survey of Intel research team [191], the rough estimated distance between two terminals connecting D2D link in *a dense urban area* is merely about 10 meters, which is much shorter than the typical distance observed in cellular systems.

2.4.1 3GPP activities on D2D channel modelling

In December 2012, a Study Item (SI) on LTE Device-to-Device Proximity Services was established by 3GPP. The D2D channel models were investigated in 3GPP RAN1 meetings. The scenarios, evaluation methodology, and simulation parameters, which are strongly related to the selection of D2D channel models, have also been investigated and reported.

Scenarios
The following are the proposed scenarios for investigation:

- Option 1: Urban macro (500m ISD) + 1 RRH/Indoor hot zone per cell – mandatory for general scenario
- Option 2: Urban macro (500m ISD) + 1 Dual stripe per cell
- Option 3: Urban macro (500m ISD), all UEs outdoor
- Option 4: Urban macro (500m ISD) + 3 RRH/Indoor hot zone per cell
- Option 5: Urban macro (1732m ISD), mandatory for public safety specific scenario
- Option 6: Urban micro (100m ISD)

Qualcomm is driving this very strongly and it is worthwhile reading their contributions R1-131412 and R1-131413. They already have their proprietary D2D solutions that will be adapted to merge into LTE technology more smoothly.

In summary, 3GPP has achieved the following agreements on the D2D channel modelling, which can be mainly summarized as follows:

- Symmetric angular spread distribution and dual mobility corrections
- Amend the ITU-R UMi/InH model to incorporate dual mobility
- TX (UE1) and RX (UE2) parameters separately with phase change per sub-path
- Direction of Travel (velocity vector) independent and random
- Doppler is determined by path AoA and AoD
- Uniform AoA spread of 104 degrees

2.4.1.1 WINNER channel models for D2D

During the process of EU/ FP7 ABSOLUTE project, the following conclusion was revealed after the investigation of all the existing channel models. For the D2D communication links, one can take advantage of Scenario A1 of WINNER II channel models, although it cannot capture all the features of D2D channels (the drawbacks will be described in the end of this subsection). The antenna heights of both BS and MS range from 1 to 2.5 meters, which are quite similar to the propagation environment encountered by D2D communications. The applicable frequency range is also from 2 to 6 GHz. Similarly, the scenarios B4 and C2 of WINNERII model have been used to evaluate the links, such as that between PLMU and UE.

The WINNER II project developed channel models suitable for different propagation environments, including indoor and outdoor, line-of-sight (LoS) and non-LoS (NLoS). In particular, WINNER II is based on a stochastic geometric modelling approach in which the radio signal is made of rays that are assumed to arrive in clusters. Each cluster corresponds to a path in the multipath environment with each cluster (or path) made of several rays (or sub-paths). Referring to the documentation explaining the MATLAB implementation of each WINNER II model [192], there are up to 24 paths with a fixed number of rays equal to 20. The models allow taking into account several propagation parameters (number of stations, number of links, antenna height, indoor/outdoor, room and street environments,

among others), as well as different RF parameters (antenna type MIMO or SISO, delay, power, angle-of-arrival and angle-of-departure of each ray). Typical outputs that can be collected with WINNER II software is the RMS delay spread of the received signal. Basically, B4 is particularly suitable for outdoor NLoS environments, whereas C2 can be used to obtain results for urban environment in NLoS conditions. Relevant parameters for B4 and C2 channel models derived from the WINNER II Interim Report are provided in Table 2-1.

Note that the wall of the building has a huge impact on the D2D channel modelling, so the discussion of D2D modelling is generally divided into three scenarios: Indoor-to-Indoor, Indoor-to-Outdoor (as well as Outdoor-to-Indoor), and Outdoor-to-Outdoor.

Table 2-1 A1, B4 and C2 WINNER II Channel Models [193]

Scenario	Definition	LoS/NLoS	Mobility [km/h]	Frequency [GHz]	Concept group	Note
A1	Indoor small office /residential	LoS/ NLoS	**0-5**	**2-6**	**LA**	
B4	Outdoor to indoor/ Outdoor typical urban	NLoS	0-5	2-6	MA	B1 or C2 to the wall/window
		$PL = PL_b + PL_{tw} + PL_{in}$ where $PL_b = PL_{B1}(d_{out} + d_{in})$ $PL_{tw} = 14 + 15(1 - \cos(\theta))^2$, $PL_{in} = 0.5 d_{in}$ $PL_{NLoS} = PL_{LoS}(d_1[m]) + 20 -$ $12.5 n_j + 10 n_j \log_{10}(d_2[m])$, with $n_j = \max(2.8 - 0.0024 d_1[m], 1.84)$ Shadow fading std: $7dB$ $3\,m \le d_{out} + d_{in} \le 1000\,m$ $h_{BS} = 10\,m$ $h_{MS} = 3n_{FL} + 1.5\,m$				
C2	Typical urban macro-cell	NLoS	0-120	2-6	MA WA	
		$PL = [44.9 - 6.55\log_{10}(h_{BS}[m])\log_{10}(d[m])]$ $+ 34.46 + 5.83\log_{10}(h_{BS}[m])$ $+ 20\log_{10}(f[GHz]/5.0)$ Shadow fading std: $8dB$ $50\,m \le d \le 5\,km$ $h_{BS} = 25\,m$ $h_{MS} = 1.5\,m$				

PL_b is the basic loss between the BS and MS along the route dout - din without wall effects

PL_{tw} is the loss through the outer wall

PL_{in} is the loss inside the building

PL_{B1} is B1 path-loss

d_{out} is the distance between the outside terminal and closest point of the wall to the inside terminal d_{in} is the distance from wall to the inside terminal

θ is the angle between the outdoor path and the normal of the wall.

n_{FL} is the number of the floor. (Ground floor is the number 1.)

h_{BS} is BS antenna height

h_{MS} is MS antenna height

d_1 and d_2 is defined for NLoS case in Figure 2-3.

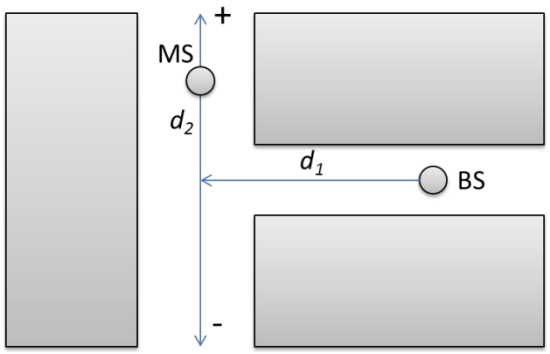


Figure 2-3 Geometry for d_1 and d_2 path-loss model for B4 NLoS

Let **H**(*t*) denote the channel matrix which is obtained after running MATLAB simulations and let $s(t) = \sqrt{P_t}\, p(t)$ denote the transmitted signal, let P_t denote the transmitted power and $p(t)$ be a suitable way of representing and OFDM signal with normalized unit power. The received signal $r(t)$ can thus be written as

$$r(t) = s(t) \otimes \mathbf{H}(t)$$

$$P_R = P_T \left|\mathbf{H}(t)\right|^2 = P_T \sum_{l=1}^{L} \left|h_l(t)\right|^2 ,$$

where $8 \le L \le 24$ is the total number of multipath components, P_R is the received power of the signal, h_l denotes the channel impulse response of the lth path in a cluster, and $\otimes$ is the convolution operation. Relying on [194] for numerical values, the achievable capacity evaluated in terms of the number of bits per transmission dimension (b) is computed as

$$b = \frac{1}{2}\log_2\left(1 + \frac{P_R}{N_0 W}\right) \qquad \text{[bits/dimension]},$$

where W denotes the signal bandwidth of LTE system, $N_0/2$ is the two-sided power spectral density of the Additive White Gaussian Noise (AWGN), and P_R is the received power. To obtain numerical results, the preferred system parameters include W = 10 MHz, P_R = 23 dBm and N_0 = -174 dBm/Hz.

Numerical results for WINNER II channel model B4 are shown in Figure 2-4. The figure shows the received power and the achievable bits/dimension. The antenna height of the PLMU is assumed to be 10 m, while that of UE is 1.5 m. The speed of a user is assumed to be 1.1 m/s and the number of traversed floors is one. It shows that the simulated distance separating transmitter and receiver is up to 100 m and the achievable bits/dimension decreases as the distance increase. For short distances, the communication links can use even a 64 QAM (b = 6), whereas it reduces to a QPSK (b = 2) approaching 100 m. Interestingly such a decrease is in line with the information conveyed by the channel quality indicator (CQI), which can be used to select a higher modulation and coding scheme (MCS) in correspondence to a higher CQI value that could occur more likely for shorter distance separations between the transmitter and the receiver.

Numerical results for the WINNER II channel model C2 are shown in Figure 2-5. The figure shows the received power and the achievable bits/dimension. Since channel model C2 is suitable for wider areas,

higher speed is considered here as well. The antenna height of the PLMU is assumed equal to 25 m, while that of the UE is 1.5 m. The speed of the user is assumed here to equal 14 m/s and the simulated distance separation between transmitter and receiver is up to 800 m. Conclusions similar to those presented in Figure 2-5 can be derived also for this case. In general, as anticipated in [195], increasing the antenna height with respect to channel model B4 has a positive effect since it allows keeping approximately the same number of bits/dimension for higher distance values.

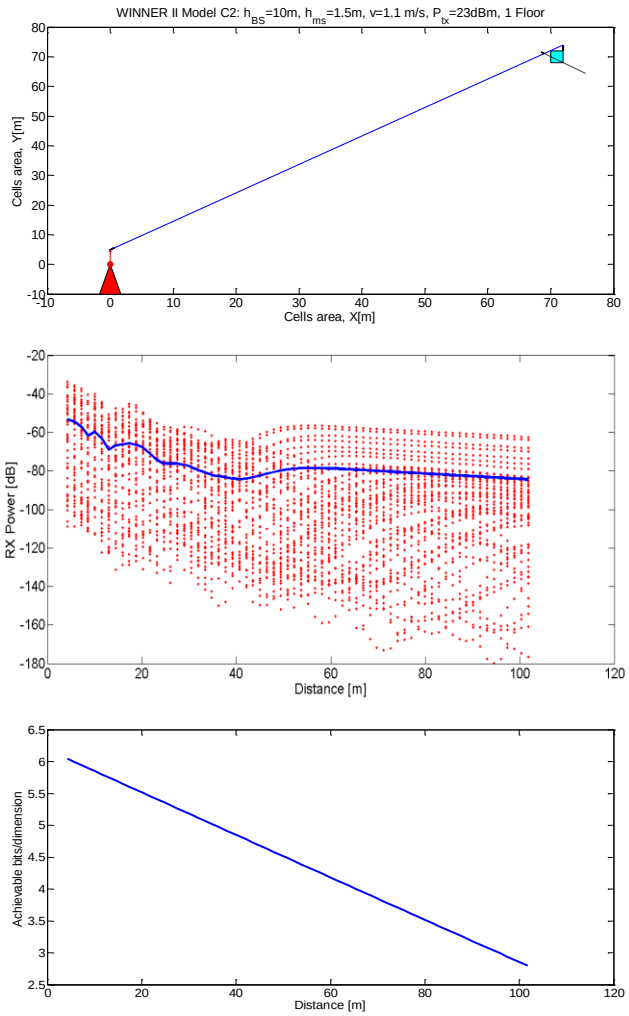


Figure 2-4 Results on WINNER II channel model B4

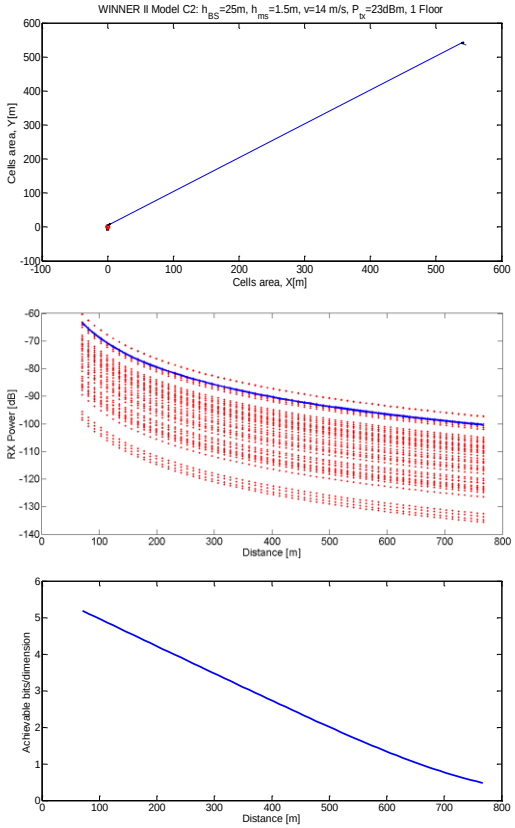


Figure 2-5 Results on WINNER II channel model C2

2.4.1.2 Comments and drawbacks about WINNER II used for D2D modelling

In WINNER II channel models, the basic principle is that for every link the large scale parameters (like angular spreads) are taken from a map. In that way, the correlation properties of those parameters are matched with those observed in measurements. However, the small scale parameters (like angles of arrival and departure) are randomly drawn from a distribution, independently for each link. This means that even close-by links have independent values for e.g. AoA and AoD, which is of course not the case in reality. This spatial inconsistency was not problematic with the quasi-stationary modelling of WINNER, but it has an impact on performance with e.g. multi-user MIMO case. The spatial inconsistency also means that the WINNER approach does not handle time evolution very well. New sets of parameters are randomly drawn at each location of a mobile, and there is no smooth transition between two locations. This means that dynamic simulations are problematic. Interpolation between two locations is of course possible. The interpolation can be done by drawing random small scale parameters, like cluster delays, powers, directions, etc., to two UE locations and linearly interpolating parameter values in between locations. A problem may result because interpolated values are always between the original values, and thus with interpolation all distributions become narrower.

WINNER models also do not specify transitions between different propagation environments (urban, rural, outdoor, indoor, etc.) or between LoS and NLoS, which also creates spatial inconsistency and unrealistic transients.

2.4.2 D2D channel measurements in EU/FP7 ABSOLUTE project

Mainly due to the lower transmitter height, the D2D channel varies greatly from the conventional terrestrial channel used to model cellular communications where the transmitters are mounted on rooftops or on towers. There have been various propagation measurements conducted to model D2D channels in the literature. They have measured the channel for different frequencies in different channel conditions and have suggested different models. Some of the related works are mentioned below. In [196], [197], and [198], the author has characterized the D2D channel for different environments and different frequencies. In [199], the authors have studied the path-loss behaviour of D2D channel in an

urban environment for millimetre-wave frequencies. In [198], the author has studied the effect of shadowing caused by human body in cellular D2D communications. However, none of any previous works to model the D2D channel in a rural area. It is important to characterize the D2D channel in a rural area for ABSOLUTE with respect to the forest fire scenario, earthquake, and flood. Hence, the ABSOLUTE project consortium has performed experiments in a rural area to characterize the D2D channel in terms of path loss exponents and the standard deviation of shadowing. The experiments have been carried out for two ISM (license-exempt) bands 922 MHz and 2.466 GHz. More information and numerical results about these measurements can be found in ABSOLUTE FP7 project deliverables and from [200][201][202].

2.4.3 Conclusion of D2D channel modelling review

As a conclusion, D2D communications have been extensively recognized as one of key technologies to meet the requirements of the upcoming 5G systems. On the other hand, accurate channel models play a vital role in the research and development of D2D communications.

However, the legacy channel models generated during the process of designing the terrestrial cellular systems, which focused on the links between base stations and mobile terminals, are hard to be applied directly to the D2D communications. That is because the propagation characteristics are strongly related to the antenna heights, the propagation distance, the mobility, and the surrounding environments of both the transmitter and receiver, among others. These factors are distinct between the conventional cellular communications and the D2D links. That is the motivation behind investigating D2D channel modelling.

In this literature review, we provided a basic vision of D2D channel models. The activities of 3GPP in the Study Item of D2D proximity services can be an important reference, where the application scenarios, the evaluation methodology and simulation parameters of D2D, have been discussed. A lot of potential D2D channel models have been investigated by the 3GPP contributors. The WINNER II channel models were also shown since some cases have similar environments as D2D, which can be used as a reference model, but some drawbacks need to be solved. Besides, the EU FP7 ABSOLUTE project have carried out some measurements on D2D channels in the rural areas for the Public Safety scenario, and the results can be used as a concrete reference for the D2D channel modelling.

2.5 Relevant standardization activities for supporting a central coordinator in LTE standards

The existing LTE networks operate based on a distributed approach, however in LTE Release 12 and 13 the new CoMP procedures were designed to allow the operation of a central coordinator, without explicitly mentioning this.

In Release 12, part of the Small Cell WI, was given support for a Central Coordinator; the text in 3GPP TS 36.300 [85], Section 16.1.9, “Inter-eNB CoMP” says:

“The task of inter-eNB CoMP is to coordinate multiple eNBs in order that the coverage of high data rates and the cell-edge throughput are improved, and also the system throughput is increased. The coordination of multiple eNBs is achieved by signalling between eNBs of hypothetical resource allocation information, CoMP hypotheses, associated with benefit metrics. Each of the signalled CoMP hypotheses is concerned with a cell belonging to either the receiving eNB, the sending eNB or their neighbour.”

The word *“neighbour”* is hiding a “C3”. Section 16.2, “RRM Architecture” starts with 16.2.1, “Centralized handling of certain RRM functions”, which includes a single word: *“void”*. The hidden support of a C3 continues through the reports related to eNB CoMP, as shown in the following text from Section 16.1.9:

“RSRP measurement reports and CSI reports may be exploited for inter-eNB CoMP. For example, the RSRP measurement reports and CSI reports can be used to determine and/or validate CoMP hypotheses

and benefit metrics. The enhanced RNTP may be used in inter-eNB CoMP to exchange information between eNBs concerning the adopted power allocation. RSRP measurement reports and CSI reports may be exploited for inter-eNB CoMP. For example, the RSRP measurement reports and CSI reports can be used to determine and/or validate CoMP hypotheses and benefit metrics."

Due to the high time-frequency resolution of the CoMP hypothesis and eRNTP reports and also to the non-incremental coding mode of the defined messages, the generated traffic is too high to be used in distributed mode. For a 20MHz channel, with 110 PRBs, and a system frame (40 subframes) the eRNTP (see 36.423 [86]) the message length can be longer than 4.4 kbit for 1 bit per PRB and 8.8 kbit for 2 bits per PRB. If the message carrying the RNTP information is repeated every 5 ms will result a data rate of 0.9/1.8 Mb/s for exchanging this information between two eNBs. For a modest number of 4 neighbour eNBs will result an up-link traffic higher than 3.6/7.2 Mb/s, which is more than what a typical up-link very high rate digital subscriber line (VDSL) will allow. Opposed to this, the centralized approach will reduce the traffic to the initial 0.9/1.8 Mb/s, which is still too high given that remains not much room for user data.

2.6 Open problems

We identify a number of open issues that are to be further investigated for the remainder of the project in collaboration with mainly WP2 but also WP5. In general, the research towards effective abstractions starts with a fundamental understanding of the coordinated RAN control procedures with consideration of the split of local and centralized control functions and corresponding abstractions. The main questions are focused on how and where the network graphs are created in relation to C3 functions:

Network graph abstractions and to which degree they can be used over different link layer technologies, which requires:

- Identification of common parameters and invariants for the purpose of exposing resources and network state information over different RATs.
- Identification of network graphs that are common for managing different RATs.
- Identification of abstractions that support programmability.

In general, finding abstractions that are meaningful for both LTE and WiFi is difficult as the possibility of creating the network graph information elements (edges and vertices) depends on the standards used and the properties of the RAT. First steps towards addressing these issues are reported in Section 4, which encompasses identification of available metrics for both LTE and WiFi technologies. Section 4 also includes suggestions on how existing and new metrics can be used for representing the network state at a higher level of abstraction. In Section 5, specific methods for creating network graphs and metrics are described in more detail.

Aggregation, storage and access of data and network graphs, which require further understanding regarding:

- Infrastructure capabilities and constraints with respect to monitoring, storage and access (i.e. available data sources);
- Application programming interfaces (APIs) for access and information exchange between network entities and controllers;
- Where network graphs (fully or partially) are stored (e.g. controllers, dedicated storage, or in-network) and how often they need to be updated and accessed.

The network graph is essentially a data structure that is aimed to provide a holistic view of a specific aspect of the network state. The metrics and methods proposed in Section 4 and Section 5 are based on highly distributed data sources combined with computationally light-weight local processing (e.g. aggregation and parameter estimation), for the purpose of providing input to network graph data structures residing in dedicated storage associated with the COHERENT controllers entities. The decentralized and hierarchical COHERENT architecture worked out in collaboration in WP2 allows for distributed and centralized storage and access of the network graphs and its elements such that

management functions can operate in close to real-time and at longer time scales with low messaging overhead.

C3 complexity with respect to the network graph abstractions and usage, which requires further discussions about:

- The level of controller-split functionality and degree of decentralization;
- The information and access requirements for different control functions operating at various levels of the protocol stack (e.g. with respect to control-loop timing);
- How network graphs are managed by a controller and between controllers in a consistent, scalable and timely manner.
- How higher-level management applications can access network graph information from one or several controllers.

The controller complexity and degree of centralization influence to which degree a network graph can be distributed, maintained and applied in various management actions carried out at long and short time scales. The COHERENT architecture and its controller components have been outlined in greater detail in D2.2 [140] which has allowed for identifying where and how network graphs are created, stored and accessed. In summary, aggregation and processing of metrics takes place as close to the data sources and the infrastructure as possible. Aggregation, data processing and metrics maintained locally in nodes having sufficient computational capabilities serve the purposes of low communication overhead in the network and support self-organizing and distributed management actions that need to be executed in real time. Network graphs are created per region and are managed directly by components being part of the regional controller (e.g. an associated network entity for offloading). At the centralized coordination level network graphs over several regions can be merged and used for management operations over longer time scales. The exact mapping depends of a network graph is specific to the management application. Section 5 describes a number of approaches related to monitoring and network graph abstractions along with suggestions how the methods can be mapped to the COHERENT architecture. Interfaces for exchanging network graph information between controllers remain to be worked out in further detail in WP2 in collaboration with WP3. Information consistency between COHERENT controllers is out of scope in WP3, but is partially investigated in WP5 for the purpose of maintaining connectivity under service disruption or performance degradations.

3. COHERENT system model and abstraction framework

In this section, the COHERENT system model is outlined and the target abstraction framework is introduced.

3.1 COHERENT network architecture

In 5G, RAN coordination and programmability are central concepts that are aimed to improve service quality, resource usage, and management efficiency, addressing the limitations of the current LTE and WLAN systems operating under highly distributed control [206]. Centralised solutions in SDN have potential to achieve the global optimal, but coming with high cost on scalability and latency. To address the scalability and latency issues, two control mechanisms in the COHERENT Architecture are designed for achieving programmable 5G RAN, namely Central Controller and Coordinator (C3) as well as Real-Time Controller (RTC) as shown in Figure 3-1. The real time here means that the time scale for the status update and control decision between RTC and Radio Transceivers (RT) is in the order of milliseconds.

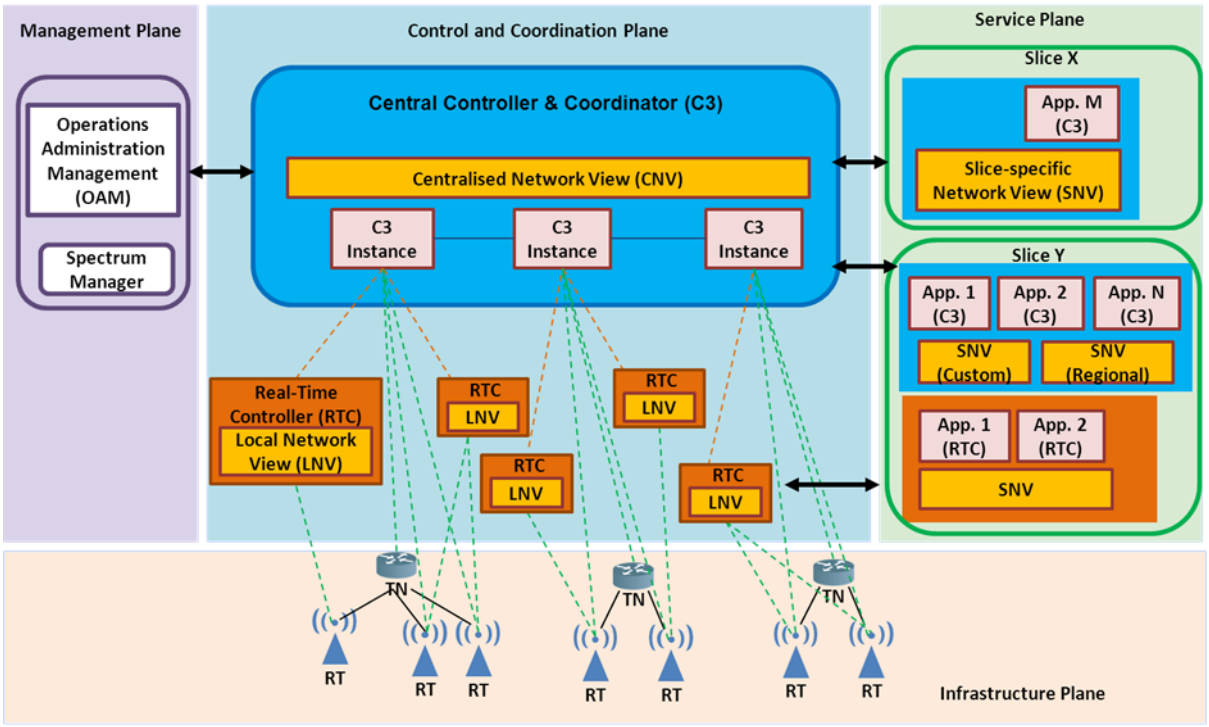


Figure 3-1 COHERENT Architecture

The C3 is a logically centralised entity[1], which provides network-wide control for the networks. The main function is to orchestrate the behaviours of network entities, mainly base stations or access points, in the RAN so that network behaviours in the RAN are harmonised. By receiving status reports from low layer entities, C3 maintains a centralised network view of the governed entities, e.g. transport nodes (TNs) and Radio Transceivers (RTs) in the RAN. Based on the centralised network view, the SDN principles are applied in the design of the C3. For overcoming scalability issues in a large and dense RAN deployment, or for performance/reliability reasons, the logically centralised C3 can be distributed with multiple C3 instances in regions so that the regional network graphs are shared across several physical control instances. The distribution abstraction **shields higher layer from state dissemination and collection,** making the distributed control problem a logically centralised one.

[1]Note that defining C3 as a logically centralised entity neither prescribes nor precludes implementation details, e.g. the federation or hierarchical connection of multiple control instances.

The biggest challenge in creating such a software defined radio access network is the inherent delay between any centralised C3 and the individual radio elements. For overcoming the delay limitation, latency-sensitive control functionalities are offloaded from the C3 to RTCs, ensuring that the radio elements are given the opportunity to adjust to rapidly varying wireless networks. By offloading latency-sensitive control functionalities between the C3 and the RTC, the C3 makes decisions that affect the logically centralised network states, while the RTC handles control decisions for latency-sensitive network functionalities in RTs with its local network view (LNV).

Inside the control entities (C3 and RTC), there are southbound and northbound interfaces. The southbound interface connects to different radio access technologies (RAT). Control applications are built upon the northbound interface, to perform the high level spectrum management, mobility management, traffic steering and network slicing functions. A network slice is defined as a collection of specific control applications and RAT configurations, which are aggregated together for particular use cases or business applications. A network slice can span all domains of the network: software programs running on cloud nodes, specific configurations of the transport network, a dedicated radio access configuration, as well as settings of the 5G devices. Some application modules in network slices may be latency-sensitive. For such a slice, these modules are located in the RTC[2]. The C3 and/or RTCs provide the required network view, namely slice-specific network view (SNV), for the network service slices so that network service slices could express desired network behaviours (by programming) without being responsible for implementing that behaviour (with hardware) themselves.

3.2 COHERENT network graph and abstraction framework

The graph comprises nodes and edges, where a node represents a single or a set of network elements, whereas an edge between a set of nodes may be used to represent aspects such as connectivity state, resource constraints, and interference coupling, depending on the nature of measurements, reporting and abstractions graph. An example of a network graph is provided in Figure 3-2.

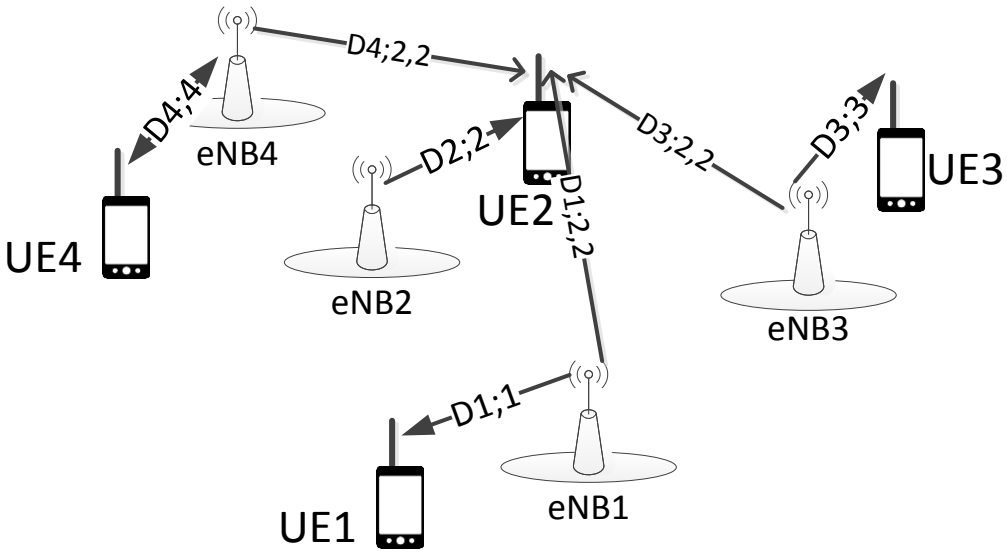


Figure 3-2 A downlink network graph

The network graph in Figure 3-2 shows the base stations (nodes or vertices) creating interference for example to UE2. The serving and interfering links are named "edges" of the network graph. The edges

[2] The examples of latency-sensitive network applications are flexible function splitting in cloud-RAN, MAC scheduling (regular, CoMP, eICIC, transmission mode selection, etc.), X2 HO decision, MAC/PHY (more generally cell) reconfiguration and most of MEC applications (localisation, augmented reality, low latency IP service, etc).

are noted with D (for downlink) followed by the index of the transmitting node, the index of the receiving UE and the index of the node serving the UE. In case of the edge link connecting an UE to a serving eNB, D is followed by the index of the node serving the UE and the index of the UE (receiver).

3.2.1 Nodes

As HMNs consist of different types of network elements, a node can be defined as a particular type of network element such as a base station, or a combination of network elements such as a base station and associated UEs comprising a small cell, or a D2D pair, etc. This broad definition will help us incorporate constraints related to intracell, intercell, and duplexing resource allocation. Following is a partial classification of the nodes, based on the list of network elements that are supported by the HMN.

The nodes can be divided into different categories. For LTE, the most prominent such categories are:

- UE (for PMR there is a distinction between hand-held and vehicular)
- eNodeB: Macro, Micro, Pico, femto or Home-eNodeB
- Relays
- Machines

Moreover, their characteristics can vary with respect to the following parameters:

- Radio access technology (RAT)
- Transmit powers
- Bandwidth (carriers and resources)
- Spectrum mask
- Subscriber group or operator
- Noise and interference floor
- Type of receiver

3.2.2 Edges

The edges between the nodes depend on the model of physical environment, details of measurements and the level of abstraction. In principle, it can be assumed that edges exist between all the nodes in the network. The edges at PHY-layer may represent the propagation and interference conditions that exist among the nodes, whereas the edges in MAC-layer may reflect the resource constraints between the nodes.

Edges can be qualified at different layers (PHY, MAC etc). There may also be dependence on the technology. Therefore, we label each edge with the time, frequency and spatial resource to which it is associated, as well as a time deadline for indicating when the information will be outdated.

3.2.3 Abstracted network graph

The network graph can be abstracted without explicit reference to the network elements represented by the nodes. In a purely cellular environment, the eNodeBs would be the only nodes in such an abstraction. The edges would then be labelled with characteristics describing averaged resource constraints, when a given resource is exploited in both cells. The abstraction is then nothing else than a clustered version of the network graph, where the distinction between UEs associated to the same cell is ignored. This allows us to distinguish between inter-cell decisions, which are handled by C3, and intra-cell decisions that can be handled by a local controller in the eNodeB, see COHERENT architecture in Section 3.1. The size of the graph is then substantially reduced, by dropping edges where the interference measures are below some predefined threshold. For example, if the resource considered is spectrum, then a threshold of acceptable background signal is defined. The edge between two nodes is considered, only if the interference caused when they use the same spectrum exceeds the defined threshold. An example is given in Figure 3-3, where the nodes are given by the cells in a homogeneous cellular system, and there is an edge between pairs of cells if the annealed signal strength exceeds the threshold. Observe that, while the graph has a strong geometric structure (nearby cells tend to cause more interference than distant ones), it is not planar, nor is it defined in geometric terms. The key idea

behind this abstraction is that a valid spectrum assignment is a proper colouring of the abstract graph, and an example of such a colouring is given in Figure 3-3. In other words, two cells are allowed to use the same spectral resource if and only if they are not connected by an edge in the graph.

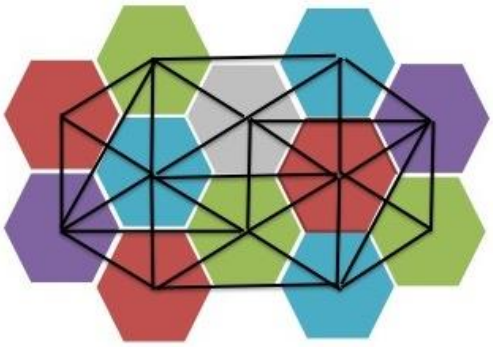


Figure 3-3 Abstracted cellular network graph in a homogeneous network, with resource assignments represented by colours

In a heterogeneous setting however, a more refined abstraction subdivides cells into sets of users, according to what other cells they experience substantial interaction with. This can be done without reference to the network elements themselves, by predefining nodes corresponding to all pairs (C,S) of a macrocell C and a small cell S, where users assigned to S can experience substantial interference from C. Once the graph is defined, network elements such as UEs can be assigned to nodes, according to measures and estimates of network parameters that are compared to predefined threshold values. An example of such an abstracted graph is given in Figure 3-4. Once again, assignment of network resources can be done via graph colouring. This separates the problem of resource assignment into two stages. An abstraction stage, in which network elements are clustered and continuous parameters are discretised, is followed by a colouring stage where a purely combinatorial problem is solved heuristically.

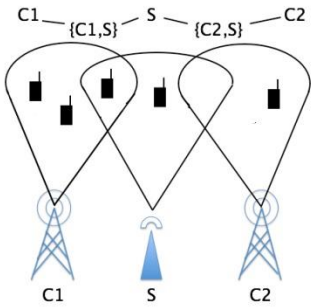


Figure 3-4 Network abstraction in a heterogeneous network, with two macro-cells and a small cell used for offloading and range extension

3.2.4 Overview of abstraction procedure

The nodes and edges of the network graph are the two basic components of the network graph. To build up the nodes and the edges of the graphs, the network information is abstracted, managed and maintained by the COHERENT controller (e.g., C3 or RTC) through the interface between the controller and RATs. For instance, when a new user joins the underlying LTE network, the configuration of this user (user identity, user capability, etc.) and the associated status (attached cell identity, tracking area code etc.) of this user are gathered and can be used to build up a node in the network graph. When such a user sends the measurement reports that contain dynamic information on the link performance, such as measured signal power and interference power, this information is used to build up the edge between this node and other nodes. All these correlated users' configurations and statuses are maintained in order to build up the network graphs. In contrast, when a network node

abruptly leaves the network, the associated configuration and status are removed accordingly and also the associated nodes and edges in the network graphs. These network graphs are used to provide abstracted information to some higher-layer applications for different purposes, e.g., mobility management, relay selection, etc. Figure 3-5 depicts how the network can be abstracted in order to build up the network graph inside the C3 instance and shows the interactions between the network management and the application. The nodes and the edges of the network graph are built based on the abstracted network information status or configuration on top of multiple underlying RATs.

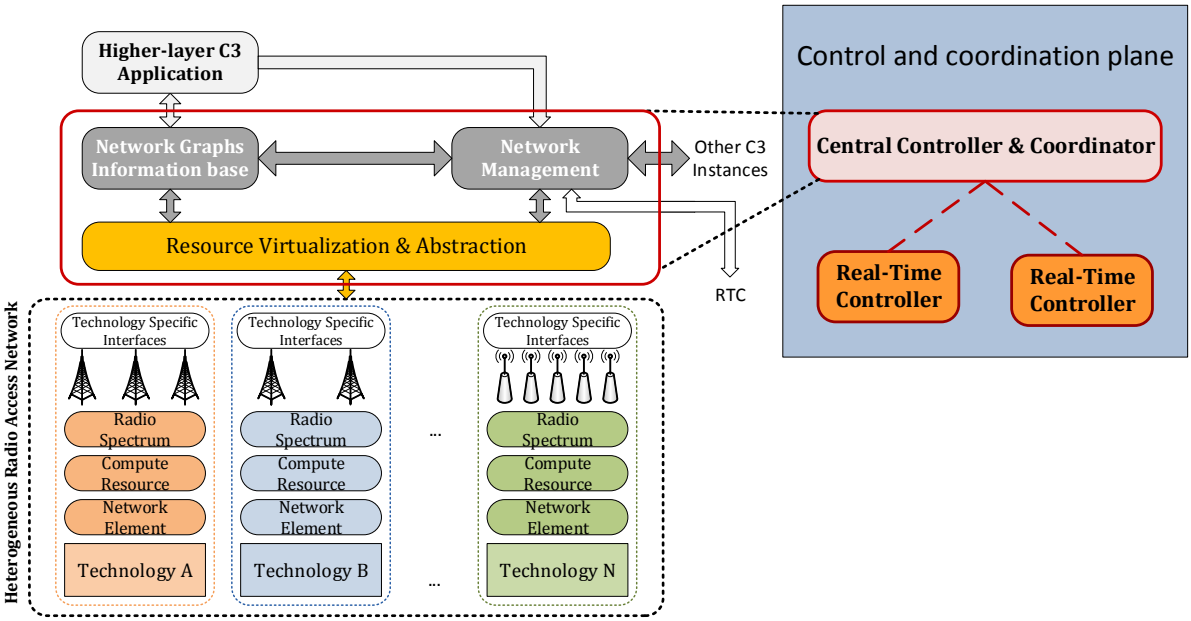


Figure 3-5 Abstraction flow for the network graph in C3 instance

As for the network graph inside the RTC, it follows the same abstraction approach as the one in C3 shown in Figure 3-6. Since the RTC only has the local view of the network, the network graph of RTC is built based only the underlying RAT and the concerned real-time network status. Moreover, the management entity of RTC interacts with the one of C3. The signalling over this interface might be originated by management entity of C3 (e.g., the UE activation/deactivation that is managed by higher-layer by C3) or the entity of RTC (e.g., CoMP, ICIC or X2 handover that are managed in real-time basis). The interaction between the management entity of the RTC and the C3 enables both network graphs of RTC and C3 to have same up-to-update common network information (e.g., status and configuration) for their individual purposes.

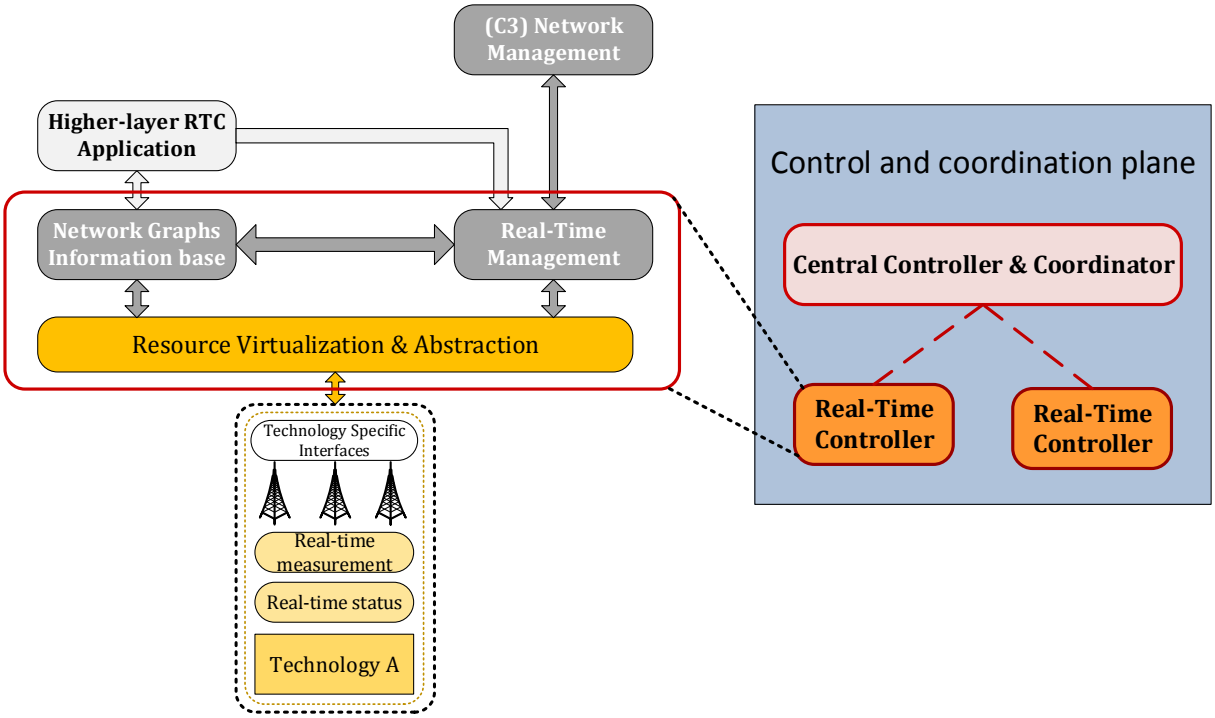


Figure 3-6 Abstraction flow for the network graph in C3 instance

4. Measurements and metrics to generate COHERENT network graph

This section outlines some important alternative metrics and measurements that could be used to generate the COHERENT network graph. The presented ideas are then tuned for specific control applications and evaluated in Section 5.

4.1 Available measurements and metrics

4.1.1 PHY layer – LTE

From a PHY layer perspective, an edge is a connection between two nodes indicating the attenuation of power transmitted at one node and received at the other. The key parameters regarding the large-scale propagation effects include path loss and the shadow fading. These parameters are static or vary over a much longer timescale, and must be considered in conjunction with instantaneous and small scale fading effects, which are dynamic in nature. The information is collected over all available resources, and used to determine the edge weights, in the form of a vector or a tensor (a vector of matrices) for scalar and MIMO channels respectively. Moreover, the gradient and the statistics of the edge weights may be considered for improved robustness.

In contrast, fast fading occurs locally in both time and space. This must be measured at and reported from each node, contributing additional information about the graph to dynamically add to the static graph that is collected globally.

LTE bandwidth can be configured by RRC according to these values, expressed in number of Resource Blocks (RB): 6, 15, 25, 50, 75, 100, which correspond respectively to the following channel bandwidth (see TS 36.101 v13.2.0 Section 5.6): 1.4, 3, 5, 10, 15, 20 MHz. Resource blocks are defined as a collection of 12 subcarriers over one subframe (at MAC level). The baseline subcarrier value of LTE is 15 kHz, such that the RB value is 180 kHz. TS 36.101 specifies available bands also for carrier aggregation (CA), proximity services (ProSe) i.e. D2D (see also Section 5.9.2), etc. [87].

4.1.1.1 Channel state information measurements in current standards

4.1.1.1.1 User end measurements

Reference signal receive power (RSRP)

3GPP definition**:** *Reference signal received power (RSRP), is defined as the linear average over the power contributions (in [W]) of the resource elements that carry cell-specific reference signals within the considered measurement frequency bandwidth. The reference point for the RSRP shall be the antenna connector of the UE.*

To measure RSRP, cell-specific reference signals transmitted by eNodeB on the first antenna port only or on the first and second antenna ports are used. If the second antenna port shall be used as well, UE is informed about it by eNodeB over broadcast channel (using System information block 3 transmitted over Broadcast Channel (BCH)). The measured power of the resource element that carries cell-specific reference signals is determined from the energy received during the useful part of the OFDM symbol, excluding the cyclic prefix. It is left up to UE vendor, whether UE has to measure on every resource element containing reference signals within the measurement period.

The following accuracy requirements were set in 3GPP 36.133:

- For intra-frequency RSRP measurements under normal conditions:
  - absolute measurement accuracy: between ± 6 and ± 8 dB
  - relative (between two intra-frequency measurements) accuracy: between ± 2 and ±3 dB
- For inter-frequency RSRP measurements under normal conditions:
  - absolute measurement accuracy: between ± 6 and ± 8 dB

- relative (between an intra-frequency measurement and inter-frequency measurement) : ±6 dB

Absolute accuracy reflects the difference between the actual and the correct measurements whereas relative accuracy reflects the error when comparing the levels of two cells. The reporting range of RSRP is defined from -140 dBm to -44 dBm with 1 dB resolution. RSRP values are reported using the mapping of measured quantity to an integer value as defined in 3GPP 36.133.

Table 4-1 RSRP measurement report mapping

Reported value	Measured quantity value	Unit
RSRP_00	RSRP < -140	dBm
RSRP_01	-140 ≤ RSRP < -139	dBm
RSRP_02	-139 ≤ RSRP < -138	dBm
…	…	…
RSRP_95	-46 ≤ RSRP < -45	dBm
RSRP_96	-45 ≤ RSRP < -44	dBm
RSRP_97	-44 ≤ RSRP	dBm

The RSRP measurements are used in UE's connected mode (for handover decision, path loss estimation used UL power control) and in idle mode (for cell selection/reselection). Starting from 3GPP Release 10, RSRP reports can also be used for adaptive component carrier configuration.

Reference signal received quality (RSRQ)

3GPP definition**:** *Reference signal received quality (RSRQ) is defined as the ratio N×RSRP/ (E-UTRA carrier RSSI), where N is the number of RB's of the E-UTRA carrier RSSI measurement bandwidth. The measurements in the numerator and denominator shall be made over the same set of resource blocks.*

The E-UTRA carrier received signal strength indicator (RSSI) parameter describes the total received power (incl. useful signal from serving cell and co-channel interference from all sources) measured on all resource elements (not just on the resource elements containing reference signals) in OFDM symbols containing reference signals for antenna port 0.

RSRQ measurement allows getting an estimate of the received signal quality considering both signal strength and interference. In the first release of LTE (Release 8), RSRQ measurement was defined only for connected mode, but starting from release 9, it can be used also in idle mode for cell reselection.

The following accuracy requirements were set in 3GPP 36.133:

- For intra-frequency RSRQ measurements under normal conditions:
  - absolute measurement accuracy: between ± 2.5 and ± 3.5dB
- For inter-frequency RSRQ measurements under normal conditions:
  - absolute measurement accuracy: between ± 2.5 and ± 3.5 dB
  - relative (between an intra-frequency measurement and inter-frequency measurement) : between ±3 and ±4 dB

The reporting range of RSRQ is defined in 3GPP 36.133 from -19.5 dB to -3 with 0.5 dB resolution.

Table 4-2 RSRQ measurement report mapping

Reported value	Measured quantity value	Unit
RSRQ_-30	RSRQ < -34	dB
RSRQ_-29	-34 ≤ RSRQ < -33.5	dB
…	…	…
RSRQ_-02	-20.5 ≤ RSRQ < -20	dB
RSRQ_-01	-20 ≤ RSRQ < -19.5	dB
RSRQ_00	RSRQ < -19.5	dB
RSRQ_01	-19.5 ≤ RSRQ < -19	dB
RSRQ_02	-19 ≤ RSRQ < -18.5	dB
…	…	…
RSRQ_32	-4 ≤ RSRQ < -3.5	dB
RSRQ_33	-3.5 ≤ RSRQ < -3	dB
RSRQ_34	-3 ≤ RSRQ	dB
RSRQ_35	-3 ≤ RSRQ < -2.5	dB
RSRQ_36	-2.5 ≤ RSRQ < -2	dB
…	…	…
RSRQ_45	2 ≤ RSRQ < 2.5	dB
RSRQ_46	2.5 ≤ RSRQ	dB

Measuring RSRQ gives information about interference in addition to the signal strength of the reference signals (measured by RSRP).

Channel quality indicator (CQI)

CQI report contains measurements of the DL radio conditions. CQI is provided to the eNodeB scheduler as a coded value of a combination of modulation and coding rate that corresponds to the largest transport block that can be received with error rate not exceeding 10 %. The exact CQI definition can be found in Section 2.7.3 in TS 36.213 [88].

"Based on an unrestricted observation interval in time unless specified otherwise in this subclause, and an unrestricted observation interval in frequency, the UE shall derive for each CQI value reported in uplink subframe *n* the highest CQI index between 1 and 15 which satisfies the following condition, or CQI index 0 if CQI index 1 does not satisfy the condition: "A single PDSCH transport block with a combination of modulation scheme and transport block size corresponding to the CQI index, and occupying a group of downlink physical resource blocks termed the CSI reference resource, could be received with a transport block error probability not exceeding 0.1."

The CSI reference resource, in the frequency domain, is defined by the group of downlink physical resource blocks corresponding to the band to which the derived CQI value relates. In the time domain, the CSI reference resource depends on a multiplicity of factors like, transmission mode (i.e. MIMO mode), configurations or reporting, ICIC or eICIC techniques, eMBMS etc., but in general it coincides with a valid DL or special subframe. The standard specifies also the conditions which the UE shall assume for its calculation (see Section 2.7.3 in TS 36.213 [88]).

It hence depends on the channel conditions, e.g. the signal-to-interference-and-noise ratio (SINR), experienced over the radio link, but also on the equipment parameters and the receiver type. Different receiver algorithms may give for the same channel condition a different CQI. Hence, CQI provides indirect information about channel state since it does not describe the channel state itself but the effect of it on the achievable throughput of the measuring UE, which depends on its receiver algorithms.

Notice also that CQI information is deduced from the PDSCH (data channel) and hence in general it is risky to use CQI to say something about the reception quality of the PDCCH, the control channel.
In LTE, CQI is coded on 4 information bits, according to the following tables which report Table 7.2.3-1 and Table 7.2.3-2 in TS 36.213 v12.8.0. The former is for reporting CQI based on QPSK, 16-QAM and 64-QAM, the latter for reporting CQI based on QPSK, 16-QAM, 64-QAM and 256-QAM.

Table 4-3 Table 7.2.3-1 in TS 36.213 v12.8.0, CQI index mapping for QPSK, 16-QAM, 64-QAM

CQI index	modulation	code rate x 1024	efficiency
0	out of range		
1	QPSK	78	0.1523
2	QPSK	120	0.2344
3	QPSK	193	0.3770
4	QPSK	308	0.6016
5	QPSK	449	0.8770
6	QPSK	602	1.1758
7	16QAM	378	1.4766
8	16QAM	490	1.9141
9	16QAM	616	2.4063
10	64QAM	466	2.7305
11	64QAM	567	3.3223
12	64QAM	666	3.9023
13	64QAM	772	4.5234
14	64QAM	873	5.1152
15	64QAM	948	5.5547

Table 4-4 Table 7.2.3-2 in TS 36.213 v12.8.0, CQI index mapping for QPSK, 16-QAM, 64-QAM and 256-QAM

CQI index	modulation	code rate x 1024	efficiency
0	out of range		
1	QPSK	78	0.1523
2	QPSK	193	0.3770
3	QPSK	449	0.8770
4	16QAM	378	1.4766
5	16QAM	490	1.9141
6	16QAM	616	2.4063
7	64QAM	466	2.7305
8	64QAM	567	3.3223
9	64QAM	666	3.9023
10	64QAM	772	4.5234
11	64QAM	873	5.1152
12	256QAM	711	5.5547
13	256QAM	797	6.2266
14	256QAM	885	6.9141
15	256QAM	948	7.4063

Moreover, a CQI index in the CQI table corresponds to a modulation order, efficiency, and an effective code rate (the effective channel code rate is defined as the number of downlink information bits (including CRC bits) divided by the number of physical channel bits on PDSCH). In order to calculate them, the transport block size must be deduced (it is the goal of the CQI calculation in fact). A CQI index is valid only if it corresponds to a couple of modulation orders and transport block sizes, inside a given number of resource blocks, which is specified in the standard. This combination must be allocable by the eNodeB, and the smallest TBS which gives the code rate closest to the one of the CQI index table is selected. In the CQI table, the column “efficiency” is the modulation order times the coding rate.

The periodicity and frequency resolution to be used by a UE to report CQI are both controlled by the eNodeB in the limits of the flexibility of the standard. In the following we will detail the granularity and periodicity of the CQI reports, up to Rel 10.

The CQI can be reported as wideband CQI (for the whole channel bandwidth) or as sub-band CQI (for a specified part of the bandwidth). Sub-band CQIs help the MAC scheduler to execute frequency-

selective scheduling. Concerning sub-band CQI, it is distinguished in eNodeB-configured and UE-selected sub-band CQI. These two types of sub-bands reporting are configured by the eNodeB RRC signaling. The definition of sub-band in fact depends on the type of reporting.

Aperiodic CQI reporting is a CQI report decided and scheduled by the eNodeB, usually on a PUSCH (data channel):

- Wideband, one CQI value for the whole bandwidth, over 4 bits.
- eNodeB-configured sub-band CQI: the wideband CQI is reported as well as all sub-bands CQI in the system transmission bandwidth. For each sub-band a differential value with respect to the wideband value is sent and coded over 2 bits only, the possible offsets being $\{ \leq -1, 0, +1, \geq +2 \}$ (see standard for exact coding). The sub-band is a group of RBs whose size depends on the system bandwidth. From the rules reported in the standard and in [88], Section 10, we can find the values in Table 3.
- UE-selected sub-band CQI: in this case the UE selects itself *M* sub-bands to be reported, and it sends the wideband CQI, the average CQI over the selected sub-bands and the CQI of the selected sub-bands, as well as their positions through a combinatorial index. The sub-band CQI is coded in two bits with respect to the wideband CQI, the possible offsets being $\{ \leq +1, +2, +3, \geq +4 \}$ (see standard for exact coding). The definition of the sub-band is different from the eNodeB-configured mode. See for the values applied to the transmission bandwidth.

Periodic CQI report instructs the UE to send the reporting on a periodic basis. The configuration of the reporting period and report type is done by the eNodeB RRC. Periodic CQI reports are sent through a PUCCH unless a PUSCH transmission has been scheduled for the UE in one of the periodic subframes. In that case the periodic CQI report is sent on the PUSCH (in Rel 8 and 9 it is not possible for a UE to have a PUCCH and PUSCH active at the same time in order to keep the PAPR of the signal as low as possible at the UE while this possibility has been introduced starting from Rel 10 but on precise conditions, please see the standard). The periodicity values which can be configured are {2, 5, 10, 16, 20, 32, 40, 64, 80, 128, 160} ms, or Off.

Only two types of reporting are possible:

- Wideband, one CQI value for the whole bandwidth, over 4 bits, like for aperiodic reporting.
- UE-selected sub-band CQI: for all downlink (PDSCH) transmission modes. The total bandwidth is divided into sub-bands which are grouped in bandwidth parts. For each bandwidth part, the UE send a CQI value over 4 bits of the selected sub-band inside the bandwidth part, plus the position of the sub-band inside the bandwidth part. Taking the values in the standard (TS 36.213 [88]), resulted in the bandwidth parts given in Table 4-6.

Table 4-5 Size and number of sub-bands for aperiodic CQI reporting in LTE

		eNodeB-configured report		UE-selected Report		
Channel Bandwidth (MHz)	**System Transmission Bandwidth (N_{RB})**	**Sub-band size (RB)**	**Number of sub-bands**	**Sub-band size (RB)**	**Number of sub-bands**	**Number of selected sub-bands**
1.4	6	1 (wideband only)	1	1 (wideband only)	1	1 (wideband only)
3	15	4	4	2	8	3
5	25	4	7	2	13	3
10	50	6	9	3	17	5
15	75	8	10	4	19	6
20	100	8	13	4	25	6

Table 4-6 Size and number of sub-bands for periodic CQI reporting in LTE

Channel Bandwidth (MHz)	**System Transmission Bandwidth (N_{RB})**	**Sub-band size (RB)**	**Number of sub-bands**	**Number of bandwidth parts (*J*)**	**Number of sub-band in bandwidth parts (Nj)**	**Position Index (*L*) Bits**
1.4	6	1 (wideband only)	1	1 (wideband only)	1	0
3	15	4	4	2	2	1
5	25	4	7	2	3-4	2
10	50	6	9	3	3	2
15	75	8	10	4	2-3	2
20	100	8	13	4	3-4	2

In case of MIMO schemes (transmission modes) with multiple codewords like spatial multiplexing if the rank indicator (RI) report is configured and is greater than 1, then CQI for two codewords must be sent. The standard specifies either separate reports or differential reports, depending on the situations and transmission modes. For details of CQI with MIMO modes in LTE and LTE-A, please refer to the standard (TS 36.213 [88]).

When eICIC schemes are active, particular attention is to be given to how CQI is measured. For that reason, with almost blank subframes (ABS) the standard specified restrictions on the measurement subframes so that the UE is able to measure the quality of the channel with the reduced interference generated by the eICIC technique.

4.1.1.1.2 eNodeB measurements

Received interference power

The uplink received interference power is measured with granularity of one resource block and represents the interference power including thermal noise over resource block. This measure was introduced for ICIC. It is exchanged between neighboring eNodeBs over X2 interface as a part of Load Information message. Neighboring eNodeBs can coordinate MAC scheduling by avoiding scheduling specific resource blocks to cell-edge UEs, when neighboring eNodeB reports high uplink interference level on these resource blocks.

Thermal noise power

Thermal Noise Power is the uplink thermal noise power measured across all resource blocks. It is used in combination with Received Interference Power to derive an Interference over Thermal (IoT) metric. IoT describes uplink interference relative to thermal noise.

Sounding reference signal (SRS)

SRS is used to estimate uplink channel quality across a wider bandwidth than the current uplink data transmission (PUSCH) or when UE has no uplink data transmissions. SRS helps the MAC scheduler to execute frequency-selective scheduling.

UL Rx SINR

UL Rx SINR is SINR measured on each reception of the uplink data transmission (PUSCH).

4.1.1.2 Positioning-supporting measurements

Starting from 3GPP Release 9 and following [166], the following measurements were defined as:

- **eNodeB Rx-Tx time difference:** the time difference between a UL radio frame and the transmission time of the corresponding DL frame
- **Angle of arrival:** the estimated angle of a UE with respect to the reference direction (geographical north)
- **Timing advance:** is used for calculating the UE distance from a cell. Type 1 (Release 9): TA = (eNode B Rx-Tx Time Difference) + (UE Rx-Tx Time Difference), Type 2 (Release 8): TA = the eNB Rx–Tx timing difference.

4.1.1.3 CSI reports

The CSI reports include three main components:

- CQI, the channel quality indicator, reported by an UE to the serving base station, and providing the code of the highest modulation and coding rate achievable with an error rate below 0.1;
- PMI, a pre-coding matrix indicator;
- MIMO rank indicator.

4.1.1.3.1 CQI reports

LTE Release 13 standardized concept supports the transmission of channel state information (CSI) measurements, based on so-called CSI Processes (see 3GPP TS 36.213 v 12.6.0 [88]), each CSI process including a non-zero power (NZP), CSI reference signal (CSI-RS) transmission and one or more CSI-RS zero power (ZP) transmissions. Based on these CSI-RS signals transmitted by the eNB, an UE can assess the channel quality indicator (CQI) including the recommended modulation and coding scheme.

Given the fact that, in deployments with multiple base stations, a CQI Report is needed for each combination of the transmitting base stations, the CQI approach conducts to numerous CSI processes. For example, if we insist that each combination of interferers should be identifiable, then for 3 interferers there are 7 possible combinations of CSI signals, conducting to 7 CSI processes. In Release 13 it was decided in 3GPP to limit the number of CSI processes for an UE to 3, thus allowing distinguishing only between combinations of 2 interferers. This will not be enough in a very dense deployment, as expected for 5G, creating a need for replacing the CQI approach when building the network graph.

It should be also noted that the energy per CSI-RS symbol is constant over the entire channel bandwidth; in case of power control per RB there will be no indication of the expected CQI or RI. In addition, in case of coordinated CSI measurements the overhead of CSI-RS extensive usage has reached approx. 17%, as shown in the 3GPP RAN3 contribution R3-150723.

4.1.1.4 Parameters concerning the frame structure

LTE specified two types of frames, Type 1 for FDD and type 2 for TDD (see TS 36.211 v12.8.0 [89]). Both frames are 10 ms long, divided in 10 subframes of 1 ms. In FDD, the subframe is divided in 2 slots of duration 0.5 ms = 15360 T_s, where T_s is the sampling time which is a constant in the LTE system: T_s = 1/(15000 x 2048) s i.e. approximately 32,5521 ns. In TDD, the frame periodicity is respected however subframes are distinguished in D (for DL, from eNB to UE), U (for UL, from the UE to the eNB) and S for "special subframe", which is a subframe for transmission direction switch. Special subframes are composed of three fields: the Downlink Pilot Timeslot (DwPTS) used in DL, a Guard Period (GP), used for switching of RF radio state (from transmitting to receiving and vice-versa) and for giving a guard time for compensation of propagation delays; the Uplink Pilot Timeslot (UpPTS) used in UL.

Table 4-7 reports 7 possible configurations of TDD frames, with their minimum periodicity as written in the standard. These configurations have an influence in the timing of PHY procedures like ACK/NACK acknowledgment, resource allocation requests, etc.

Table 4-7 UL-DL configuration for Type 2 frame (TDD), as in Table 4.2-2 in TS 36.211 v12.8.0

Uplink-downlink configuration	Downlink-to-Uplink Switch-point periodicity	Subframe number									
		0	1	2	3	4	5	6	7	8	9
0	5 ms	D	S	U	U	U	D	S	U	U	U
1	5 ms	D	S	U	U	D	D	S	U	U	D
2	5 ms	D	S	U	D	D	D	S	U	D	D
3	10 ms	D	S	U	U	U	D	D	D	D	D
4	10 ms	D	S	U	U	D	D	D	D	D	D
5	10 ms	D	S	U	D	D	D	D	D	D	D
6	5 ms	D	S	U	U	U	D	S	U	U	D

The special subframe can be tuned to different values according to the operational needs of the network (see Table 4.2-1 in [89]). With this configuration it is possible to accommodate different guard time, for as a function of the area of the deployed cells. Notice also that, when relay nodes are present, specific rules, in particular for TDD, are present in order to organize the transmission.

In the frequency domain and without Carrier Aggregation (CA), the configurable channel bandwidths of LTE are {1.4, 3, 5, 10, 15, 20} corresponding to {6, 15, 25, 50, 75, 100} RB. Another important parameter at the PHY layer is the cyclic prefix, which is defined in the following table. The value of 7.5 kHz is to be used only for subframes for eMBMS.

Table 4-8 Subcarrier spacing and corresponding CP as in Table 6.12-1 in TS 36.211 v12.8.0. CP length is given in sampling times

Configuration		Cyclic prefix length $N_{\mathrm{CP},l}$
Normal cyclic prefix	$\Delta f = 15\,\mathrm{kHz}$	160 for $l = 0$ 144 for $l = 1,2,\ldots 6$
Extended cyclic prefix	$\Delta f = 15\,\mathrm{kHz}$	512 for $l = 0,1,\ldots 5$
	$\Delta f = 7.5\,\mathrm{kHz}$	1024 for $l = 0,1,2$

Transmission bandwidth of the system is sent over the Master Block Information (MIB). For D2D communications, named inSlide Link (SL) communications in LTE, there is also a master information block called MIB-SL, which is sent with a periodicity of 40 ms and which contains the bandwidth, the TDD configuration of the SL frame, flag for incoverage operation, etc. For more details, please see Section 5.8.2.3. The configurable bandwidth are {6, 15, 25, 50, 75, 100} RB as for the global MIB. In general System Information (SI) can change with a certain periodicity by following a certain procedure which notifies to the UEs the next change beforehand (please check [90] for further details).

4.1.2 PHY layer – WiFi

4.1.2.1 Received signal strength indicator (RSSI)

The IEEE 802.11 standard does not specify in detail how RSSI shall be computed by WiFi-compliant devices. More specifically according to the standard the RSSI shall measure the RF energy received by the PHY layer. As a result RSSI is not a good indicator of the desired signal power level since it indicates the sum of the desired signal + noise + interference power.

RSSI measurements in a WiFi system are given in arbitrary units as an 8-bits positive integer. As an example, Cisco Systems cards report 101 different power levels (between 0 and 100). Atheros (now Qualcomm) Wi-Fi chipsets instead return RSSI values between 0 and 127 (0x7f) with 128 (0x80) indicating an invalid value. There is no standardized way of converting RSSI readings and power level in either mW or dBm.

Different vendors and even different releases of WiFi drivers can (and in general do) use different assumption regarding ranges and granularity. Moreover, RSSI measurements are acquired only upon reception of a frame and are taken only on the preamble of such frame not on the entire frame. The preamble is always transmitted at the basic rate (1-2 Mb/s for IEEE 802.11b, 6 Mb/s for IEEE 802.11a/g, MCS0 for IEEE 802.11n/ac).

Since the way of measuring the RSSI is so vague, it is not possible to compare RSSIs from different clients and perhaps even from different channels or PHYs within the same client (i.e. it cannot be used as a comparison between DSSS and OFDM PHYs). Due to the reasons above, RSSI is a poor indicator for typical network management tasks such as handover or load balancing.

4.1.2.2 Received channel power indicator (RCPI)

With the 802.11k amendment to the 802.11 standard, a new channel quality indicator has been introduced, namely the Received Channel Power Indicator (RCPI). The RCPI is an indicator of the received power in a given channel measured at the antenna connector. The RCPI is measured over the entire frame and is a monotonically increasing logarithmic function of the received power in dB. The RCPI is reported again as an 8-bit positive integer (i.e. unsigned) in the range between 0 and 255. There is no unit of measurement associated to the RCPI. An RCPI of 0 (zero) corresponds to a received power of lower or equal to -82dBm while an RCPI of 255 corresponds to a received power higher or equal to -18.5 dBm.

4.1.2.3 Wireless LAN radio measurements (802.11k)

The 802.11k standard introduces the following measurements:

Beacon report. A beacon report allows an AP to poll a wireless client for the list of networks or BSSIDs from which it can detect a beacon or a probe response.

Frame report. Each frame report element contains one or more Frame Report quadruplets, each consisting of the Number of Frames, BSSID and Transmit Address. Each quadruplet summarizes the traffic from one transmit address.

Channel load. If a station accepts a Channel Load Request it shall respond with a Radio Measurement Report. It specifies channel busy condition that the measuring station's view of what happens on specified channel. Channel Number, Channel Band, Actual Measurement Start Time, Measurement Duration and Channel Busy Fraction are included in Channel Load Report.

Noise histogram. If a station accepts a Noise Histogram Request it shall respond with a Radio Measurement Report frame containing one Measurement (Noise Histogram) Report element. The Noise histogram report in the radio measurement category includes only non-802.11 energy in its result by sampling the channel only when CCA indicates that no 802.11 signal is present.

Station (STA) statistics. A STA receiving a STA Statistics Request shall respond with a Radio Measurement Report frame containing one or more Measurement (STA Statistics) Report elements. Each element contains a group of STA Statistics.

Neighbour report. A Neighbour report is sent by an AP and it contains information on known neighbours of the AP. A wireless client requesting a Neighbour report shall send a Neighbour Report Request frame to the AP it is associated with. An AP accepting a Neighbour Report Request shall respond with a Neighbour Report Response frame containing one or more Neighbour Report elements. This information may be used by a STA when scanning for target APs during BSS-transitions or for other purposes.

RCPI. The RCPI element indicates the received channel power indication of the last received packet from a given station. Within a probe response the RCPI element would carry the RCPI value measured on the probe request. This is essentially a way for the AP to ask a wireless client to perform some active measurements.

4.1.3 MAC layer – LTE

4.1.3.1 Scheduling and possible abstractions

While typically the scheduling in a LTE system is done at each eNB, in COHERENT a Central Controller and Coordinator (C3) is considered which, subject to backhaul capacity and delay, can coordinate the resource allocation between R-TPs (real transmission points). The C3 can even provide the full scheduling information. In a mobile system, a real R-TP is a transmission point including at least one antenna designed for at least one frequency band allocated to the mobile service.

A scheduler, function of information regarding the scheduling requests and the available resources, assigns the UE to be served by one or more cells, where a cell is defined per node and frequency range. The exact time-frequency-power-space resources to be used for a packet transmission or reception can be scheduled by a node or by the C3. Hence, the scheduling is done at two levels:

- Serving node selection;
- Time-frequency-power resource allocation for each serving node.

4.1.3.1.1 Scheduling granularity

In LTE the scheduling granularity in time domain is in general 1 ms (subframe duration) while the frequency granularity is a PRB (physical resource block) or a subband or a full channel. To transform this information in a RAT-independent mode should be mentioned the time granularity e.g. 1 ms, 0.9 ms, LTE_full_subframe, and the frequency granularity (e.g. 180 kHz, LTE_PRB, LTE_Subband, LTE_reduced_subband, etc.). Alternatively, the subframes could be, as in LTE, defined by the system frame number and the subframe number. This allows, for synchronized networks, a practical good time reference.

4.1.4 MAC layer – WiFi

4.1.4.1 Scheduling granularity: time

Broadly speaking, there are two main families of strategies to allocate resources in a wireless network: scheduled access and random access. In the former case, resources for a wireless link are allocated in the time, frequency, and space domains. In the latter case, a common random access scheme for medium access is used by all participating wireless clients in order to reduce collisions. LTE belongs to the former family and uses OFDMA as (scheduled) medium access scheme while WiFi belongs to the latter family and exploits CSMA/CA as (random) medium access scheme.

4.1.4.2 Scheduling granularity: space

The minimum allocation unit in a WiFi system is the channel identified by a frequency band, a time interval, and the AP at which it is available. In order to have a consistent naming with LTE system, we decide to rename the WiFi channel as resource block.

Each resource block is fully described by a 2–tuple <f, t>, where f is the frequency band, and t is the time slot. A frequency band is a 2–tuple <c, b> where c and b are, respectively, the center frequency and the bandwidth.

For example, the resource pool made available by an 802.11n AP tuned on channel 36 and supporting 40 MHz–wide channels is represented by the tuple ((36, HT 40), ∞). The prefix HT is used to indicate that this band supports the high throughput modulation and coding scheme (MCS). Resource Blocks can also be blacklisted preventing the COHERENT C3 from using them.

Each WiFi AP or client in the network can support as many resource blocks as the number of its WiFi interfaces. For example a dual band WiFI AP would expose to the COHERENT controller the following list of resource blocks: ((6, HT 20), ∞) and ((36, HT 20), ∞). This models the fact that the AP can support at the same time two operating bands. The other channels supported on each band would be exposed to the COHERENT controller as black listed resource blocks modelling the fact that the AP cannot be tuned on more than two channels at the same time.

Notice that the proposed model does not forbid the same resource block to be assigned to multiple Light Virtual Access Points (LVAP) in that this could in general result in a valid resource allocation scheme if, for example, the LVAPs are sufficiently separated in space or if suitable inter–cell interference coordination (ICIC) schemes are employed. Similarly, wireless networks using random access protocols, such as CSMA/CA, effectively schedule multiple transmissions on the same resources in frequency and time with the aid of suitable back–off and retransmission schemes to handle collisions.

Notice how the same formulation can be used also to model the channels and bands supported by a wireless client. For example, a resource request could be represented by the following tuple ((1, 20), ∞). This allows us to express resource allocation problems as an intersection between the resource blocks available in the network, and the resource blocks supported or requested by a client. Information on the link quality experienced by the requesting client on the matching resource blocks can be used to further filter the set of candidate resource blocks according to application–level parameters. A non–empty intersection set of resource blocks signifies that a valid solution for the resource allocation problem has been found.

The final set could be composed of multiple resource blocks possibly scheduled at different APs and on different frequency bands or timeslots. The support for such a scenario depends on the actual implementation of the client radio interface. For example, in an LTE network, a client could accept multiple resource blocks possibly scheduled at different eNBs and on different frequencies modelling the technique known as cooperative multi–point, or CoMP. This model also effectively decouples uplink and downlink allowing clients to be scheduled at different APs on the uplink and on the downlink directions (if the feature is supported by the link–layer technology).

An example of a simple resource allocation scenario for a WiFi WLAN is shown in the Table below. Here Pn are the network resource blocks and Pl are the resource blocks supported by a client. It is easy to see that the intersection Pn ∩ Pl produces a non–empty set composed of two resource blocks scheduled at two different APs (W1, W2). Due to the fact that WiFi does not allow scheduling one client on more than one AP, the final resource allocation decision will be a single resource block selected using criteria such as the channel quality experienced by the client on the matching resource blocks and/or the specific application–level requirements.

Table 4-9 Resource block tuples

Network	Client	Intersection
W 1 ((6, HT 20), ∞)	L 1 ((36, HT 20), ∞)	W 1 ((36, HT 20), ∞)
W 1 ((36, HT 20), ∞)	L 1 ((48, HT 20), ∞)	W 2 ((36, HT 20), ∞)
W 2 ((1, HT 20), ∞)	L 1 ((54, HT 20), ∞)	
W 2 ((36, HT 20), ∞)		
W 3 ((11, 20), ∞)		
W 3 ((36, 20), ∞)		

4.1.4.3 Abstraction: channel quality map

From the perspective of a client, it is not important to which Wireless Termination Point (WTP) it is attached but what communication QoS it can obtain. Such information translates, at the physical layer, into the transmission efficiency of the radio channel linking interference with parameters such as packet error rate. The channel quality and interference map allow the control logic to reason about the channel quality and interference experienced by clients and to assign resources accordingly.

The channel quality map abstraction provides network programmers with a full view of the network state in terms of channel quality between clients and APs over the available resource blocks. Let G = (V, E) be a directed graph, where $V = V_{AP} \cup V_{STA}$ is the set of APs and wireless clients in the network, and E is the set of edges or links. An edge $e_{n,m,i} \in E$ with $n, m \in V$ exists if m is within the communication range of n over the resource block $i \in P$. A weight $q(e_{n,m,i})$ is assigned to each link representing the channel quality between the two nodes.

The channel quality map abstractions can be used to select the resource blocks that can satisfy the requirements of a client by intersecting the set of available resource blocks in the network (P_N) with the requested resource blocks (P_L). The set of available resource blocks is obtained as: $P_N = W_1 \cup W_2 \cup \cdots \cup W_N$. The matching resource blocks M are then given by: $M = P_N \cap P_L$. The list of resource blocks M' that satisfies a certain interference level condition, such as the signal to interference plus noise ratio (SINR) between the client n and the WTP m on the Resource Block i being greater than a certain threshold t, is given by: $M' = \{i \in M : SINR(e_{n,m,i}) > t\}$ where $SINR(e_{n,m,i})$ can be estimated via the edge weights in the Channel Quality Map. M' is the empty set if a valid resource allocation is not found.

Figure 4-1 sketches a sample channel quality map. Notice how in this case RSSI has been taken as a measure of channel quality. The map depicted here is currently implemented in the COHERENT SDK and has been leveraged in order to implement a mobility management control application for WiFi networks. The map is built using as an input the Wireless LAN radio measurements reports presented in Section 4.1.2.3. In particular, a monitor interface created on top of each physical radio available at a WiFi AP is used to extract the signal strength field present in the radio tap header of every decoded WiFi frame. In order to distinguish between WiFi APs and WiFi Stations, the To–DS, From–DS, frame type, and frame sub–type fields present in the 802.11 header are used. For each neighbour within the decoding range, an agent running within each WiFi AP computes the average of the RSSI over windows of 500 ms, an exponential weighted moving average and an *N*–point smoothed moving average are also maintained. The latter two filters have been selected because they can reduce the noise while being responsive to RSSI changes. Such a property is useful when dealing with fast– fading affected RSSI signals. At the same time, fast response allows to promptly react to changes in the channel conditions.

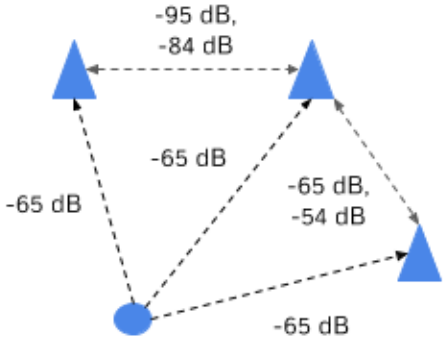


Figure 4-1 An example of RSSI Map

Table 4-10 summarizes the RSSI map request message parameters sent by the controller to the WiFi AP over the COHERENT C3 southbound interface (the details of the encoding will be reported in D2.3).

Table 4-10 Southbound interface: RSSI map request

Parameter	Type	Description	Example
addrs	Ethernet Address	Neighboring stations/Aps to track	'FF:FF:FF:FF:FF:FF'
block	Resource Block	The channel to monitor	('04:f0:21:09:f9:96', 36, L20)

In the above example specifying ff:ff:ff:ff:ff:ff will return the RSSI of any station within the decoding range of WTP 04:F0:21:09:F9:96 on channel 36. The channel quality map tracks the RSSI level of any active WiFi device including the ones belonging to networks that are not under the administrative domain of the COHERENT C3. This includes also wireless clients that are not associated to any network but have their wireless interface active (WiFi clients periodically broadcast Probe Request messages in order to discover available networks). A sample output of the primitive is reported below as a JSON document. In this case the station a0:d3:c1:a8:e4:c3 is a neighbour of the WTP 04:f0:21:09:f9:96 on the 802.11a channel 36. The report includes, besides the previously described averages, also the total number of frames received since the query was created (hist_packets) together with the average (last_rssi_avg), the standard deviation (last_rssi_std), and the number (last_packets) of RSSI measurements taken during the last observation window.

{ "a0 :d3: c1 : a8 : e4 : c3": { "ewma rssi": −82, " hist packets ": 15810, " last packets ": 10, " last rssi avg ": −79, " last rssi std ": 7 "sma rssi": −82, } }

The channel quality map can be accessed at the frame–level granularity providing the network programmer with a real–time picture of all link–layer events. Each WTP tracks the following meta–data associated to link–layer events:

- Transmitter address. The MAC address of the transmitter.
- TSFT. The 802.11 MAC's 64-bit time synchronization function timer. Each frame received by the radio interface is timestamped with a 1 µsec resolution clock by the 802.11 driver.
- Sequence, The 802.11 MAC's 16-bit sequence number. This counter is incremented by the transmitter after a successful transmission.
- The frame RSSI (in dB), rate (in Mb/s), length (in bytes), and duration (in µsec). The collected traces are then periodically delivered to the SD–RAN Controller where they are synchronized to a common time reference.

Notice that, only successful unique frames transmissions are recorded, i.e. frames with incorrect checksum and/or Physical Layer Convergence Protocol (PLCP) header as well as retransmitted frames are ignored. The collected meta–data can be exploited for several purposes.

Table 4-11 Southbound interface: Transmission summary request

Parameter	Type	Description	Example
addrs	Ethernet Address	Neighboring stations/Aps to track	'FF:FF:FF:FF:FF:FF'
limit	Int	Number of transmission reports to be generated	-1
every	Int	Reporting interval (in ms)	2000
block	Resource Block	The channel to monitor	('04:f0:21:09:f9:96', 36, L20)

The statement above generates a periodic callback when a traffic trace has been received from a WiFi AP. The traffic traces include all the link–layer events within the decoding range of the WiFi AP. The limit parameters instruct the WiFi AP to send only a specified number of reports after which the operation is stopped and all the allocated data structures on the WiFi AP are freed. Specifying −1 results in the WiFi AP sending traffic traces forever. Specifying ff:ff:ff:ff:ff:ff as addrs will generate transmission summaries for all stations in the network. Notice that the entire data structure containing a frame meta–information is 18–bytes long. A saturated 54–Mb/s channel can deliver up to 2336 frames/s[3]. In such a scenario the system would generate 42048 bytes of meta–data per second per radio interface which correspond to a signalling bandwidth between the WTP and SD–RAN Controller of about 336 kb/s which is negligible considering the widespread adoption of Gigabit Ethernet in enterprise networks.

Finally, the RSSI primitive allows the programmer to trigger a callback the first time the RSSI of a station verifies a certain condition at any WiFi AP in network.

Table 4-12 Southbound interface (RSSI trigger request)

Parameter	Type	Description	Example
addrs	Ethernet Address	Neighboring stations/Aps to track	'FF:FF:FF:FF:FF:FF'
relation	Enum('LT', 'GT', 'EQ')	Condition to be verified	'LT'
value	Int	The RSSI level	-70
block	Resource Block	The channel to monitor	('04:f0:21:09:f9:96', 36, L20)

After the trigger has fired the first time and as long as the RSSI remains below −70 dBm, the callback method is not called again by the same WiFi AP, however the same callback may be triggered by other WiFi APs. Specifying ff:ff:ff:ff:ff:ff as lvaps will trigger the callback when the RSSI of any client at any WiFi AP is below −70 dBm.

4.1.4.4 Abstraction: link statistics map

Figure 4-2 sketches a sample Link Statistics Map. This is a variant of the channel quality map where edges are annotated with the link delivery statistics. For each supported MCS, the map reports the delivery probability, and the link throughput.

[3]At 54 Mb/s an 802.11a radio encodes 216 bits/symbol. The OFDM encoding then adds 6 more bits at the end of the frame, so a maximum length frame of 1536 bytes becomes a string of 12288 bits plus 6 trailing bits. The total 12294 bits can be encoded with 57 symbols each requiring 6μsec for transmission. As a result, ignoring backoff, the minimum time required for transmitting this frame is: DIFS (34μsec) + DATA (248μsec) + ACK (24μsec) = 322 μsec, which corresponds to 2336 frames/second.

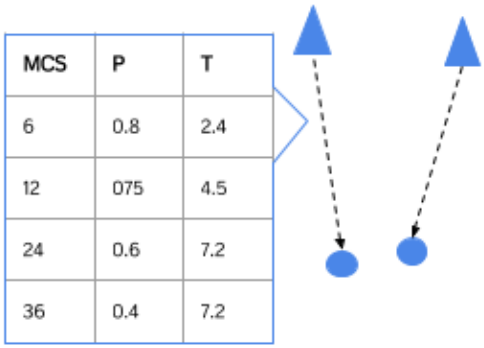


Figure 4-2 An example of link statistics map

Table 4-13 Southbound interface (link statistics map request)

Parameter	Type	Description	Example
addrs	Ethernet Address	Neighbouring stations to track	'04:11:af:4f:34:0d'

For each supported MCS the average delivery probability and the expected throughput in the last observation window are reported. Moreover, also the total numbers of successful and failed transmissions for the wireless clients are reported.

4.1.4.5 Abstraction: traffic matrix

The traffic matrix abstraction allows programmers to track the traffic exchanged by a certain wireless client and to use binning in order to aggregate such information by frame length (useful in wireless networks due to the fact that short packets incur higher transmission overheads). An example for traffic matrix graph is shown in Figure 4-3.

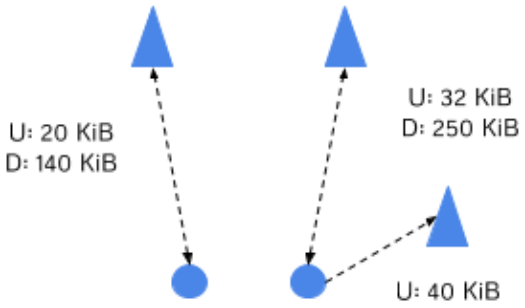


Figure 4-3 An example of traffic matrix map

Table 4-14 Southbound interface (link statistics)

Parameter	Type	Description	Example
addrs	Ethernet Address	Neighboring stations to track	'04:11:af:4f:34:0d'
bins	list	Bins to be used for classification.	[512, 1472, 8192]

4.1.4.6 Resource allocation model

Links in a wired network, e.g. a switched Ethernet LAN, are essentially deterministic and the status of a port in a switch is binary, i.e. active or not active. While some Ethernet switches can select the transmission rate (10, 100, 1000 Mb/s), this feature is aimed at reducing power consumption when the traffic load is low and not as a mechanism for coping with fluctuations in the channel quality. In

contrast, links in a wireless network are stochastic and, as a result, the physical layer parameters that characterize the radio link between a client and an AP, such as transmission power, modulation and coding schemes, and MIMO configuration must be adapted according to the actual channel conditions.

Such level of adaptation requires real–time coordination between LVAPs and WTPs and can only be implemented near the air interface. The radio port abstraction allows the COHERENT Controller to reconfigure or replace a certain control policy if its optimal operating conditions are not met.

A port is defined by a 3–tuple <p, m, a> where p is the transmission power, m is the set of available MCS, and a is the MIMO configuration (number of spatial streams). For example, in the case of an 802.11n network, assigning the port configuration: <30,(0:7),1> to a client means that the AP will use a fixed transmission power of 30 dBm, the set of MCS between 0 and 7, and single antenna configuration for its communication with the LVAP. The port abstraction allows fast timescale adaptations (MCS adaptation in this example) to be delegated to a local controller located near the AP or to the AP itself. Finally, since a Port specifies the configuration of the link between an AP and a client, an AP will have as many Port configurations as the number of client it is currently managing. Conversely a client will have a single Port configuration (this is due to the fact that a client can be associated to one AP at the time).

A traffic graph can be derived from all the Ports defined in a network modelling the application level traffic between client and APs and vice-versa. This graph essentially results in the traffic matrix for a network and can be effectively used in mobility management and load balancing applications.

4.2 Possible new metrics and approaches for constructing network graphs

Some ideas of possible new metrics for constructing network graphs are presented in this subsection. The ideas are then detailed and tuned for relevant control applications in Section 5.

4.2.1 Probabilistic input to network graphs

In addition to abstracting over existing low-level measurements, the network graph should also represent the network state through the use of more sophisticated metrics that provide reliable network observations obtained in a scalable and resource-efficient way. The latter can be achieved by developing distributed mechanisms capable of measuring and aggregating metrics at various levels in the network. Measurements and light-weight analytics performed locally in the equipment (e.g. eNB) and close to the data source, can effectively support self-organization and local coordination among network entities while reducing the signaling overhead. Further, probabilistic models can generally provide more reliable representations of the local network behavior in a compact form, as uncertainty in the observations can be taken into account. We see that distributed and probabilistic models are specifically useful for self-organizing systems and distributed control mechanisms, which should be scalable, agile and stable in order to be effective [40]. The output of the distributed mechanisms and node-local analytics can be used as inputs for analysis at higher levels, for the purpose of supporting mobility management and network optimization. Moreover, we see that probabilistic models in many cases can support increased programmability and higher-level abstractions through the means of configuration based on probabilistic requirements rather than in terms of low-level parameters.

We investigate distributed approaches for producing computationally light-weight models of the network state which can be represented in a compact form as part of the network graph (see Sections 5.2 and 5.6). More specifically, we investigate probabilistic abstractions for representing and balancing load in LTE systems and representing link performance in WiFi networks based on available metrics such as RSSI and PDR. The probabilistic approach under development is based on aggregation and modelling of measurements stored in node-local data counters that are used to estimate a model of the observed network performance. The probabilistic approach generally offers three main benefits: 1) local processing with low signalling and measurement overhead; 2) high-granular observations of the network performance at various time-scales; and, 3) representative models that can account for uncertainty compared to raw counter readings or momentaneous observations. The two-step approach described in [150][203][204] encompasses: 1) local aggregation of measurements and/or readings from

logs or counters at short time scales (e.g. seconds or subsecond level); followed by: 2) analytics carried out either directly on the node (depends on computational capabilities) or externally at longer time-scales (seconds or minutes). The viability of the approach is currently evaluated using real data for modelling and representing aspects of link performance at a higher-level of abstraction in wired [203] and wireless networks (Section 5.2). Further, we employ and elaborate on the principles of estimating and using risk as a metric for distributed load balancing in Section 5.6 based on [150][203].

4.2.2 Coupling loss

We target a method for creating the network graph which is scalable and can easily accommodate a variable number of active interferers. The C3 should add the interference from different sources to get the total interference at receiver and for this it should use the power in W.

In a network graph the coupling loss could be used as a weight of each branch. For deducting the coupling loss we consider the simplest form of a link budget, in dB:

$$P_R = P_T + GT - PL - OL + GR,$$

where P_R, P_T are respectively the transmit and receive powers, *GT*, *GR* are respectively the transmit and receive antenna gains, and *PL*, *OL* are respectively path loss and other losses (connectors, cables, body absorption, fading, etc.).

We define the linear coupling loss (CL) as the power P_T (W) at the antenna connector of the transmitter divided by the power P_R *(W)* at the antenna connector of the receiver:

$$CL(\mathrm{W}) = P_T(\mathrm{W}) / P_R(\mathrm{W})$$

Relative to the path loss (PL), CL in dB is given by the logarithmic equation:

$$CL(\mathrm{dB}) = PL(\mathrm{dB}) - GT(\mathrm{dB}) - GR(\mathrm{dB}) + OL(\mathrm{dB})$$

CL can be considered as an abstraction of the wireless link, as includes many values of parameters that are not needed to be known, such as path loss, transmit antenna gain, receive antenna gain, other losses. Instead, it is sufficient to measure the received power at the antenna connector at the receiving node and use the known transmitted electrical power before the antenna connector at the transmitting node for calculating it.

4.3 Management of measurements and metrics

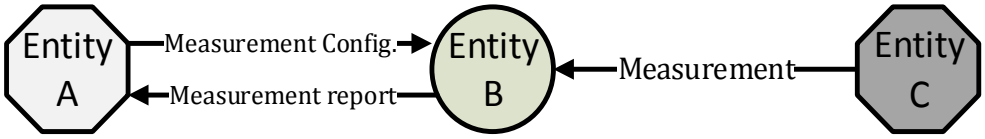
In previous sections, we provide a series of measurement and metrics in order to abstract the network into several different network graphs. However, to well-manage these entities efficiently in the COHERENT C3, each metric and measurement shall be accompanied with the following general entries:

- Type: Status or configuration
- Measured RAT: LTE or WiFi
- Last update time stamp and valid time stamp
  - This information is used to manage the validity time for the metrics, the last update time stamp is marked based on the configuration of this measurement event.
  - For the metrics with unknown valid durations, the valid time stamp is not set since they are updated until some specific events.
- Entity identities in triplet: Identified by the unique identities of the entity (unique identity for LTE UE, LTE eNB, WiFi AP or WiFi client). Based on these identities, the direction of the metrics (Downlink, Uplink, Sidelink) can be formulated.
  - Identity of target measurement entity: The target measurement entity could be the same as the measurement executor entity. Moreover, it could be empty for the non-specific

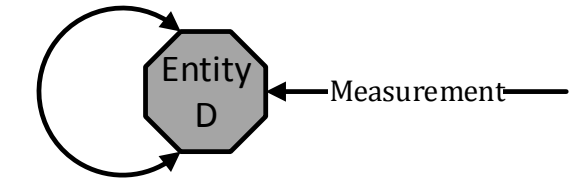
measurement target case, e.g., thermal noise power at LTE eNB, SINR at LTE UE, noise histogram, etc.

- Identity of measurement configurer: The network entity that commands the measurement that is performed can be the same as the executor entity if the measurement is self-configured. Moreover, it can also be the same as target measurement entity, for instance, a LTE eNB can configure the LTE UE to measure the RSRP sent from the same LTE eNB.
- Identity of measurement executor: The entity that does the measurement based on the configuration from the configurer.

In Figure 4-4, the identities of three different entities are shown, the measurement is configured by entity A to entity B in order to measure the status from C to B. In this case, the measured status is maintained with the entity identities triplet as: (Entity A, Entity B, Entity C). As we stated above, entity A, B and C could be the same in partial or in total. In the following we show another example of measuring the thermal noise at the LTE eNB when the measurement target is empty and the measurement configurer and measurement executor are the same due to its self-configured characteristic. Hence the entity identities triplet is: (Entity D, Entity D, -).



a) Three entities in the measurement events



b) Single entity in the measurement event

Figure 4-4 Measurement entities triplet

Besides the common entries stated above, some other system-related and content-related indexes are included in each own content individually. For the specific system-related index, such as the associated release version of the RAT (LTE Rel-8, Rel-9, etc.) and the protocol version of the RAT (IEEE 802.11a, 802.11g, etc.), the layer where the measurement takes effect (Physical layer, MAC layer, RLC layer, PDCP layer, X2/S1 layer, etc.) are included in its content. As for the content-related index, such as the specific sub-identities of the measurement (e.g., Receiver antenna index, Physical resource block index, etc.), link direction of the measurement (e.g., uplink, downlink, sidelink, etc.), measurement duration and measurement periodicity are also incorporated in the content.

5. Evaluation of abstraction opportunities in selected COHERENT control applications

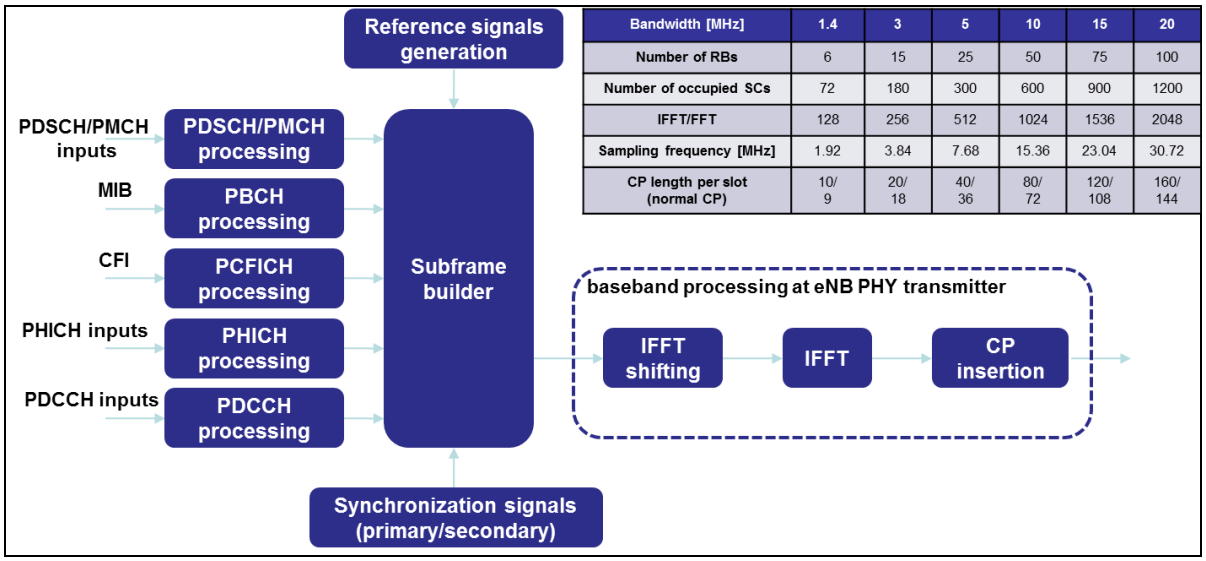
In this section, some interesting case examples, where the network graph concept can be applied, are presented. Specifically, we investigate LTE downlink performance, channel quality indicators for WiFi, distributed and coordinated MIMO systems, inter-node interference modelling, load balancing, cognitive radios, D2D communications, mobility management, and coverage extension. For some of the presented topics, the work is only preliminary and meant to be further investigated in the upcoming deliverables.

5.1 LTE downlink PHY transmission model and link-level performance

The LTE PHY technical specifications are designed to accommodate channel bandwidths from 1.4MHz to 20MHz [185]. OFDM was selected as the modulation scheme for LTE downlink transmissions because of its robustness in the presence of severe multipath fading. Specifically, in this scheme, the high-rate stream of data symbols is first serial-to-parallel converted for modulation onto M parallel subcarriers. This increases the symbol duration per subcarrier by a factor of approximately M, so that it becomes significantly longer than the channel delay spread. Downlink multiplexing is accomplished via OFDMA. As defined in 3GPP LTE Rel. 9 technical specifications, LTE downlink supports 6 physical channels and 2 types of physical signals [186], in which physical channels are used to convey the information from higher layers and physical signals are used for downlink cell search and synchronization, and channel estimation.

5.1.1 LTE downlink PHY transmitter

The diagram of LTE downlink PHY transmitter and the main OFDM modulation parameters are shown in Figure 5-1. The transmitter consists of 6 downlink physical channel processing, 2 types of physical signals, 1 subframe builder, and baseband signal processing.



Bandwidth [MHz]	1.4	3	5	10	15	20
Number of RBs	6	15	25	50	75	100
Number of occupied SCs	72	180	300	600	900	1200
IFFT/FFT	128	256	512	1024	1536	2048
Sampling frequency [MHz]	1.92	3.84	7.68	15.36	23.04	30.72
CP length per slot (normal CP)	10/ 9	20/ 18	40/ 36	80/ 72	120/ 108	160/ 144

Figure 5-1 LTE downlink PHY transmitter

LTE Downlink physical channels include:

- Physical Broadcast Channel, PBCH
- Physical Control Format Indicator Channel, PCFICH
- Physical Hybrid ARQ Indicator Channel, PHICH
- Physical Downlink Control Channel, PDCCH
- Physical Downlink Shared Channel, PDSCH

- Physical Multicast Channel, PMCH

The PBCH carry part of the system information and is required by the terminal in order to access the radio network. The payload to the PBCH processing is called master information block (MIB). The size of MIB on the PBCH is 40 bits. MIB includes the PHICH configuration of the cell, downlink bandwidth, and system frame number (SFN). In addition, the number of transmit antenna ports at eNodeB are implicitly transmitted with PBCH block. The PBCH may use single antenna transmission and transmit diversity with QPSK modulation scheme.

The PCFICH carry information about the number of OFDM symbols used for transmission of PDCCHs in a downlink subframe. The payload to the PCFICH processing is called control format indicator (CFI). The size of CFI on the PCFICH is 2 bits. In order to make transmitted information sufficiently robust, the 2-bit CFI is encoded, scrambled, and QPSK modulated so that it forms 16 symbols. The PCFICH is transmitted on the same set of antenna ports as the PBCH; with transmit diversity being applied if more than one antenna is applied.

The PHICH carry information about the hybrid ARQ ACK/NACK to indicate to the terminal whether a transport block in the UL should be retransmitted or not. HARQ input (HARQ indicator) is set to 0 for a positive ACK and 1 for a negative (NACK). This information is repeated in each of 3 BPSK symbols. Then PHICH is spread through one complex orthogonal Walsh sequence so that 12 BPSK symbols are generated. With different orthogonal sequences, multiple PHICH could be multiplexed in one PHICH group. The PHICH processing may use single antenna transmission or transmit diversity.

The PDCCH carry scheduling assignments and other control information. The payload to PDCCH is so called downlink control information (DCI) format. There are total 11 DCI formats defined in Release 9 and 13 DCI formats defined in LTE-Advanced. The PDCCH is transmitted on an aggregation of one or several consecutive control channel elements (CCEs), where a control channel element corresponds to 9 resource element groups. PDCCH may use single antenna transmission or transmit diversity with QPSK modulation scheme.

The PMCH carries multimedia broadcast and multicast services (MBMS). It is transmitted together with MBSFN reference signal. No transmit diversity is specified for PMCH processing. Layer mapping and precoding shall be done assuming a single antenna port.

The PDSCH carries the data from the eNodeB to User Equipment (UE). It is the PHY channel used for unicast transmission, but also for transmission of paging information and system information blocks (SIB). In addition, the PDSCH can be transmitted on different antenna ports together with UE-specific reference signal in order to achieve the transmit beamforming. Different modulation schemes can be applied for PMCH and PDSCH, such as QPSK, 16QAM, and 64QAM.

LTE Downlink physical signals include:

- Reference Signals
- Synchronization Signals

Five types of downlink reference signals (RS) are defined in LTE-Advanced, namely, cell-specific RS (CRS), MBSFN RS, UE-Specific RS (D-RS), positioning RS (PRS), and channel state information (CSI) RS. The main features of five types of downlink RS are shown in Table 5-1.

Table 5-1 Downlink reference signals

Reference signals	Antenna port	Cyclic Prefix	Subcarrier Spacing
CRS	0, 1, 2, 3	Normal/Extended	15KHz
MBSFN RS	4	Extended	15KHz or 7.5KHz
D-RS	5	Normal/Extended	15KHz
D-RS	7, 8,…,14	Normal/Extended	15KHz

PRS	6	Normal/Extended	15KHz
CSI RS	15, 16,…,22	Normal/Extended	15KHz

Two types of downlink synchronization signals are defined:

- Primary synchronization signal (PSS)
- Secondary synchronization signal (SSS)

The synchronization signals represent LTE physical layer cell identity, given by,

$$N_{\mathrm{ID}}^{\mathrm{cell}} = 3N_{\mathrm{ID}}^{(1)} + N_{\mathrm{ID}}^{(2)}$$

with the cell ID group identifier and the identity within one cell ID group.

Subframe builder is used to construct one LTE downlink subframe. LTE technical specification defines two types of downlink subframe, one is normal downlink subframe and the other is special subframe. Note that special subframe is only used for Frame structure type 2, i.e. for time division duplexing (TDD), and it is used by the UE to switch from reception to transmission or vice versa. Different physical channels have different resource mappings in downlink subframe. Taking normal Cyclic Prefix as example, the PBCH signals are located in OFDM symbols #7, #8, #9, #10 and they are sent over the 72 center subcarriers. The reason for this configuration is that a terminal may not know the downlink cell bandwidth when receiving the PBCH. The concept of Resource Element Group (REG) is defined in LTE specification for control signal mapping. The total number of REGs is calculated based on the channel bandwidth and CFI value. The PCFICH signals are located in OFDM symbol #0 and they are sent over the 4 REGs. The PHICH signal mapping depends on PHICH group and PHICH duration configured by higher layers and the signals are mapped in a various number of REGs. The remaining number of REGs is used for the PDCCH signal mapping. The PMCH signals are mapped in the full channel bandwidth and the PDSCH signals are mapped in the bandwidth scheduled based on different resource allocation types [187]. The physical signal mapping has the higher priority over the physical channel mapping. In other words, no physical signals can be overwritten by physical channel processing outputs. The detailed description of physical signal mapping can be found in [186]. After applying IFFT processing to each OFDM symbol in one downlink subframe, the OFDM baseband signal is generated. In order to avoid Inter-Symbol Interference (ISI) due to the multipath impairment, the CP is inserted at the beginning of each OFDM baseband signal.

5.1.2 LTE downlink PHY receiver

When UE receives the signals sent over the wireless channel, it first performs the time and frequency synchronization and then detects cell identity, CP type and Half Frame (HF) identity over the received signals based on synchronization signals. UE carries out different types of RSRP or RSRQ measurements for the purpose of cell (re-)selection or handover. After removing the CP and applying FFT processing to the synchronized baseband signals, UE starts to decode control information and system or user data. Meanwhile, it performs periodic CSI measurement and report triggered by higher layers or aperiodic CSI measurement and report triggered by DCI format 0 or Random Access Response (RAR) grant. The diagram of LTE downlink PHY receiver is shown in Figure 5-2. Note that all PHY modules included in the UE receiver are processed based on frequency domain data.

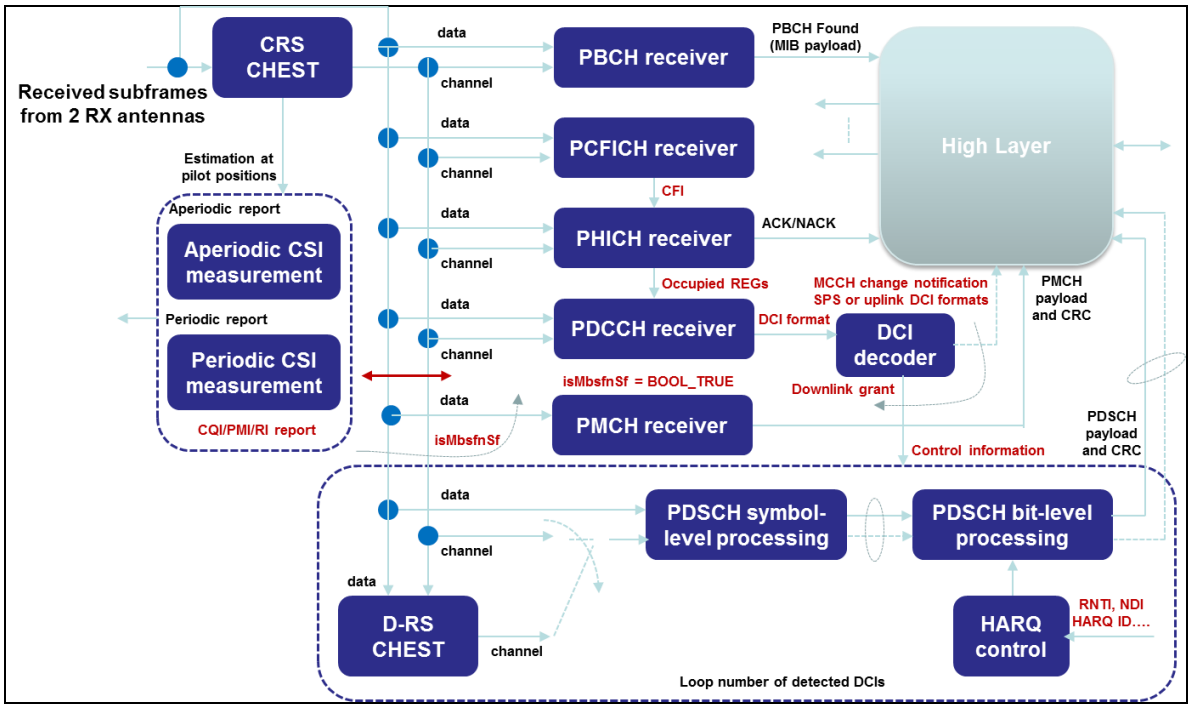


Figure 5-2 LTE downlink PHY receiver

At UE receiver, each subcarrier component is expressed as the product of the transmitted signal and the channel frequency response at the particular subcarrier. Therefore, there is a need to estimate the channel so as to recover the transmitted data. In general, the downlink LTE channel is estimated using reference signals known at both the transmitter and receiver side. Once the channel estimates at RS locations are obtained, the interpolation techniques (for instance, Wiener 2X1D interpolation [188]) can be used to obtain channel response at the non-pilot positions. As shown in Figure 5-3, CRS-based channel estimation (CHEST) is the first module in diagram of UE PHY receiver. It estimates the channels and outputs its estimates to physical channel processing and CSI measurement.

Figure 5-3 shows the control block diagram of CRS-based channel estimation and its parameter estimations. The components shaded in color "blue" indicate processing for CRS-based channel estimation. The three blocks shaded in color "gray" show the processing for the parameter estimation. The *Delay* 'D' is the duration of one subframe after which the appropriate taps are selected based on the delay spread, Doppler spread and SNR estimation [189].

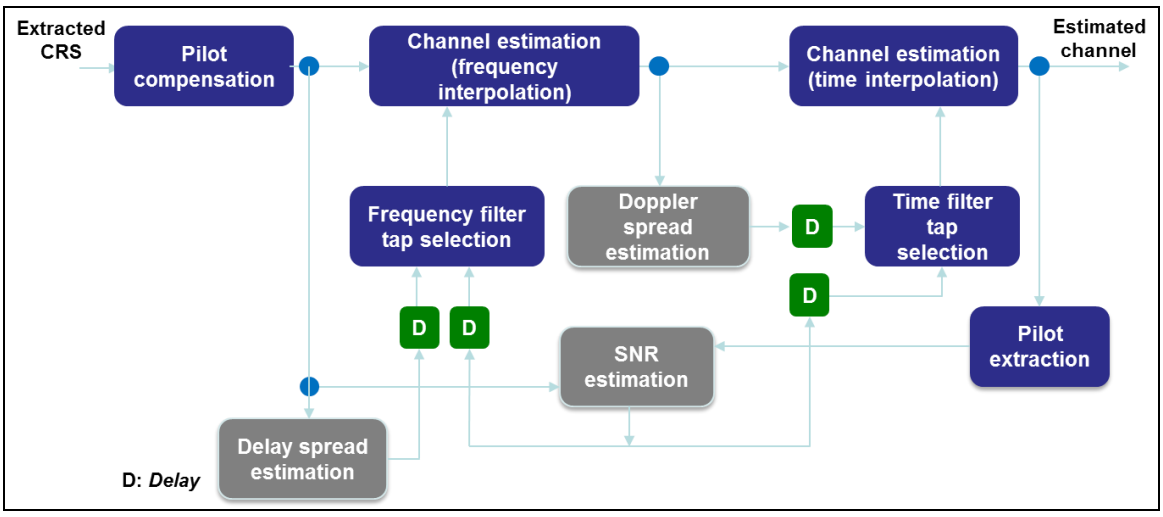


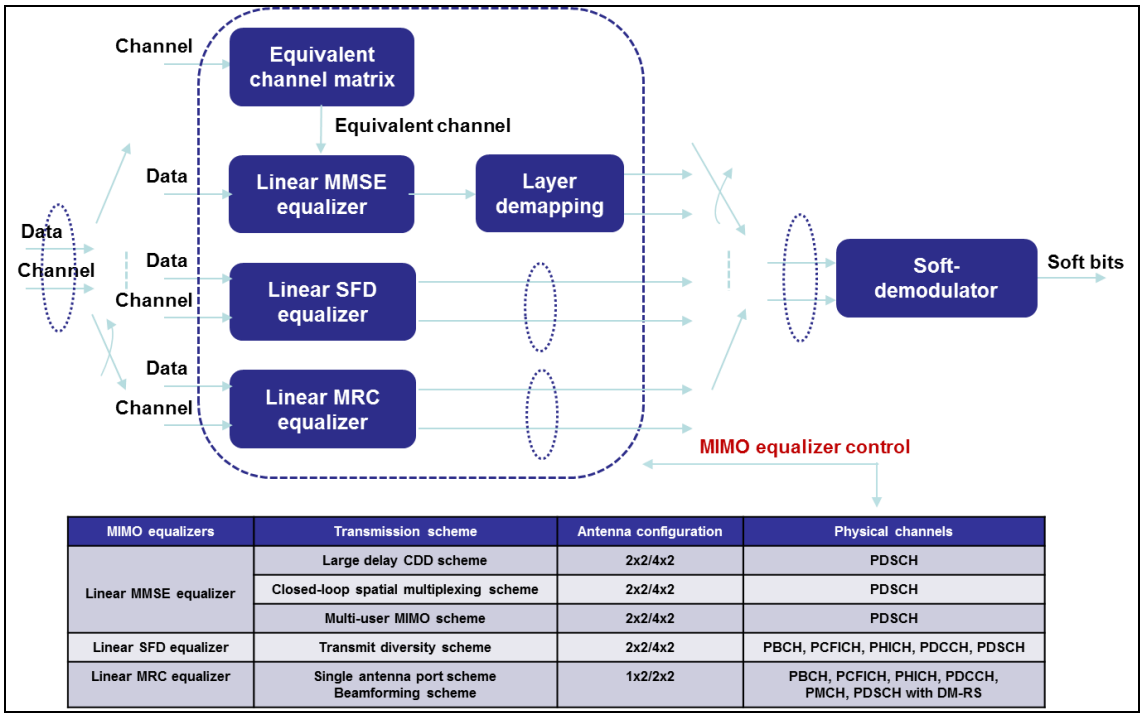
Figure 5-3 Control block diagram of CRS-based channel estimation

The received signals at the location of CRS are extracted and passed to the so-called pilot compensation component. The channel is first estimated at the pilot locations. The compensated channel pilots are fed as inputs into 3 components, namely, delay spread estimation, channel estimation (frequency interpolation), and SNR estimation. The filter taps for frequency interpolation are chosen based on the estimated delay spread and SNR. These are fed into the Frequency filter tap selection component to generate the suitable filter coefficients needed to carry out the channel estimation (interpolation) in frequency direction. The interpolated channel estimates in frequency direction are passed into the Doppler spread estimation component to estimate the Doppler spread. The Doppler spread and SNR estimates are then passed into the Time filter tap selection component to select the appropriate number of taps needed to perform channel estimation in the time direction. The "Estimated channel" at the output of channel estimation component represents the final output for the channel estimation.

Based upon the estimated channel at CRS subcarrier position, periodic CSI measurement or aperiodic CSI measurement is performed in the dedicated downlink subframe and is reported to higher layer. Note that CSI measurement includes estimations of CQI, precoding matrix indicator (PMI), and rank indicator (RI) at UE PHY receiver. According to [187], the UE is requested to report the largest CQI out of 15 possible values corresponding to a specific modulation and coding scheme (MCS) for which the UE can guarantee a transport block error probability that does not exceed 10%. The selection of CQI that is valid for the entire reporting bandwidth requires some kind of averaging of the individual instantaneous subcarrier SINRs. In [190], the exponential effective SINR mapping (EESM) technique is introduced that represents a method for the purpose of computing equivalent wideband SNR estimates in the context of system level simulations. Based on the simple information measure the EESM is defined as

$$SINR_{EESM} = -\beta_{MCS} \cdot ln\left(\frac{1}{N_{sc}} \sum_{k=0}^{N_{sc}-1} e^{-\frac{\gamma(k)}{\beta_{MCS}}}\right)$$

where β_{MCS} is a design parameter that needs to be carefully chosen and N_{sc} denotes the number of subcarriers that are used for the computation of the average.



MIMO equalizers	Transmission scheme	Antenna configuration	Physical channels
Linear MMSE equalizer	Large delay CDD scheme	2x2/4x2	PDSCH
	Closed-loop spatial multiplexing scheme	2x2/4x2	PDSCH
	Multi-user MIMO scheme	2x2/4x2	PDSCH
Linear SFD equalizer	Transmit diversity scheme	2x2/4x2	PBCH, PCFICH, PHICH, PDCCH, PDSCH
Linear MRC equalizer	Single antenna port scheme Beamforming scheme	1x2/2x2	PBCH, PCFICH, PHICH, PDCCH, PMCH, PDSCH with DM-RS

Figure 5-4 MIMO receivers at UE PHY receiver

Physical channel receiver conducts the inverse operation of physical channel processes at transmitter side. PBCH receiver decodes the MIB payload and sends it to higher layer. CFI value can be obtained from PCFICH receiver. PHICH receiver calculates ACK/NACK information and the occupied number of REGs, which are used for PDCCH signal extraction. PDCCH decoders all possible DCI formats by blindly monitoring common search space and UE-specific search space. DCI decoder parses all DCI formats in order to obtain downlink grant associated with control information for PDSCH decode and/or uplink grant for physical uplink shared channel (PUSCH) transmissions. PMCH receiver is called if subframe is configured as MBSFN subframe; otherwise, PDSCH receiver decodes system information, or paging information, and user data based upon the control information. To decode control and PMCH/PDSCH data, different MIMO techniques could be applied according to antenna configuration and transmission schemes. Figure 5-4 shows the diagram of MIMO receivers at UE PHY receiver. Advanced MIMO receivers can be applied for LTE-Advanced in order to achieve a good tradeoff between the link performance and implementation complexity.

5.1.3 LTE PHY link-level simulations

In order to verify LTE downlink PHY link-level performance according to the requirements in [185], the following test bench shown in Figure 3-6 is constructed. The test bench consists of 4 main blocks, namely, test case (TC), downlink PHY transmitter, channel models, and downlink PHY receiver. The final output from this test bench is the throughput for the specific test case at certain defined SNR in [185]. The test case includes 3 sub-blocks, TC setting, eNodeB TC, and UE TC, in which TC setting defines all needed system and subframe configuration parameters. eNodeB TC generates control and data payload to downlink PHY transmitter and the UE TC generates configuration parameters given by radio resource control (RRC) signaling to downlink PHY receiver. The LTE baseband signals generated by downlink PHY transmitter are sent over LTE channels models [185]. After applying channel and additive white Gaussian noise (AWGN) to the input signal, the outputs from LTE channel mode block are sent to downlink PHY receiver. It can decode the data and calculate the CRC per subframe for throughput calculation. For the case of CSI testing, the CSI measurements need to be reported from downlink UE receiver to the transmitter side in order to select proper CQI value, PMI, and RI value.

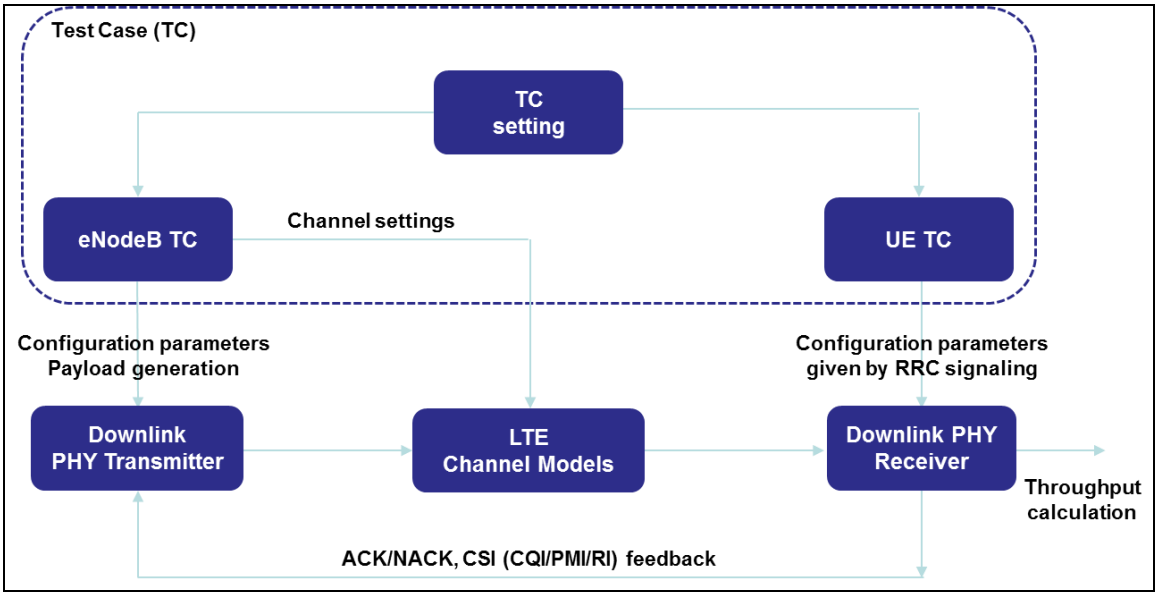


Figure 5-5 Test bench for PHY link-level modelling

LTE technical specification [185] defines multi-path fading propagation conditions and 4 channel models with different Delay profiles and Doppler spreads. The 4 channel models are named as

- Extended Pedestrian A model (EPA)
- Extended Vehicular A model (EVA)

- Extended Typical Urban model (ETU)
- MBSFN channel model

The delay profiles of channel model are selected to be representative of low, medium and high delay spread environments. Figure 3-7 shows the time impulse response of above 4 channel models.

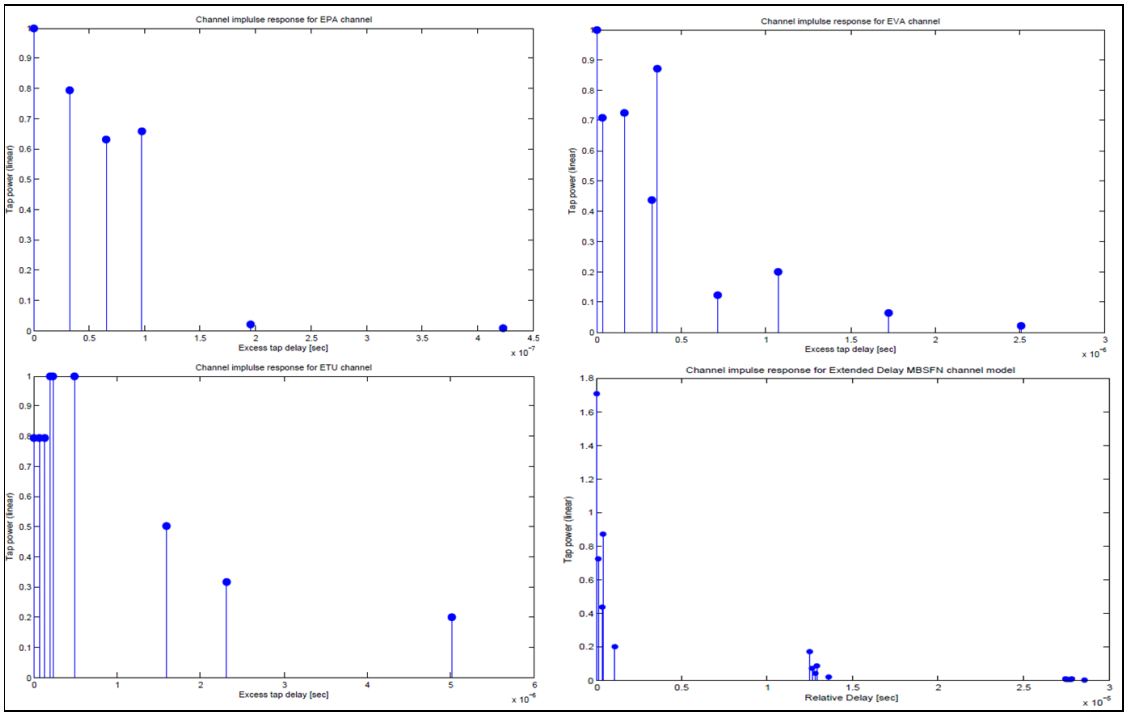


Figure 5-6 Time impulse responses for different channel models

Using CA (CommAgility) test bench and LTE downlink PHY software, link-level simulations are carried out for the test cases defined in [185]. Assuming 30.72MHz sampling frequency, the selected results are given in Table 5-2 (OL denotes Open-Loop and CL denotes Closed-Loop)

Table 5-2 Summary of results

FDD (Frame Structure Type 1)						
Bandwidth	**Modulation**	**Channel Model**	**Transmission Scheme**	**SNR**	**Reference Throughput**	**Result Throughput**
10MHz	QPSK	ETU 300Hz	SIMO 1x2	0.0	70%	90.67%
20MHz	64QAM	EVA 5Hz	SIMO 1x2	17.6	70%	98.89%
10MHz	16QAM	EVA 5Hz	TD 2x2	6.8	70%	91.78%
10MHz	16QAM	EVA 70Hz	OL 2x2	13.0	70%	90.22%
10MHz	16QAM	EVA 70Hz	OL 4x2	14.3	70%	95.61%
10MHz	64QAM	EPA 5Hz	CL 4x2	14.7	70%	91.22%
TDD (Frame Structure Type 2)						
Bandwidth	**Modulation**	**Channel Model**	**Transmission Scheme**	**SNR**	**Reference Throughput**	**Result Throughput**
10MHz	QPSK	ETU 300Hz	SIMO 1x2	-0.2	70%	90.10%
20MHz	64QAM	ETA 5Hz	SIMO 1x2	17.6	70%	98.59%
10MHz	16QAM	EVA 5Hz	TD 2x2	6.8	70%	94.55%
10MHz	16QAM	EVA 70Hz	OL 2x2	13.1	70%	85.66%
10MHz	16QAM	EVA 70Hz	OL 4x2	14.2	70%	92.38%
10MHz	64QAM	EPA 5Hz	CL 4x2	14.7	70%	90.24%

5.2 Modelling of channel quality indicators for WiFi

5.2.1 Problem description

Programmable software-defined infrastructures enable dynamic and adaptive networks with seamless client migration between APs as well as resource and energy efficiency, but require effective representation of the network state. In current WiFi-standards, raw readings or averages of RSSI measurements can be obtained by polling the node entities, but is limiting in terms of increased communication overhead. Besides, raw measurements or averages do not provide reliable estimates of the network state as they do not capture the properties of the underlying distribution. To address the aforementioned limitations, we propose a probabilistic and distributed aggregation model capable of providing a high level abstraction of the link quality in terms of probabilistic representations of the observed network state.

Specifically, we investigate the possibility of building a scalable channel quality map (CQM) that serves as an abstraction of the network graph and represents the estimated channel quality of the links in a WiFi-network. Based on this channel quality map, the aim is to support network management operations such as optimization regarding mobility (e.g. handover) and energy efficiency, in addition to maintaining connectivity and quality of service.

The indoor wireless environment is highly dynamic with mobile stations being subject to varying propagation loss, multipath fading and shadowing. Even links between stationary devices are subject to noise, interference and moving scatterers. The temporal properties of these links are essential to study and model, to get accurate characterizations of links, which can then be used for a range of prediction and management applications.

The approach developed is centered on WiFi technology and use case 1.SR - RAT Sharing Support described in D2.1 [15] - the aggregation and modelling principles described here are, however, relevant for more than one use case. In order to support distributed monitoring the COHERENT architecture need to fulfill requirements #28 [15].

5.2.2 Proposed solution

We aim at building a network graph representing link quality between wireless and programmable node entities in a network. Our proposed solution to this is centred on a distributed and probabilistic approach for building a network graph or channel quality map (CQM) by modelling link quality indicators.

Characterizing a link requires the definition of a metric for link quality which could either be directly measured or estimated from alternate indicators. Throughput or goodput are the desired metrics that quantify link quality for most applications. A link quality estimator is thus gauged by its ability to accurately estimate the time varying throughput on the link. Received signal strength indicator (RSSI), signal to noise ratio (SNR), bit error rate (BER), and packet delivery ratio (PDR) are common indicators used in prior research for acquiring robust localizations of wireless nodes and for modelling link performance [141][142][143][144][145]. Each indicator has its advantages and limitations. We have begun our study of probabilistic characterization of temporal variations in WLAN links by using RSSI in combination with PDR to create signal strength maps. Previous research has shown that RSSI alone is insufficient to characterize link quality since it only measured over preamble of a frame and also does not capture the presence of interference [146][147][148]. We hence use it as a starting point from which we will extend the estimation metric to a tuple of RSSI and PDR, which will extend the signal strength map (SSM) to a more complete and accurate CQM.

In the COHERENT framework, a wireless and programmable node entity has the capability to deploy one or several software-defined access point (AP) controller applications. The CQM represents the observed link quality between clients (STA) and programmable node entities (APs) and is aimed to support management decisions. Supported management operations are typically related to client migration or (de-)instantiation of AP-controller applications under varying conditions, such as when the user concentration increases or decreases at a certain location.

We assume that we only have control over programmable node entities and that clients are not controllable. The distributed aggregation model consists of local monitoring components deployed in the programmable node entities. A local monitoring component aggregates raw RSSI and packet delivery measurements for the purpose of computationally light-weight analysis for a given time-scale.

We have so far investigated real-world measurements of RSSI values and packet delivery rates from a WiFi-testbed consisting of programmable node entities (i.e. APs) and clients such as laptops and smartphones. The data are collected from controllable node entities which sniff transmitted frames from AP-STA and STA-AP. Upon decoding the received frames, we obtain the information of transmitter MAC-address, RSSI at the received AP and sequence number of the frame, which are used as input to our distributed aggregation model. Hence, each programmable node entity maintains a local CQM representing the locally perceived network state, which can be used by various management functions in the COHERENT architecture.

5.2.2.1 Algorithm description

Preliminary investigations of the data indicate that probabilistic aggregation can be performed by modelling a random RSSI having a known distribution and then estimating its parameters using sample moments. We are currently investigating the feasibility of the following approach using data from real-world deployments.

In each programmable node entity, for each link it observes, counters ${c_i}^k$ (e.g. sum) of the observed RSSI and additionally also the PDR, are measured over a pre-defined observation interval indexed by i and link indexed by k. The length of the observation interval should be appropriate for the rate at which the link is varying and the latency requirement of the application at the controller. Our requirements and observations indicate that a time scale of seconds or milliseconds is appropriate depending on the network traffic conditions. In our initial experiments on the data at hand the time scale varies between 1-10 s. Given the updated counters, the aim is to estimate the model p_i for RSSI and PDR, respectively, for the purpose of expressing the link performance in terms of a joint probability model ξ_i based on the percentiles derived from the cumulative distribution functions (CDF) F_{RSSI} and F_{PDR}. In general, the CDF provides the probability of making an observation below a certain value, $P(X \leq x)$. Hence, it is possible to assess the probability (or risk) to observe RSSI and PDR indications below tolerable levels. Given the complementary information properties of RSSI and PDR, we believe that the link performance can be expressed as a joint probabilistic model ξ_i based on the product of both percentiles - see Section 5.2.2.2 for further details.

For both RSSI and PDR the following algorithm per estimator is carried out:

1. Upon reception of a frame from a transmitter within observation interval i: update counters with the new observation ${x_i}^k$ and update the number of samples $n_i = n_i + 1$
2. At the end of time interval i, update and store model parameters p_i based on maintained counters; reset counters for next observation interval $i + 1$.
3. Compute percentiles $F_{RSSI}(p_i^{RSSI})$ and $F_{PDR}(p_i^{PDR})$, respectively.
4. Compute ξ_i based on F_{RSSI} and F_{PDR}

The above outlines a generic framework in which specific parameter estimation approaches can be plugged depending on the properties of the data and the computational capabilities of the node entities.

5.2.2.2 Network graphs and abstraction aspects

5.2.2.2.1 Measurements

Vlavianos et al. [141] do a detailed study of the appropriate metric to use for link quality and describe the conditions under which each is useful. They examine the correlation of each metric to the transmission rate used in 802.11a and 802.11g WiFi measurements. They demonstrate that RSSI, while being the easiest metric to measure, does not capture the effects of interference in the network since

RSSI is available only for successfully received frames. The PDR, although a good estimate is specific to the size of the frame and rate of transmission, both of which vary in the network. Recognizing the strengths and limitations of each metric our idea is to combine them in an efficient manner for an accurate CQM.

The measurements currently investigated for representing a CQM for WiFi networks are RSSI and PDR (based on monitoring the frame sequence number per link) which are observed at the MAC layer. Within each observation interval i the measurements are stored upon decoding of received frames. The approach can be used within the current standards for wireless technologies. The data under investigation have been collected in a testbed with 802.11g programmable node entities. Measurements and updates of parameter estimates after each observation interval i is carried out locally in each programmable node entity.

5.2.2.2.2 Abstraction metrics

The probabilistic monitoring function will provide input to the CQM (or rather SSM) in terms of parameter estimates that represent the learned model in a compact form. The possible output includes:

- Joint probability model metric ξ_i
- Descriptive statistics, such as sample mean and variance of the metrics used (RSSI and PDR in our case)
- Reports on performance changes (e.g. degradations)

In general, estimating the parameters of known distributions enables more advanced analysis of the link performance compared to keeping track of the average or an exponentially weighted moving average (EWMA). Once the estimated models have been acquired for both the RSSI and PDR, it is possible to 1) derive descriptive statistics, such as sample mean and variance; and 2) assess the link performance by studying the (empirical) CDF or probability mass function (PMF) at various percentiles.

In the case of the RSSI model, the second option can be used to create a probabilistic threshold or metric that reflects the tolerable risk φ_{RSSI} of having a signal strength that is persistently below a certain RSSI value. The properties of the (empirical) CDF allow for creating a signal strength degradation detector - given an RSSI-value r (e.g. $r = -85$ dBm), the $F_{RSSI}(r; p_i)$ and $0 < \varphi_{RSSI} < 1$ (e.g. $\varphi_{RSSI} = 0.1$), the risk of having signal strength degradation can be tested relative to the condition $\varphi_{RSSI} < F_{RSSI}(r; p_i)$.

In the case of the PDR model, the probability mass corresponds to the $F_{PDR}(p_i)$, representing the probability of successfully receiving a packet. In a similar way as in the RSSI case, a threshold on the tolerable packet loss on a link can be defined as $0 < \varphi_{PDR} < 1$, with a corresponding detector defined as $\varphi_{PDR} < 1 - F_{PDR}$.

Neither the RSSI nor the PDR models provide sufficient information about the link quality alone, but can be combined to a single probabilistic metric based on the joint probability ξdefined as:

$$\xi_i = (1 - F_{RSSI}(r; p_i))F_{PDR}(p_i),$$

where the total probability of observing RSSIs above a certain threshold and a certain packet delivery ratio is computed. Here, $\xi_i = 1$ corresponds to a perfect link. The combination allows for detecting when the overall link performance is degrading. By studying the individual RSSI and PDR percentiles, the main contribution to the link performance degradation can be identified. For example, a low ξ_i may be caused by a weak RSSI (and thereby also low PDR) indicating failure or misconfigured equipment, or just a link with low signal. In other cases, a low ξ_i may be influenced mainly by low PDR due to e.g. interference or congestion while the RSSI remains high.

Estimation approach: In the ideal case, the RSSI and PDR can be modelled by estimating parametric distributions given that the data can be fitted to known density functions. In general, parametric

modelling is more information effective compared to non-parametric models, in the sense of a better capture of the underlying statistical behaviour which can be used for optimization and prediction purposes. Empirical observations of the datasets indicate that RSSI measurements in each observation interval could be modelled in terms of normal distribution (in combination with adequate continuity corrections since the data points are discrete) for larger time scales (e.g. 20s), but further analysis is required to confirm this, specifically for smaller time scales at which the function will operate. Moreover, observations reported in the literature do not provide conclusive results regarding the RSSI distribution [149]. Additionally, both RSSI and PDR are discrete valued which further limits the use of continuous parametric distributions. Hence, we will employ a non-parametric modelling approach that does not require assumptions about the underlying generating process of the measurements and that can be applied to discrete values, in order to estimate the percentiles directly.

To compute the percentile for the RSSI measurements at a given tolerance RSSI level *r*, we estimate the RSSI percentile $P(X \leq r) = F_{RSSI}(r) = \sum_i \zeta_i / n$ as the ratio of observations less than r and where ζ_i is an indicator variable and n is the number of samples in a time window:

$$\zeta_i = \begin{cases} 1, & x \leq r \\ 0, & x > r \end{cases}$$

Further, the PDR is per definition the ratio between the number of received and sent packets [141], meaning that we estimate $P(X = 1) = F_{PDR} = \sum x / n$ where x is derived from the difference between observed packet sequence numbers.

The number of data points in each time window may vary from very few to many which influence the estimates in each interval. Moreover, the observed links are relatively noisy which leads to large variations in the measurements. In order to obtain "smoother" estimates that also provide more reliable results when the number of samples is small in a time window, we estimate the percentiles by the use of Bayesian estimators such that a current estimate is a weighted combination of the current and prior parameter estimates. We assume that the observations for ζ for RSSI and x for PDR, respectively, follow a Bernoulli distribution with parameter F_{RSSI} and F_{PDR}, respectively, and that the Bernoulli parameters F_{RSSI} and F_{PDR} also follow a distribution (a prior). Based on using the Beta distribution as a prior over the Bernoulli parameters, the Bayesian percentile estimates in both cases correspond to:

$$E[F_i] = \frac{\nu_{i-1} F_{i-1} + n\bar{x}}{\nu_{i-1} + n}$$

where is υ can be considered as weight influencing the importance of recent data - over time the effect is that older estimates decay exponentially, similarly to an EWMA. The $E[F_i]$ can then be plugged into the expression for the link performance model ξ_i.

5.2.2.2.3 Implementation of the network graph

Our approach emphasizes distributed methods. The parameters for link performance (ξ_i) are estimated locally at each programmable node entity and communicated to the controller as required by the application either on request or based on triggers. Given the decentralized control architecture and the monitoring components as described in D2.2 [140] we see that the monitoring approach can be split into three components operating at different levels of the architecture in order to achieve resource-efficiency and timeliness.

As shown in Figure 5-7, the parameter estimation based on the RSSI and PDR measurement counters is carried out locally in each node entity (or station). The obtained parameter estimates can then be used for deriving probabilistic link quality metrics and for detection of changes in the link quality. If the computational capacity of the node entity allows it, the processing of the parameter estimates can take place directly on the node and reported upon request or when a change (e.g. degradation) has been detected, meaning that a node-local CQM is maintained. Alternatively, the parameter estimates can be pushed to a dedicated node entity handling the CQM for further processing at regular intervals. The CQM is in most cases assumed to be maintained by relevant functions residing in a controller. The

entries a CQM hold the estimates and the probabilistic metric of the observed RSSI and PDR per receiver-transmitter pair (identified by the MAC addresses). Additional descriptive statistics such as moments, sample mean and variance may be stored as well. Functionality for maintaining the CQM and using the information for various management and controller actions (e.g. migration of a station to another access point) may include additional processing for computing the probabilistic metrics in the case this is not done locally in the nodes. Finally, CQMs may be merged for representing the link quality over several regions in terms of a centralized network view [140].

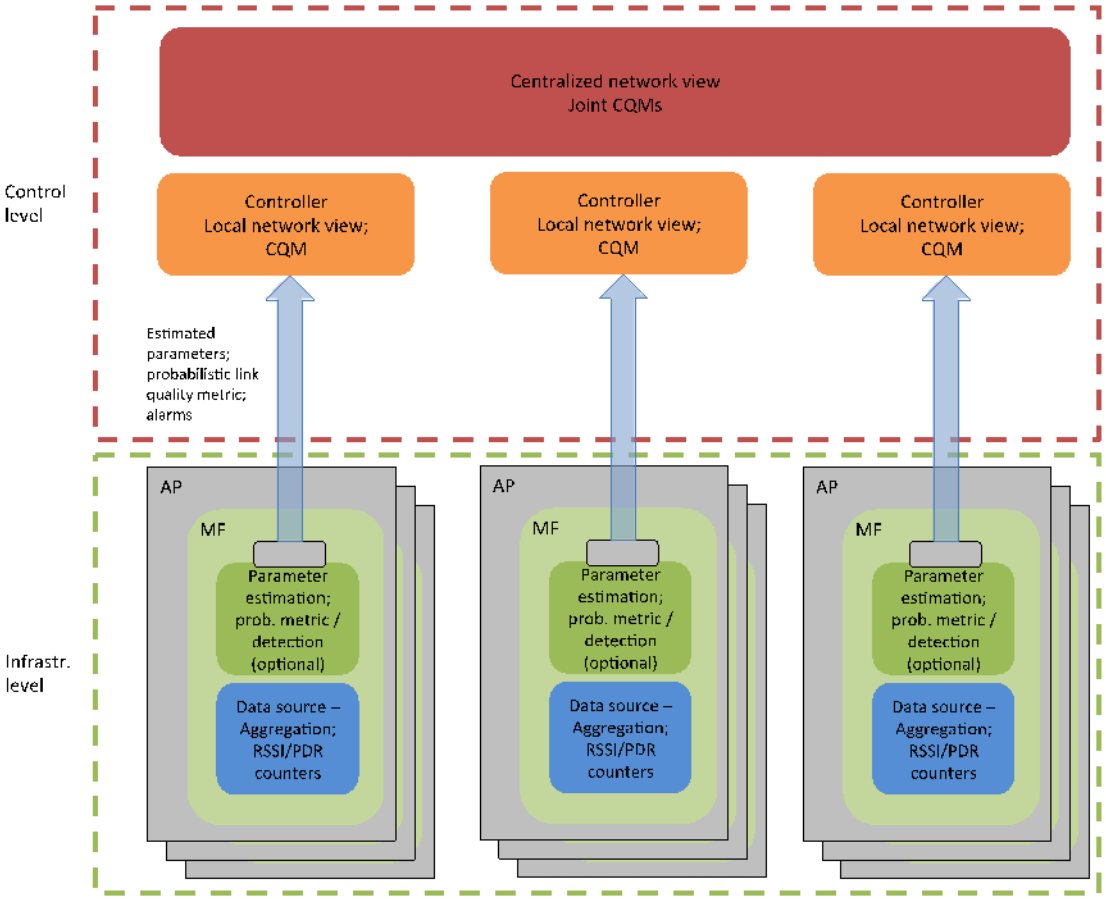


Figure 5-7 A conceptual overview of monitoring functions (MFs) implementing local aggregation and probabilistic modelling inside APs

5.2.3 Evaluation

5.2.3.1 Evaluation methodology

Our evaluation so far is based on qualitative analysis of real world data collected from an enterprise WiFi network. We have defined a metric ξ_i that incorporates quantile RSSI obtained by using a threshold on the RSSI CDF, and PDR obtained from sequence numbers. The metric ξ_i is a probabilistic estimate of the link quality with $\xi_i = 1$ indicating a perfect link. A threshold on ξ_i is used to identify link degradation.

5.2.3.2 Results

Figure 5-8, Figure 5-9, and Figure 5-10 depict, respectively, the plots of Bayesian estimated time series ξ, $(1 - F_{RSSI})$ and F_{PDR}. The quantile RSSI plot is obtained by choosing an RSSI threshold r that, given a data rate and no interference, results in high PDR on the link. The threshold value is based on the theoretical relationship between SNR and BER extended to a relationship between RSSI and PDR for a known frame size and data rate. This RSSI threshold does not account for interference which is

motivation for including sample PDR to capture interference. Our demonstrative choice of -86 dBm should be adjusted for each WiFi supported data rate. Looking at Figure 5-9 we see that the quantile RSSI drops low a little before 2000 s in the time series. This degradation in link is accompanied by a drop in the PDR metric as well. Looking at Figure 5-10 we see another degradation in the link after around 3000 s. Notice that this drop in PDR is not accompanied by a drop in the quantile RSSI metric. This is likely a scenario where the RSSI of the link is high, but interference causes PDR to drop. The combined metric ξ reacts to both degradation in RSSI and in PDR and provides a single metric that can be used to monitor link quality. A suggested threshold probability for a good link (good being defined by the combined metric) is indicated by the red horizontal line on the plot. This can be used to trigger management or control actions when the link degrades below this threshold.

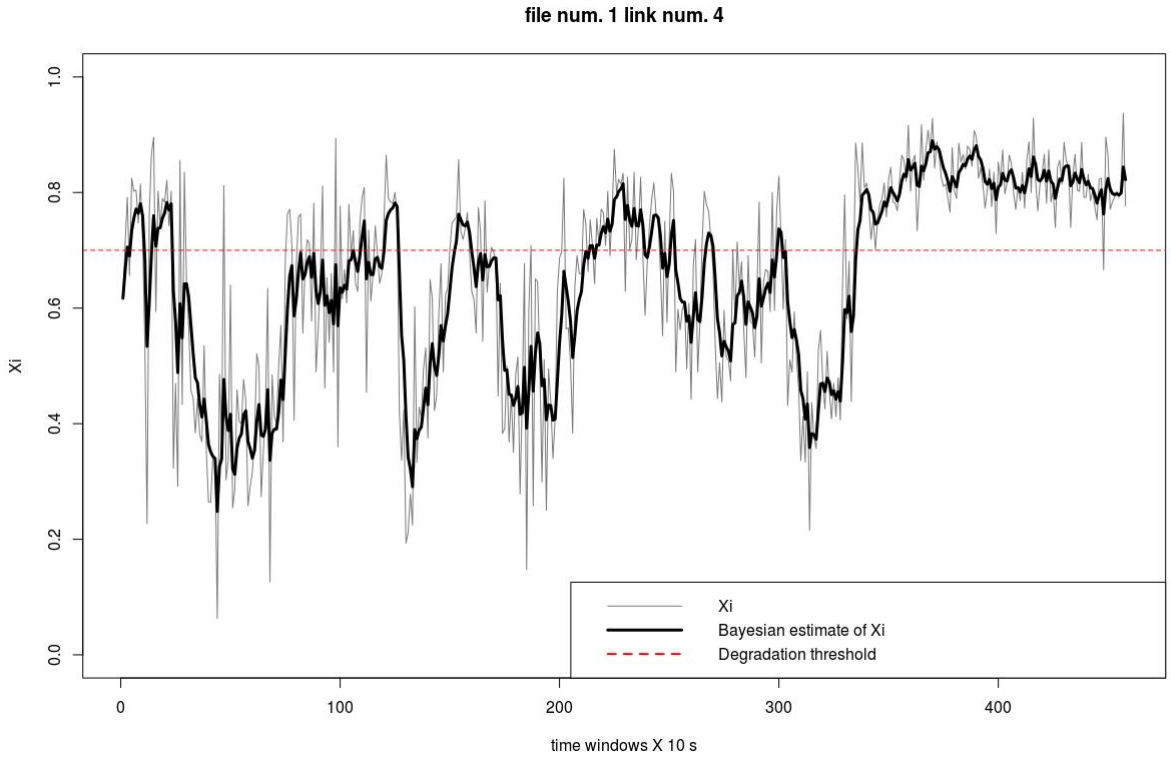


Figure 5-8 Plot of ξ_i as a time series over a sample link, with window size of 10 s. The black overlay curve on each plot is the smoothed Bayesian estimation of the metric in the background.

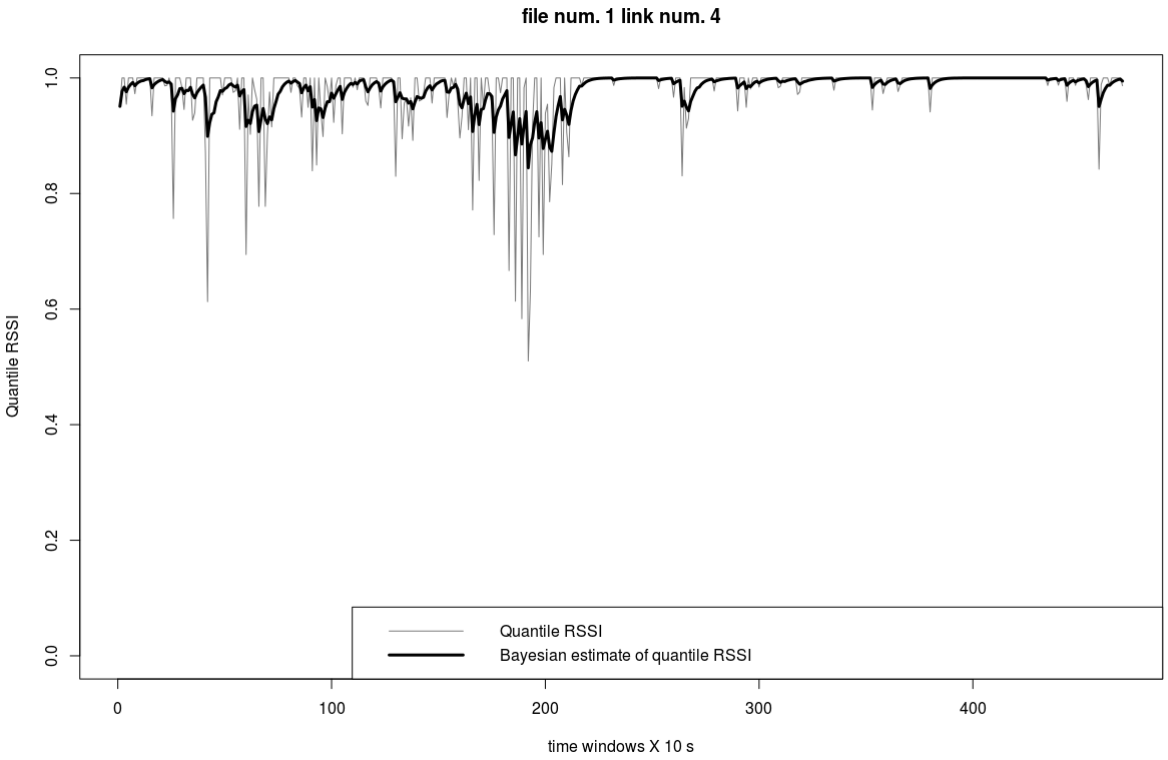


Figure 5-9 Plot of quantile RSSI as a time series over a sample link, with window size of 10 s. The black overlay curve on each plot is the smoothed Bayesian estimation of the metric in the background.

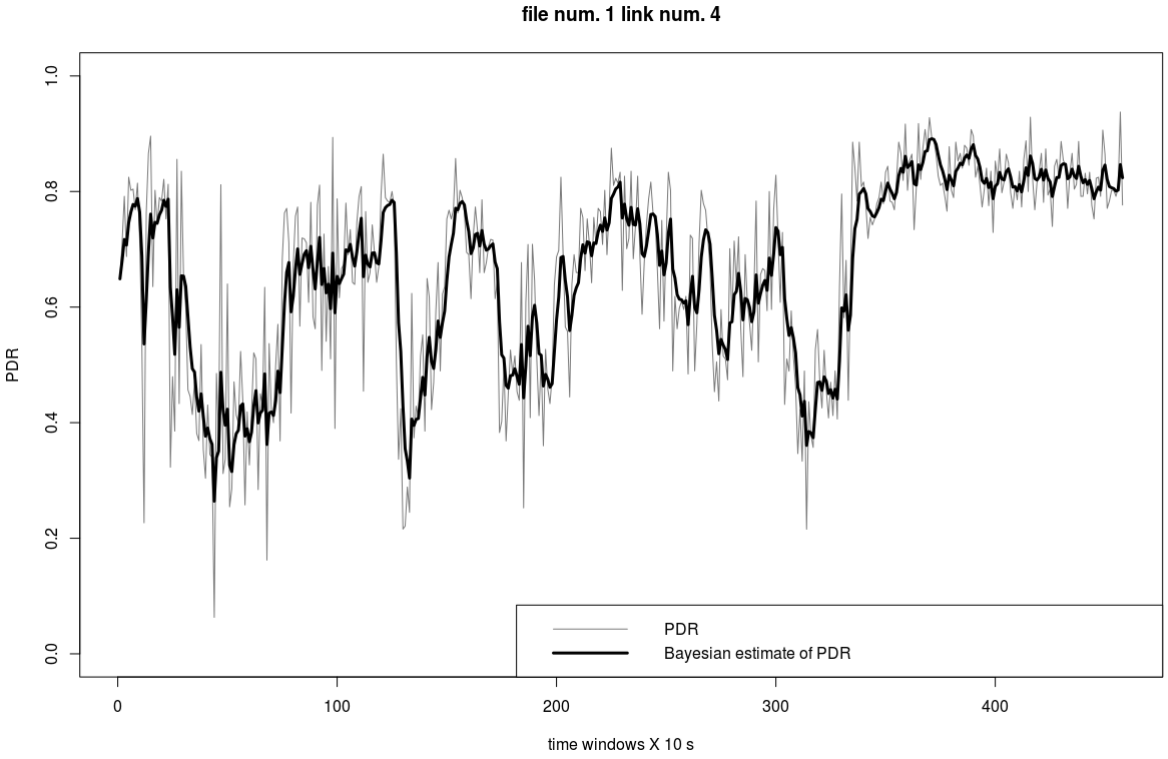


Figure 5-10 Plot of F_{PDR} as a time series over a sample link, with window size of 10 s. The black overlay curve on each plot is the smoothed Bayesian estimation of the metric in the background.

5.2.4 Final discussion and summary

The work throughout the COHERENT project is focused on probabilistic modelling and encompasses to understand: the statistical properties of various channel quality indicators obtained from a WiFi testbed as well as identification and analysis of other available metrics that can be used to represent the network state in a CQM; how to observe and aggregate the measurements in a resource-efficient and scalable manner; architectural aspects of such a distributed modelling mechanism; and, evaluation of the proposed mechanism.

Next, we will continue to investigate the performance of the outlined algorithm and obtain quantitative analysis based on further measurements in the WiFi-testbed in combination with developing an adaptive measurement mechanism. We see that an adaptive measurement mechanism is necessary for autonomously performing active measurement with varying intensity when the traffic on a link under observation is too sparse per time interval to provide a reliable channel quality estimate. Evaluation results are planned to be reported in D3.2.

We also plan to improve our current link quality indicator ξ, by extending PDR with bit error rate (BER) or other equivalent metrics. PDR is dependent on the size of frames and rate of transmission over the link. BER would be a more generic metric that can be used to estimate throughput on a link for a known frame size and rate.

5.3 Distributed antenna system/massive MIMO

5.3.1 Problem description

With decreasing cell sizes at high-density traffic hotspots, the number of cell-edge areas increases dramatically. As the users move throughout hot-spot areas covered by multiple base stations, they will be always located at cell-edge at some moments and experience performance degradation due to high interference. Mobile operators are looking for efficient densification techniques while the users are looking for consistent service in hot spot areas.

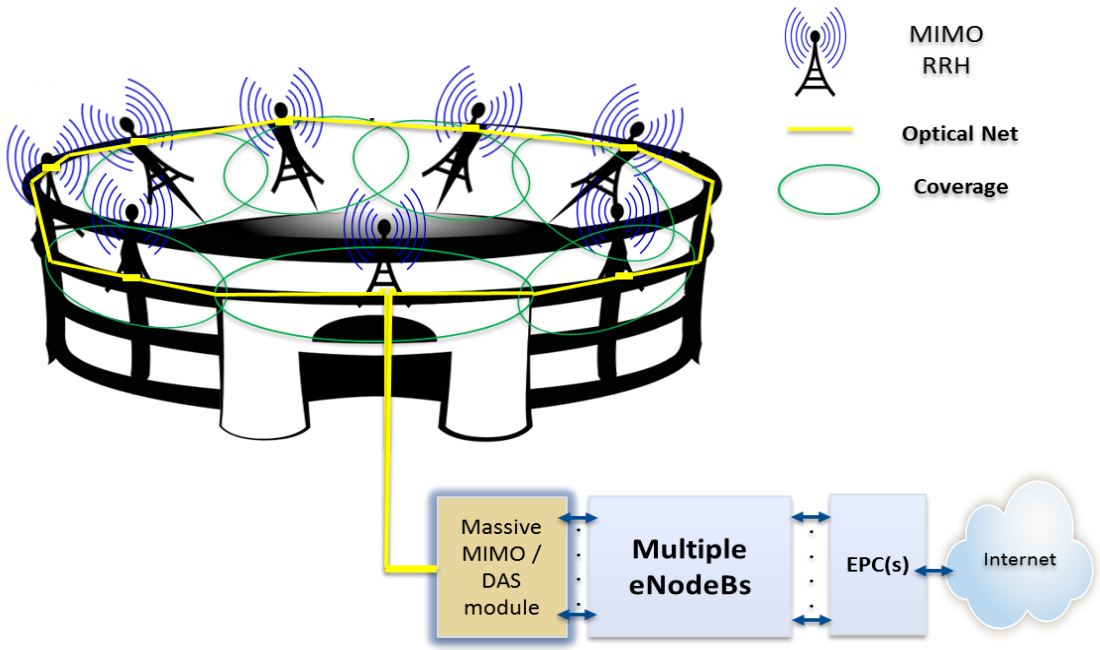


Figure 5-11 Possible DAS setup in a stadium

5.3.2 Modelling

The problem is addressed in the project by using Massive or very-large MIMO eventually combined with distributed antenna system or base station coordination (e.g. CoMP and interference alignment techniques). Massive MIMO contributes to interference elimination and thus will provide a significant

improvement to cell-edge users' throughput, contributing to their consistent mobile experience. For mobile operators, it will allow a more intense reuse of the spectrum in dense areas and so increasing cell spectral efficiency and improving cell coverage. Furthermore, the combination of small cells with Massive MIMO presents new opportunities to reduce energy consumption and associated energy costs. Massive MIMO is considered among the key enablers to achieve the challenging 5G requirements.

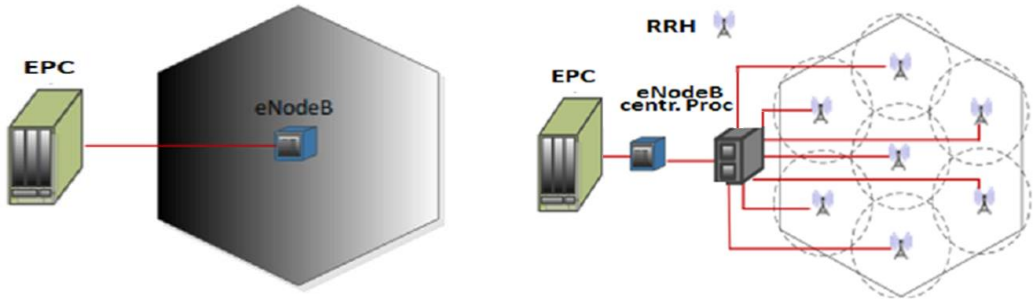


Figure 5-12 Collocated antenna cell and distributed antenna cell

This potential technology will be investigated, modelled theoretically and simulated for the possible distributed antenna system (DAS) scenarios. The necessary parameters for the network graph w.r.t. DAS will be abstracted as well. The description below is a summarization of the work accomplished so far, which comprises a detailed literature review and initial findings on modelling of DAS.

Figure 5-11 describes a possible DAS setup in a stadium, where the density of users is quite large. In this setup the coverage of each remote radio head (RRH) is made quite small by transmitting very small amount of power, thereby eliminating interference, to accommodate many RRHs for serving the high density of users. The DAS can be setup in different ways like cooperative RRHs and dynamic clustering.

DAS promises to deliver viable solutions for the future wireless challenges using multiple RRHs in a distributed architecture, thereby supporting rapid frequency re-use, decreasing the interference and increased coverage. DAS architecture consists of Multiple RRH's connected to a centralized base station via an optical fiber network. By combining multiple base stations (BSs) into a server we increase demand for resource control. It has to cope not only with processing latency in servers but also with latency and throughput on fronthaul interface [95]. A fronthaul link, an interface between RRH and base band unit (BBU), is a new element that does not exist in a traditional base station. The fronthaul connection is usually served by common public radio interface (CPRI) or open base station architecture initiative (OBSAI) protocol. The fronthaul link capacity requirement highly depends on the functional split between RRH and BBU. For example, it requests approximate 1 Gbps fronthaul capacity in order to fully transport all air-interface raw data samples for a single UE with 75 Mbps peak rate over the uplink. This gap between the air-interface user throughput and fronthaul capacity can be highly reduced if more functions are employed at the RRH side; for example, applying the FFT operation at the RRH side can almost half the data rate. Besides the challenge of the capacity, the timing constraint due to LTE HARQ processing is stringent when we consider the fronthaul link. That is because we also take the latency due to the round-trip fronthaul transmission delay into account. A suggested value of the one-way fronthaul transmission latency shall be less than 250 μs (NGMN, 2015), so in the worst case, the available processing time is reduced by 500 μs due to the round-trip transportation between RRH and BBU.

In DAS setting, a single RAN engine simultaneously operates multiple RRHs. This is similar to classical MIMO system except that geographically distributed RRHs (Figure 5-12) give an additional degree of freedom, usually referred to as *macro diversity*. In DAS setting we are interested whether to use linear diversity pre-coding or a simple transmit antenna selection (TAS). A linear pre-coding is a fashionable way to tap into diversity. However, in a DAS the antennas are located far away from each other. User has significant path loss difference towards different antennas and it is preferable to direct all the transmission power to the strongest RRH only. As a natural extension of TAS, we also consider space-

division-multiplexing (SDM) where as many users as RRHs are simultaneously scheduled. In SDM, each RRH serves one user. In a typical DAS, the RRHs have minimal intelligence of their own; essentially they are an extension of the base station antenna ports.

Previous work with single antenna RRHs has established the benefits of DAS for improving indoor coverage [96, 97], reducing outage in the downlink [98, 99], reducing outage in the uplink [100], and increasing capacity [101]. DAS works with different multiple access technologies [102, 103]. The application of multiple user MIMO to DAS has received less attention. Simulation results in [104] for several different transmission techniques, incorporating realistic channel parameters and interference, show that orthogonal transmission between RRUs is preferred in some cases. Results in [105] show that multiuser MIMO DAS outperforms selecting the best user in the presence of out-of-cell interference. While the results in [106] consider scheduling, only three single antenna RRHs are considered and expressions for ergodic rates are not provided. The RRHs may have a single antenna or multiple antennas while the subscriber has a single receive antenna.

The current work mainly focuses on the details of implementation of LTE PHY on a TI (Texas instruments) DSP and the corresponding processing blocks that need to be adapted to support the functionality of DAS. Effort has been made to implement a commercial LTE system on DSP and ARM boards. The LTE PHY layer is ported on a commercial TI DSP which gives flexibility to modify or enhance the functionalities towards DAS.

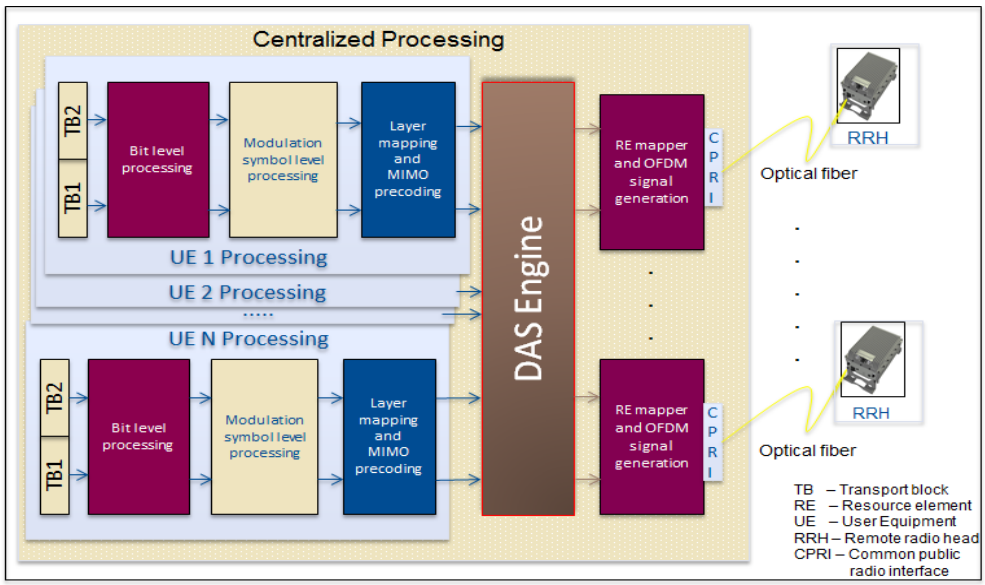


Figure 5-13 DLSCH transmitter processing chain adaptation for DAS [107]

Figure 5-13 represents a PHY level, DAS architecture with centralized processing unit, where the PHY level processing for all the UEs are done, replacing the ubiquitous base stations or cellular architecture, connected to many RRHs'. In [107] without loss of generality LTE downlink shared channel processing for transmission has been taken as a use case to analyze realistically the necessary adaptations required for DLSCH processing blocks to incorporate it in the DAS. This analysis is quite valid and can easily be extended for all downlink and uplink channels. For the LTE downlink shared channel processing chain, a case with 2 TBs per UE per subframe is assumed. After all the TBs received by the PHY, it starts per user and per TB related processing at bit level like CRC, code block segmentation, Turbo coding, rate matching and code block concatenation, it is forwarded to modulation symbol level processing which includes scrambling with a pseudo random sequence and modulation mapper. The generated modulation symbols for that user, i.e., the symbols from all the TBs pertaining to that UE are passed through layer mapping and MIMO precoding blocks. In the LTE specification, after the MIMO precoding the symbols are mapped into the allocated resource elements depending on the resource allocation type and number of RBs and the OFDM signal is generated. But in the DAS, the outputs of the MIMO precoder for all the UEs allocated in that subframe could be further processed by DAS

Engine block. It is obvious that one can also incorporate the MIMO precoding block into the DAS Engine itself, but to make clear the adaptations w.r.t. the existing LTE standard the MIMO precoder and DAS precoder are shown separately.

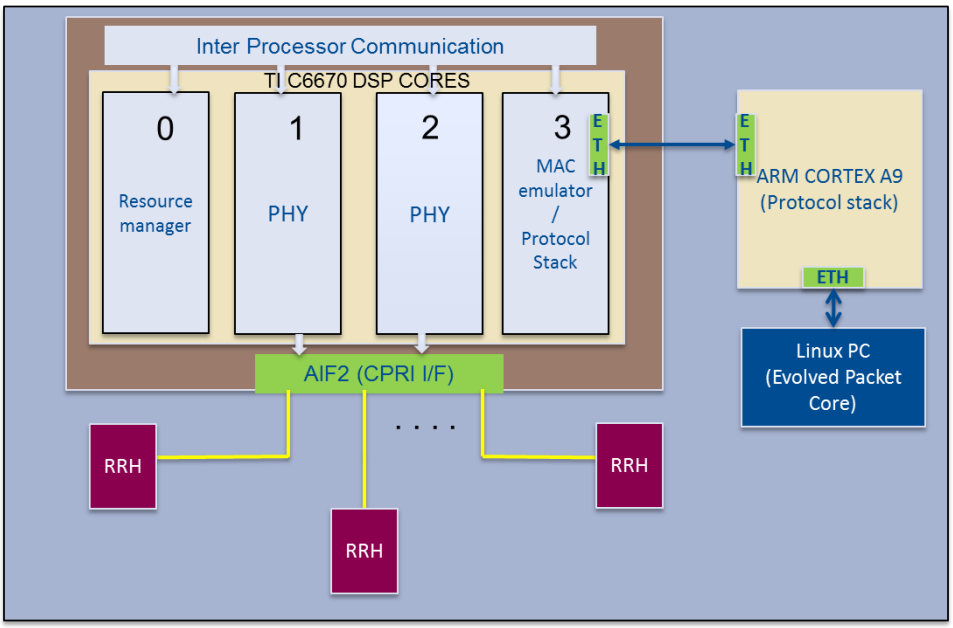


Figure 5-14 Hardware implementation blocks [107]

As described in Figure 5-14, the PHY layer processing blocks were ported onto the DSP, which communicate with Protocol stack, running on an ARM board via Gigabit Ethernet. The Antenna Interface component on the DSP receives the IQ samples from the DSP and converts them into CPRI frames that facilitate the transfer of baseband signal to the RRHs via optical cables. Protocol stack will communicate with EPC, running on a Linux PC. Although DSP is capable of processing for many users, without loss of generality, as a first step we have selected a use case with only 2 UEs per subframe with 2 RRHs. Implementation of the processing blocks on a DSP gives us high level of flexibility to analyze, modify, append new features and measurements. Real-time tracing and debugging on CCS, a TI proprietary debugger provides immense opportunities to fine tune or even select the correct set of precoding vectors in real-time. The current LTE setup can communicate with commercial LTE UEs. Depending on the precoding techniques used within the DAS precoder, a commercial UE can be used to communicate if no further processing on the UE is required. In case of further processing on the UE is required, we employ a Keysight MXA LTE analyzer to capture IQ samples, which can be decoded offline using MATLAB.

5.3.3 Measurement setup

In Figure 5-15, on the left side the LTE eNodeB testbed in a microTCA chassis is depicted and on the right side Keysight's MXA LTE signal analyzer is used for measuring the quality of the signal received. The microTCA chassis contains the TI DSP board and RRH board apart from the microTCA controller card. The PHY layer processing blocks are ported onto the DSP. The DSP communicates with protocol stack, which is running on an ARM board, via Gigabit Ethernet. The antenna interface (AIF2), a hardware accelerator on the DSP receives the time-domain IQ samples from the DSP and converts them into CPRI frames that facilitate the transfer of baseband signal to the RRHs via optical cables. Protocol stack communicates with evolved packet core (EPC), running on a Linux PC.

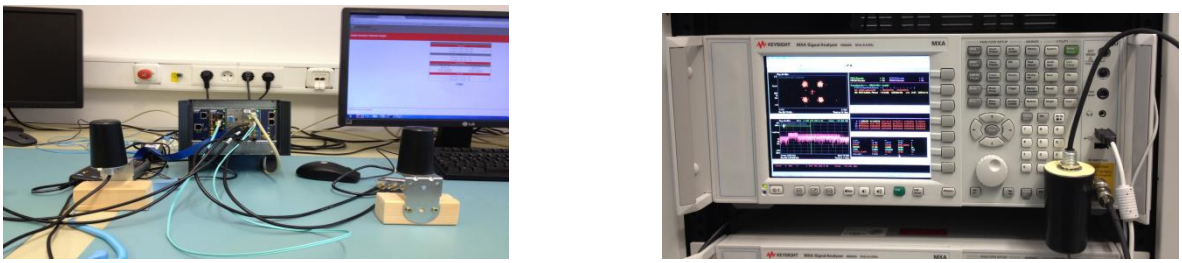


Figure 5-15 Measurement setup

5.3.4 Results

Keysight's VSA software running on MXA signal analyzer is capable of real-time capturing and decoding of time domain LTE signals over the air. The VSA software can show very useful information like IQ constellation plots, channel equalization, spectrum, LTE channel decoding, error vector magnitude (EVM) for each LTE channel etc., that are quite useful for validating and analyzing. The implemented DAS technique is called "antenna selection". For the measurements, 2 transmit antennas at eNodeB, 2 UEs with SISO transmission, QPSK and overlapping resource allocation for both the UEs (RBs 0 to 7) are considered with different radio network temporary identifier (RNTI) for the UEs having been configured. Therefore, 2 Tx antennas are serving 2 different UEs. Different measurement trials [107] have been done, in a closed room, by altering the distance between the two transmit antennas and observing the quality of decoded signals i.e. EVM on the LTE analyzer.

In the first trial as shown in Figure 5-16, single transmit antenna case is considered for reference. EVM measured was 6.5% at a receive power of -69.2 dBm.

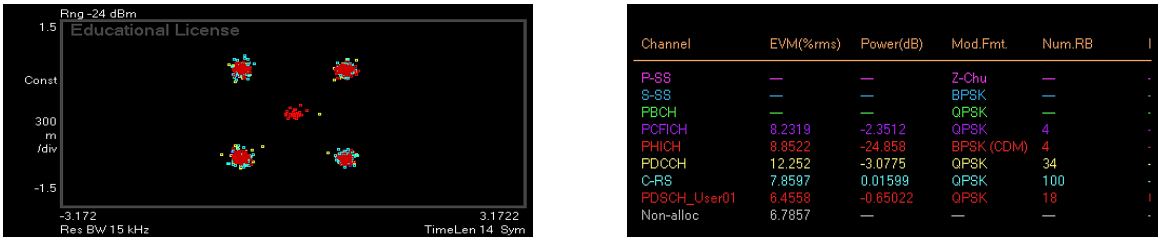


Figure 5-16 Constellation and EVM - Single transmit antenna case

In the second trial as shown in Figure 5-17, the second antenna is also activated. The two antennas are separated by about 300 cm from each other. The measured EVM is about 11.5%, which is still acceptable for decoding the QPSK modulation.

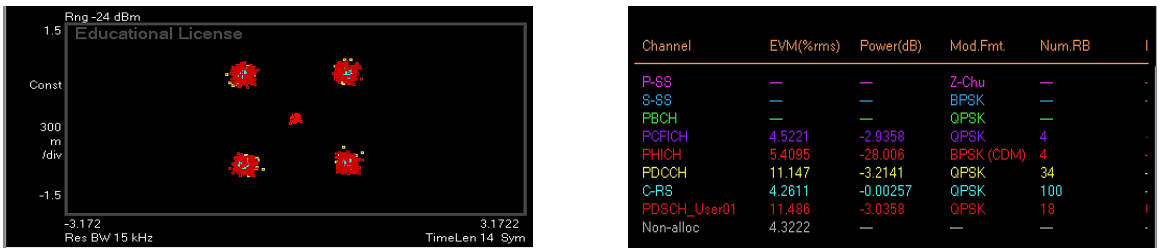
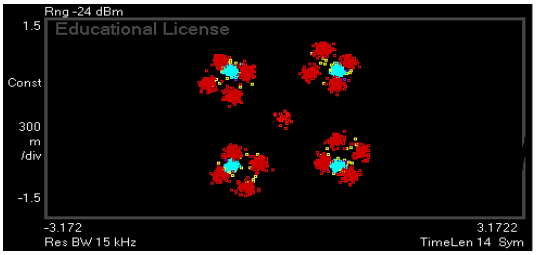


Figure 5-17 Constellation and EVM – Two antennas separated by 300 cm

In the last trial as shown in Figure 5-18, the two antennas are separated by about 200 cm from each other. The measured EVM is about 30.5%. Here the performance is severely degraded.



Channel	EVM(%rms)	Power(dB)	Mod.Fmt.	Num.RB
P-SS	—	—	Z-Chu	—
S-SS	—	—	BPSK	—
PBCH	—	—	QPSK	—
PCFICH	9.523	-2.8601	QPSK	4
PHICH	8.3747	-24.322	BPSK (CDM)	4
PDCCH	16.595	-3.1901	QPSK	34
C-RS	5.6998	0.00707	QPSK	100
PDSCH_User01	30.857	-3.7456	QPSK	18
Non-alloc	5.5698	—	—	—

Figure 5-18 Constellation and EVM – Two antennas separated by 200 cm

It is evident from the trials that, the transmit antenna selection is dependent on the separation distance between the two antennas and with enough separation distance antenna selection technique can be deployed in DAS.

5.3.5 Usage of abstraction

In the COHERENT architecture, the eNodeBs are connected to the COHERENT controller for transmitting and receiving necessary information in the format of abstracted network graphs. The abstracted network graphs basically describe all the participating nodes in that network. The edge of these graphs defines the key parameters pertaining to the interconnecting nodes. In the DAS deployment with antenna selection technique, CQI and SINR per RB allocated are some of the useful parameters from the physical layer perspective for construction of network graphs.

5.4 Coordinated multipoint transmission

Coordinated multi-point transmission and reception (CoMP) has been identified as a promising technology to improve cell-edge user performance and cell coverage in 3GPP LTE-Advanced (LTE-A) system. The focus of the current work is on investigating the downlink CoMP transmission schemes mainly considered in the literature: joint processing and coordinated scheduling and beamforming. In Figure 5-19, (a) represents the joint transmission (JT), (b) represents dynamic cell selection (DCS), and (c) represents coordinated beamforming. RRE means radio resource element. These techniques are described in this section. For this technology, necessary parameters to be abstracted for network graph would be investigated. In the CoMP transmission the related control channels, including the physical downlink control channel (PDCCH), are transmitted only from the serving cell regardless of the transmission scheme [109]. It is targeted to apply the abstraction of network graphs in future deliverable D3.2 of COHERENT project.

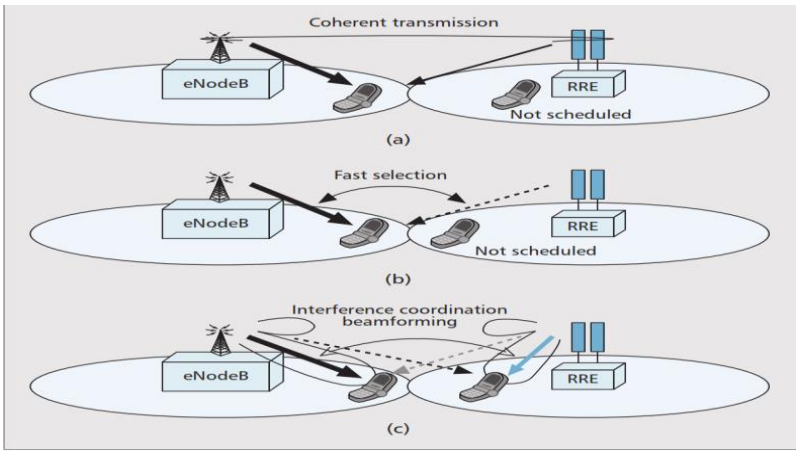


Figure 5-19 CoMP downlink transmission schemes [108]

5.4.1 Joint processing

Joint processing is further categorized into joint transmission (JT) and dynamic cell selection (DCS). In JT the same resource block (RB) of the PDSCH is transmitted from multiple cells associated with a UE – specific demodulation reference signal (US-RS) among coordinated cells (i.e., from non-serving cells as well as the serving cell). In DCS an RB of the PDSCH associated with a UE specific reference signals, is transmitted from one cell among the coordinated cells, and the cell transmitting the PDSCH with the minimum path loss is dynamically selected through the fast scheduling at the central BS.

5.4.2 Coordinated scheduling and beamforming

In a coordinated scheduling and beamforming (CS, CB) approach, an RB of the PDSCH is transmitted only from the serving cell together with the PDCCH. Hence, an RB is assigned to the UE with CS/CB by scheduling of the serving cell. However, scheduling and beamforming is coordinated among multiple coordinated cells. In this case transmit beamforming weights for each UE set are generated to reduce the unnecessary interference to other UE scheduled within the coordinated cells. Therefore, in particular the cell-edge user throughput can be improved due to increase in received SINR.

5.5 Dynamic inter-node OOB interference modelling

5.5.1 Inter-system interference characterization

The inter-system interference depends significantly on the modulation techniques used in both systems. As it is foreseen that the 5G system will be based on some multicarrier PHY we will focus on such a scenario, i.e., interference caused by this system to another system and interference caused by another system to this system. Some results on the interference between LTE (OFDM-based) and UMTS and GSM can be found in (ETSI TR 1 [208]) and (ETSI TR 2 [209]).

Standards give some requirements on transmitted signal spectrum shape and requirements on the interference rejection capabilities at the receiver, e.g., in ETSI TR 4 [210] and ETSI TR 4 [211]. In ETSI TR 4 chapter 6.6, the requirements on the LTE BS TX are defined that help to limit interference power at other systems adjacent in frequency. The low Out-of-Band (OOB) radiation power can be obtained first by limiting the occupied bandwidth. Bandwidth of the system is defined as a continuous frequency range where at least 99% of the total transmit power is emitted. The bandwidth is limited in the standard. Secondly, adjacent channel leakage ratio (ACLR) is limited. It is defined as a ratio of power transmitted in a given channel to the power transmitted in the adjacent (in frequency) channel. Importantly, in Section 6.6.2.1-1, the in-band power is calculated over “channel bandwidth” and OOB radiation is calculated over “transmission bandwidth. As defined in Table 5.6-1 for channel bandwidths 1.4, 3, 5, 10, 15, 20 MHz the transmission bandwidths are 1.08, 2.7, 4.5, 9, 13.5, 18 MHz. In all the cases the minimum required ACLR equals 45 dB. Additionally, absolute limit on the power leaking into adjacent channel is stated, e.g., -32 dBm/MHz for Local Area BS and -50 dBm/MHz for Home BS. Additional unwanted emission limits dependent on the Base Station type and transmission band are given in Section 6.6.3 using dBm measured using a given filter bandwidth, e.g. 100 kHz. It is defined like a spectrum emission mask, i.e., requirement changes with the frequency offset from the occupied bandwidth. It is defined for the frequency range spanning 10 MHz lower or higher than the channel bandwidth. For all frequencies outside this range but within a band 9 kHz-12.75 GHz “spurious emission” limits are defined in Section 6.6.4. Maximum (over frequency) transmitted power (dBm) averaged in time is defined using a given measurement filter bandwidth, e.g. 1 kHz, 1 MHz. In the same section, additionally, some extra limits are stated for some specific frequency ranges, e.g., occupied by GSM900, or UL band for FDD BS operation in Section 6.6.4.2.1.

From the perspective of receiver standardisation “Dynamic range” limits (defined in Section 7.3 of in ETSI TR 4) are interesting. This is a measure of capability of a considered system to operate (achieving at least 95% of the maximum throughput) under interference (modelled as a white noise) present in the same channel. Similarly as previously, the interference or noise power is defined over transmission bandwidth. As an example, for the Local Area BS and 5MHz channel bandwidth, the receiver has to operate for wanted signal power equal to -62.2 dBm and interference signal power equal to -74.5 dBm

assuming Fixed Reference Channel A2-3. Simple calculations show that the receiver should work for an SIR equal or higher than 12.3 dB.

In Section 7.4 requirements on "In-channel selectivity" are defined. In this case it is assumed that two E-UTRA signals (from 2 UEs) are received in adjacent RBs (non-overlapping). Both signals are assumed to be time-synchronized and the BS has to be able to receive a weaker one achieving at least 95% of the maximal throughput. For Local Area BS the signal of -92 dBm (over 15 RBs) should be successfully received for the interference of -73 dB (over 10 RBs), i.e., power of a single RB of an interferer is 20.76 dB higher than the victim RB power.

While the "In-channel selectivity" deals with signals transmitted on separate spectrum resources but received within the BS RX bandwidth, the "Narrowband blocking" and Adjacent Channel Selectivity (ACS) defined in Section 7.5 specify robustness to the interfering signal transmitted outside of the BS RX channel bandwidth. Narrowband interference is an E-UTRA signal utilizing only a single RB. It is assumed that the wanted signal is a few dB stronger, e.g., 6 for Local Area BS and 5 MHz channel, than the sensitivity level (defined in Section 7.2). In this case Narrowband interference of -41 dBm has to be acceptable while -44 dBm in the case of interference from fully utilized E-UTRA transmission.

For the E-UTRA interfering signals being more distant in the frequency (e.g. up to 20 MHz from the wanted channel edge) the "Blocking" requirements are defined. For the interfering signals more distant in the frequency blocking requirements are defined assuming Continuous Wave interfering signal.

The last parameter that is worth to be mentioned is "receiver intermodulation". In the case of two or more signals being transmitted nearby, some nonlinear operations can be observed in the BS frontend causing high effective interference power. According to Table 7.8.1-2 it is assumed that in the band adjacent to the wanted E-UTRA transmission two interfering signals are transmitted, i.e., CW and another E-UTRA signal. The frequency separation is chosen in such a way that intermodulation product falls within interfered E-UTRA band. The protection mechanisms in the UE are similar as the one presented for BS, as such not shown.

It is important to mention that the above presented specification aims at reducing effective interference between different transmitters and receivers. These values have been derived by means of calculations, simulations, or measurements as shown e.g. in ETSI TR 1. In general, as shown in Figure 5-20 below, the effective interference is dependent on the characteristics of the TX and RX nodes. The TX (System A) transmits some signal utilizing a given band with a given spectrum shape. While keeping the given in-band spectrum characteristic (e.g. power), the OOB radiation power (measured sometimes using ACLR) is decreased in order to decrease effective interference power at system B RX.

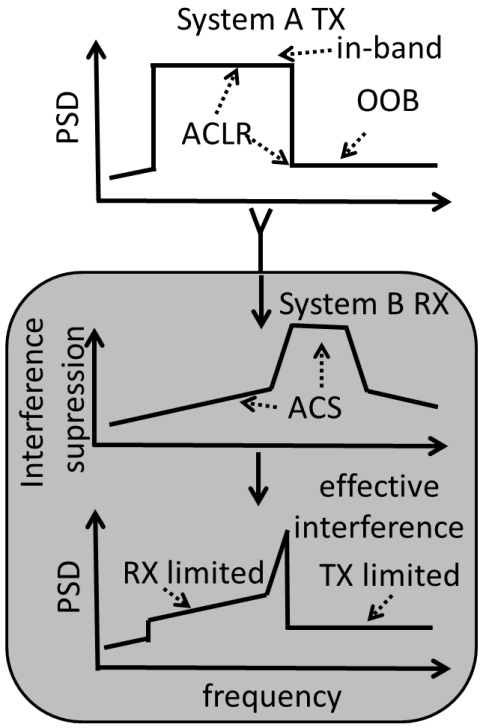


Figure 5-20 TX and RX characteristic influence on the effective interference

However, for very low OOB radiation power the dominating factor causing effective interference is limited selectivity of System B receiver (shown as an ACS and filter characteristic in the middle of Figure 5-20). The combination of TX (ACLR) and RX (ACS) characteristics is typically described by means of Adjacent Channel Interference Ratio (ACIR). ACIR can be defined as a ratio of mean interference power at the victim RX antenna port (total power in the first plot in Figure 5-20) to the effective interference power degrading transmission (assumed 0^{th} in-band attenuation of useful signal-aggregated power in the bottom of Figure 5-20).

The existing standards (as shown above) have static requirements on ACLR/ACS. However, as the number of systems operating in adjacent bands increases (as well as the number of a given system configurations), it is foreseen that a more dynamic approach to interference modelling and protection is required.

5.5.2 Influence of PHY modulation techniques on inter-node interference

Most of the recently introduced wireless communications systems are based on OFDM modulation, e.g., IEEE 802.11n, and LTE. It is foreseen that a 5G system will be based on some multicarrier modulation scheme as well, e.g. OFDM or filter bank multicarrier (FBMC) [84].

There are different sources of inter-node interference in multicarrier PHY-based systems. If different nodes utilize the same frequency resources in the same time the interference will be dependent mostly on the transmission power and pathloss (there is no Code Division Multiple Access like in UMTS). However, if both nodes use different time or frequency resources, interference can be significantly reduced.

In the case of OFDM the dominating source of inter-node interference is a rectangular pulse shape used. Theoretically, if both systems use OFDM modulation all nodes can be frequency and time synchronized. While accurate (small fraction of OFDM subcarrier spacing, i.e., 15 kHz in LTE) frequency synchronization is required, cyclic prefix allows tempering time synchronization requirements. If all users utilize different time and frequency resources, all transmitted signals are orthogonal causing no inter-node interference. In practice different nodes can have synchronized local oscillators (thanks to GPS reference clock) but some carrier frequency offset will occur because of the

Doppler Effect. Moreover, propagation time delay between different nodes can violate the orthogonality between different subcarriers. As such, in the case of limited coordination between nodes, inter-node interference can occur. In the worst case, the power of interference decreases with squared frequency distance between utilized subcarriers. It is caused by sinc-shaped frequency response of a single subcarrier. However, many signal processing algorithms have been designed in order to reduce OOB radiation of OFDM [84].

In the case of FBMC transmission the time domain impulse is not rectangular. It is designed in order to be "well localized" in time and frequency domain. Although in the case of perfect synchronization pulses on different subcarriers are not orthogonal (differently from OFDM), interference is negligible. Importantly, FBMC pulses have lower side lobes power than frequency domain response of a single OFDM subcarrier. As such if the nodes are not accurately synchronized in time or frequency, significantly lower interference power is expected. Exact values depend on a chosen prototype filter. It is foreseen that FBMC-based systems can operate without time coordination (asynchronously) in adjacent frequency bands without increasing interference power (in comparison to synchronous transmission).

The other sources of inter-node interference are distortions caused by analogue frontends in TX and RX side. The most common problem is nonlinear distortion in power amplifiers. Both OFDM and FBMC have high transmitted signal amplitude variations. While such a system passes through real front-end, clipping of the strongest samples can occur. In the frequency domain it results in an increased OOB radiation power (TX side) and intermodulation products of interference signal falling into wanted signal band (RX side). However, there are different techniques to decrease the influence of this problem, e.g., front-end predistortion, Peak-to-Average Power Ratio reduction, and Input Back-Off increase. As shown in [84], in the presence of nonlinear front-end and high energy efficiency requirement (low IBO value) both OFDM and FBMC provide similar ACIR values. The other analog front-end-originating distortions can be, e.g. IQ imbalance or local oscillator leakage.

It is foreseen that based on a given modulation scheme parameters and front-end model the expected ACIR value can be calculated before the transmission and used e.g. for resource allocation.

5.5.3 Proposed solution and results

It is shown above that there are many different constraints provided in standards in order to limit interference between different nodes utilizing different frequency resources. However, as the number of different wireless systems used in a given band increases (considering dynamic spectrum utilization) and the number of possible configurations of a given standard increases, the protection of all systems from interference becomes harder. On the other hand, standards should be designed in order to protect different devices from harmful interference in the worst case scenario. As such, even in the absence of this "worst case" scenario, wireless devices have to significantly limit OOB radiation (TX) or increase receiver selectivity (RX) sacrificing some resources, e.g. throughput, power, computational power, time.

Two interesting solutions to dynamic inter-node interference modelling has been presented [109, 110]. In [109], it has been shown that Spectrum Emission Mask can be calculated dynamically based on the assumed victim receiver type (ACS known) in order to maximize achievable throughput. It showed that the achievable throughput is much higher than in the case of static SEM/ACLR requirements (common in standards). The gain comes from utilizing path loss to the victim receiver. The interference is limited significantly (limiting throughput too) only when the path loss is relatively low. Additionally, the proposed method assumes some minimum OOB radiation power (caused e.g. by the power amplifier nonlinearity) keeping it close to the real frontend transmission.

In [110], the concept of a known ACS is used in order to design spectrum shaping algorithm at the OFDM transmitter. Such a transmitter maximizes ACIR that is closer to optimality than the typically used ACLR maximization. The ACIR value can be increased (reducing interference at the victim receiver) by sacrificing some system resources, i.e., throughput or computational power.

Here, an approach to identification of interference coupling between two systems, namely GSM and LTE operating in adjacent bands is presented. It is foreseen that a mobile network operator want to deploy 5G/LTE-like network using the same bandwidth as GSM carrier [212]. The migration from GSM standard may take some time as there are still some users/applications utilizing this standard deployed. While LTE transmits subcarriers in blocks of 12 called Resource Blocks (RBs), GSM uses a carrier of 200 kHz bandwidth. As presented in Sections 5.5.1 and 5.5.2, the effective interference depends both on the RX filter characteristic, i.e., ACS, and TX spectrum shaping, i.e., ACLR/SEM. In the case of GSM the SEM is defined in [175] and ACS has been calculated in [176]. For the LTE we assume that the GSM carrier operates in the band of LTE carrier, as such limits put on TX/RX in standard, presented briefly in Section 5.5.1 are not appropriate. This is a more flexible, yet challenging scenario, from the perspective of future HetNets. As such, interference observed at a given frequency can be calculated investigating properties of IDFT and DFT used in the LTE transmitter and the receiver, respectively. By assuming digital modulator/demodulator implementation and considering the approach presented in [177], the PSD of *n*-th subcarrier at pulsation ω in TX is

$$PSD_{OFDM-TX}(\omega,n)=\frac{1}{N(N+N_{CP})}\left(\frac{\sin\left(\left(\frac{\omega}{2}-\frac{\pi n}{N}\right)(N+N_{CP})\right)}{\sin\left(\frac{\omega}{2}-\frac{\pi n}{N}\right)}\right)^2,$$

where N is size of IDFT used and N_{CP} is number of Cyclic Prefix samples. The magnitude of frequency response in RX is

$$|H_{OFDM-RX}(\omega,n)|^2=\frac{1}{N}\left(\frac{\sin\left(\left(\frac{\omega}{2}-\frac{\pi n}{N}\right)N\right)}{\sin\left(\frac{\omega}{2}-\frac{\pi n}{N}\right)}\right)^2.$$

The overall interference caused by *n*-th subcarrier to GSM receiver can be calculated as

$$P_{OFDM-GSM}(n)=\frac{\alpha_{OFDM-GSM}}{2\pi}\int_{-\pi}^{\pi}PSD_{OFDM-TX}(\omega,n)\,|H_{GSM-RX}(\omega)|^2 d\omega,$$

where $\alpha_{OFDM-GSM}$ is a coefficient dependent on pathloss, antenna gains between OFDM TX and GSM RX, and $|H_{GSM-RX}(\omega)|^2$ is magnitude of GSM RX filter based on ACS defined in [176]. The overall interference power caused to *n*-th subcarrier by the GSM carrier can be calculated as

$$P_{GSM-OFDM}(n)=\frac{\alpha_{GSM-OFDM}}{2\pi}\int_{-\pi}^{\pi}PSD_{GSM-TX}(\omega)\,|H_{OFDM-RX}(\omega,n)|^2 d\omega,$$

where $\alpha_{GSM-OFDM}$ is a coefficient dependent on pathloss, antenna gains between GSM TX and OFDM RX, and $PSD_{GSM-TX}(\omega)$ is PSD of GSM carrier (here assumed worst case, i.e. SEM presented in [175]). In both the above presented cases the interference caused by/to all active subcarriers has to be added. The respective PSDs or magnitudes of RX filters frequency responses are presented in Figure 5-21. It is assumed GSM carrier occupies RB 0 while LTE uses only RB 3. OFDM has high power sidelobes both at the TX side (left axis) and RX side (right axis) that dominate the interference coupling between both systems.

The final metric that can be useful to populate COHERENT network graph can be a pair of ACIR values presented as a function of RB shift in Figure 5-22. The ACIR of interference caused to GSM RX is normalized in order to obtain 0 dB ACIR for RB overlapping in frequency. The received interference power $P_{interf-GSM\,RX}$ caused by a given RB can be calculated as

$$P_{interf-GSM\,RX}\,[dBm]=P_{OFDM\,TX\,RB}\,[dBm]+G_{TX}[dBi]+G_{RX}[dBi]-PL_{OFDM-GSM}[dB]-ACIR\,[dB],$$

where $P_{OFDM\,TX\,RB}$ is power transmitted in a single RB at OFDM transmitter, G_{TX} and G_{RX} are transmitter and receiver gain, $PL_{OFDM-GSM}$ denotes pathloss between OFDM TX and GSM RX and ACIR is a value taken from Figure 5-22. In the case of interference caused to a given OFDM RB by GSM TX the above equation can be rewritten as

$$P_{interf-OFDM\ RX}\ [dBm] = P_{GSM\ TX}\ [dBm] + G_{TX}[dBi] + G_{RX}[dBi] - PL_{GSM-OFDM}[dB] - ACIR\ [dB],$$

where $P_{interf-OFDM\ RX}$ is interference caused to a given RB, $P_{GSM\ TX}$ is transmit power of GSM carrier, $PL_{GSM-OFDM}$ is pathloss between GSM TX and OFDM RX. In this case ACIR equals 0 dB when the whole received interference power is integrated over whole OFDM band. In the case of a selected RB overlapping in frequency with GSM carrier (offset equal 0) ACIR equals 1.13 dB. The GSM carrier has wider bandwidth than a single OFDM RB.

This model characterizes only waveform specific TX/RX. It does not consider nonlinearity distortion (increasing interference) or some additional filtering applied at TX/RX (decreasing interference) as these are vendor-specific. In a more general case it is possible that each vendor provides its ACLR/ ACS (recall that 1/ACIR = 1/ACLR + 1/ACS) values that can be used to improve the network graph.

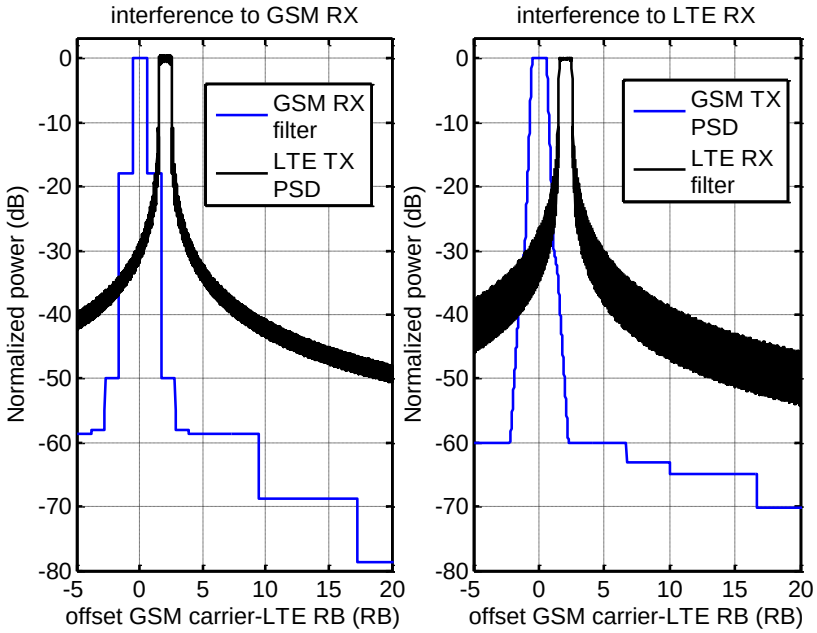


Figure 5-21 Plots of coexistence between single occupied RB and GSM carrier

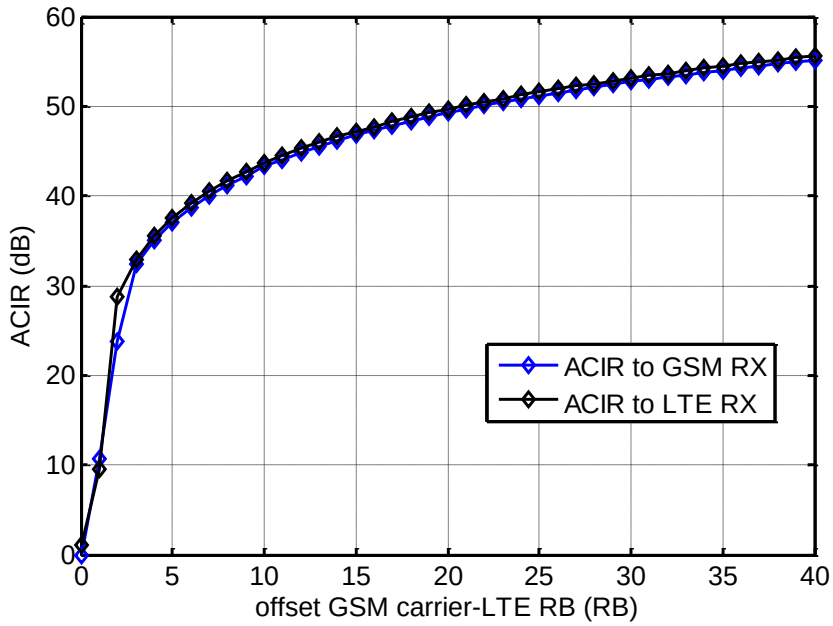


Figure 5-22 Normalized ACIR values for interference cause by/to GSM carrier to/by single RB

5.6 Load balancing

Load balancing in a heterogeneous network consists in distributing the load in the network by associating users to BSs in lightly loaded tiers, such as macro, pico and femto cells, relay BS and D2D communication pairs. Each tier differs in transmit power, cell size, deployment density, bandwidth, backhaul, cost, ease-of-deployment and propagation characteristics [111].

Cell association is based on metrics such as maximum RSSI, maximum SINR, maximum achievable rate, nearest distance and strongest channel gain. In HetNets, such cell association results in unbalanced load distributions due to different power in BSs and propagation characteristics [113]. To balance the load of the cells, a bias is added to the metric used for cell selection such that more users can be associated to a cell, in particular to a low-power BS [113]. This technique is known as cell range expansion (CRE).

Possible deployments for load balancing are co-channel or out-of-band transmissions. In co-channel deployments the cells operate in the same frequency band. In out-of-band deployments the cells operate in at least two frequency bands. One notable example for this last case is the use of licensed and unlicensed bands.

Load is balanced with the objective of maximizing a metric of the network, such as resource utilization, fairness, waiting time in queues, or throughput. For this purpose an optimization problem is defined in terms of maximizing a sum utility of the system, in which, for co-channel deployments, user association and user scheduling are coupled due to the load and interference in the network. The optimal user association and scheduling is a complex, NP hard, combinatorial optimization problem. In [114], the problem is converted to a convex problem by assuming a full buffer model, and relaxing the binary constraint on one-to-one cell-user association to a real number, allowing users to associate to all the BSs in the system. The solution gives an upper bound in performance.

Co-channel interference in HetNets is managed by means of intercell interference coordination (ICIC). In ICIC, time and/or frequency resources are muted for transmission, thus mitigating interference [115]. In the particular case of LTE, the granularity of these resources involves resource blocks in frequency domain, and sub-frames in time domain [116, 117]. LTE Rel. 10, under the category of enhanced ICIC (eICIC), defines almost-blank-subframes (ABS), in which signaling remains in the subframe and data is muted. An example of co-channel interference in UL is the case of macro cell users interfering small cell BSs. An example for DL is in the case of femto-cells, in which a closed subscribed group (CSG) policy forbids non-member users to associate to the femto cell. These users are served by the macro cell, and receive strong interference from the femto cell.

Load balancing has been analysed from different perspectives [112], such as, relaxed optimization [114], Markov decision processes [118], game theory [119] [120], CRE [117], and stochastic geometry [113]. According to [112], the factors that should be considered in a comprehensive study of load balancing include: backhaul bottleneck, mobility, UE capability, asymmetric DL and UL, and D2D offloading.

5.6.1 Enhanced distributed load balancing

The proposed load balancing approach in [121] is a distributed on-line mechanism capable of autonomously assigning target load values for each involved node, based on the observed load situation in its environment. The target load estimates together with the actual loads are used as input to a mechanism that redistributes and balances load within the network among adjacent nodes. Compared to previous work [122, 123, 124, 125], the approach proposed dynamically assigns target loads in a distributed manner under varying network conditions.

We will enhance the load balancing mechanism by investigating several aspects that improves the current mechanism and the support to the COHERENT network graph, including enhanced target functions for load balancing and development of probabilistic thresholds for triggering distributed updates of target loads.

The proposed load balancing approach described in [121] is based on a distributed on-line mechanism capable of autonomously assigning target load values for each involved node, based on the observed load situation in its environment. The difference between the actual load and the target is used as input to an actuation mechanism (LTE CRE) that redistributes and balances load among adjacent nodes. Compared to previous work [122] - [125], the approach dynamically assigns target values in a highly distributed manner and scales and adapts well to varying network conditions.

Within COHERENT, we are working on enhancing the balancing mechanism by investigating several potential improvements on the mechanism, and instrument it to support and exploit the COHERENT network graph, including specifying enhanced target functions and developing probabilistic thresholds for triggering and adapting the distributed mechanism, using information stored in and derived from the network graph. The first key findings and preliminary results are presented below.

We also include an overview of the kind of metrics and abstractions we can supply to the network graph and an outlook on using data available in or obtained through the graph to control the behavior of the distributed algorithm.

5.6.1.1 Technical background

The mechanism described in [121] is an on-line distributed and self-adaptive balancer of radio resource load for use in an environment with a mix of RAN nodes from different tiers. The actuation mechanism used, is dynamic reconfiguration of the CRE parameter within each involved cell, but the way of deriving suitable values for the parameters is novel, and consists of a highly distributed and scalable algorithm, where the controller responsible for each cell *i* maintains a local neighborhood of other cells with which *i* had recent handovers. Moreover, the controller also maintains a *target load value* for the cell which is computed as the average of the target loads in the neighborhood of *i* and an exponentially decaying mean (EWMA)

$$\hat{l}_j = \frac{w\hat{l}_{j-1} + l_j}{w+1}$$

of the load of *i*, where l_j is the observed momentary load, $\hat{l}_{j-1}$ the previous (EWMA) estimate, and *w* is a weight, determining how slowly the influence of older momentary loads decays. The *difference* between the target value and load EWMA is then scaled to a range of CRE values deemed acceptable w.r.t. QoS of individual users, normally ~ ± 3 dB. This is done, taking into account the range of differences between EWMA loads and target values of nodes in the neighborhood, so that the range of acceptable CRE values gives a maximum actuation latitude.

As showed by simulations described in [121], the mechanism works well for a range of traffic and mobility scenarios, in the sense that it does improve overall load balance significantly, and that load is effectively shifted to smaller cells in typical HetNet scenarios but works to offload them in situations where they instead become overloaded, e.g. flash-crowd scenarios. Figure 5-23 shows load distribution examples for the unbalanced and balanced cases.

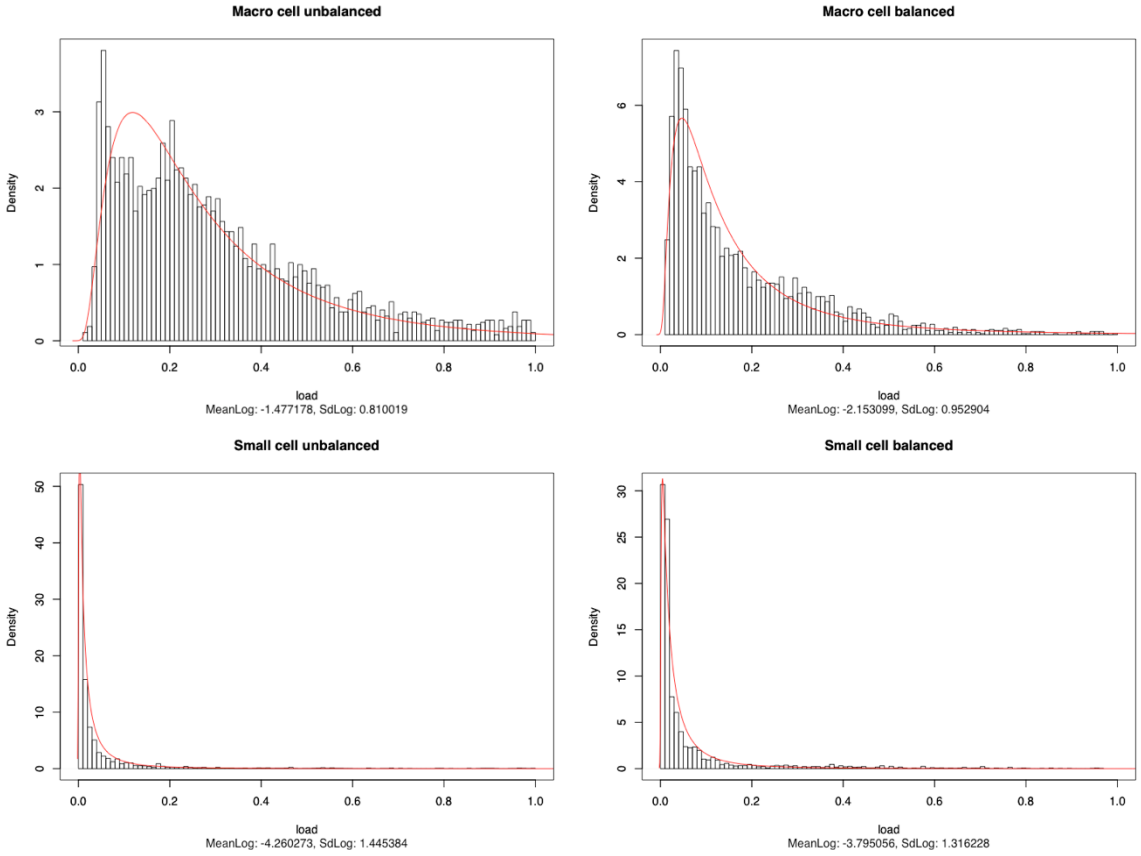


Figure 5-23 Load distributions for unbalanced and balanced cases for a macro cell and an adjacent small cell

A disadvantage of this mechanism is that it tends to use the full range of CRE offsets even when none of the involved cells are highly loaded, which potentially impacts QoS for UEs forced to connect to a node with less SINR than the best one. In these situations, the load balance is not an ideal metric of the system performance. At least it has to be weighed against other metrics of system efficiency.

Another limitation is a lack of a more global view of the network state. The design of the mechanism allows each involved node a limited view of the entire network since its target metric is computed using the *target metric of its neighbors*, which in turn is based on the target metrics of *their* neighbors, etc. In certain types of scenarios, this mechanism may prove to be insufficient, or react too slowly to sudden changes, and for these, the more global view provided by the network graph can be utilized. In order to exploit that opportunity we first need to consider what kind of metrics each cell should provide to the network graph, and what kind of processing of that information is needed to support improved decisions. To investigate these issues, we will explore a number of scenarios where the current mechanism does not work optimally.

In this report, however, we will mainly address the issue of unnecessarily adjusting CRE for systems of lightly loaded cells, i.e. a first step to consider, more generally, the QoE impact of the mechanism.

5.6.1.2 Proposed solution

The first approach we propose to address this issue is to not balance the absolute (EWMA) load of cells, but instead estimate and balance the *risk* of reaching states of high loads.

An alternative would be to modify how we scale the difference of the metric between adjacent nodes into the range of CRE values, but doing this in a way that preserves the latitude of the actuation

mechanism in situations with high load is not trivial. Instead we propose to characterize the distribution of the observed load values so that we may estimate the risk of reaching a given threshold. Depending somewhat on the model we use for this characterization, and accuracy of the estimation method, we should be able to balance the risk of reaching these high risk states using the full range of available CRE values, but avoiding assigning significantly different CRE values to cells in situations where the risk of overload is insignificant for all of them.

5.6.1.2.1 Solution / algorithm description

The risk model we propose is related to the one used in [150], which uses the amount of mass in the CDF of the probability distribution beyond a fixed threshold corresponding to or close to the capacity of the utilized resource. In order to accurately estimate this risk, we need to model the empirical distribution well enough to compute the probability of observing loads higher than the threshold. As a first approach, we will assume that the load distribution is lognormal. Currently, the only reason for this assumption is that we have experience with this type of model and can quickly produce a prototype and experiment with model parameters. In addition, in our simulator, we did actually observe (again, see Figure 5-23) empirical distributions which have a clear lognormal signature. This was a bit of a surprise since even though we have seen numerous examples of lognormal throughput distributions in recordings from both land lines, and in WCDMA data, most of the event generating processes in the simulator samples from more heavy-tailed distributions, Pareto in the case of the traffic, and truncated power-law in the mobility model. If load distributions seen in the simulator are examples of the type of smoothing we see in real traffic traces, or something else, is too early to say.

To verify or falsify this model assumption we will have to investigate empirical load statistics for collections of cells. A next step will be to adjust the method to those findings, e.g. modelling the distribution with other parametric methods, or inspecting quantiles of the empirical distribution directly, if necessary. Meanwhile we can report (in Section 5.6.1.3 below) on the results using this model within our simulator.

As noted above, the balancer attempts to regulate the CRE of each individual cell in order to equalize a given target metric. In the previous version of the mechanism, the metric was based on the EWMA of the momentary radio resource usage. Replacing the momentary load metric, which in the simulator is actually the mean over whole seconds, with a continuously updated estimate of the risk of exceeding 90% of the cell capacity turned out to be relatively straight-forward, using previous results. For evaluating and tuning the new mechanism, however, we had to add several new plots and metrics. The key new metric is computed as follows:

First compute the load EWMA as in calculation of $\hat{l}_j$ above, and compute in addition a corresponding exponentially decaying variance measure as

$$\widehat{S}_j^2 = \frac{\mathbf{w}\widehat{S}_{j-1}^2 + \frac{w}{w+1}(l_j - \hat{l}_j)}{w+1}$$

where l_j, $\hat{S}_{j-1}^2$ and w are as in calculation of $\hat{l}_j$ and $\hat{S}_j^2$ are the consecutive variance estimates.

Then estimate the parameters of an (assumed) lognormal load distribution using the method of moments (MoM):

$$\begin{cases} \widehat{\mu}_j = \ln(\hat{l}_j) - \frac{1}{2}\widehat{\sigma}_j^2 \\ \widehat{\sigma}_j^2 = \ln(1 + \frac{\widehat{S}_j^2}{\hat{l}_j^2}) \end{cases}$$

and use the CDF of the estimated distribution of the load L to extract a momentary risk estimate $\hat{c}_j$ of observing loads over a given fraction (here 90%) of the cell capacity r.

$$\hat{c}_j = P(L > 0.9r) = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left[\frac{\ln(0.9r) - \hat{\mu}_j}{\sqrt{2}\hat{\sigma}_j}\right]$$

We do not need to maintain any EWMA of the risks, since the decay of each observation is already encoded in the parameter estimates. This derivation should be more properly checked for Bayesian accuracy.

5.6.1.2.2 Network graphs and abstraction aspects

The metrics listed below are tentative at this time. We will continue work to prioritize which metrics are most useful to our own purposes, but also encourage feedback on what could be most useful for other purposes within the project.

5.6.1.2.2.1 Measurements

- Delivered to the graph:
  1. Raw momentary radio resource consumption for each cell (probably undesirable because of high data rate, see below)
  2. Corresponding additional measures for other resources, e.g. backhaul capacity utilization, and base station compute load

5.6.1.2.2.2 Abstraction Metrics

- Delivered to the graph:

  1. EWMA relative radio resource load for each cell
  2. Absolute local balance for each cell, i.e. difference between target and one or more EWMAs of actual load
  3. Local load balance (i.e. variance of the above measures) in the neighborhood of each cell
  4. Same data for any given quantile of load distribution
  5. If distribution turns out to fit well to a parametric distribution, estimation of its parameters for each cell, and risk of observing an arbitrary load level
  6. Detections and alarms of e.g. developing flash crowds, possibly also of network partitioning, based on computed neighborhood and node pair protocol used to compute target loads
  7. Triggering startup of backup equipment, or allocation of additional spectrum, if possible

- Read from the network graph (examples and hypothesis only, at this time)
  1. CQI data, for setting the range of acceptable CRE values in a neighborhood
  2. More global load data to adjust e.g. temporal properties of the actuation mechanism

5.6.1.2.2.3 Implementation of the network graph

The use case most relevant to how we plan to use and contribute to the network graph, is “Use Case 1.RS: RAN sharing among heterogeneous mobile networks in Scenario 1: Network cooperation and inter-operability”, as described in deliverable D2.1 [135]. Within this context, we expect to deliver at least a subset of the metrics mentioned above for use in controlling the distributed mechanism.

It is currently too early to specify exactly what metrics will be used, and how information will be represented and processed in the graph to be useful for our purposes, but as a coarse first outline we envision an architecture that supports regional controllers where metrics generated by individual nodes can be combined and correlated with e.g. CQI measurements, and represented to support the configuration decisions in the local controllers.

5.6.1.3 Assessment / Evaluation

5.6.1.3.1 Evaluation methodology

The mechanism previously described was evaluated on achieved load balance only, and taking into account QoS impact only by limiting the range of CRE-values allowed. For the improved mechanism we will instead need a measure that takes the actually achieved performance of the system into account. A metric that does capture an important aspect of this is the total throughput of the system, but since that also depends on the state of the dynamic traffic and mobility models in the simulator we will focus use the percentage of total throughput demand satisfied by the system, and it changes over time.

The traffic model generates demands from each user sampled from a Pareto distribution, which expressed as a request for a throughput in kb/s. The length of the duration over this throughput is required is sampled from another Pareto distribution, while the following idle time is log-normal. These distributions were fitted to WCDMA data made available to us in ~2012, but can easily updated given e.g. reliable statistics on LTE or 5G data. Since the radio model is currently based on a limited number of LTE base station stats, and we are interested to stress the system into high load states, we have increased the mean of the throughput Pareto sampling mechanism with a single adjustment of the slope parameter, and scaled further by adding UEs, which is coarse, but does appear to give clear indication of the difference between the older mechanism, and the new one.

The throughput demand is uncorrelated with the UE location and may therefore, due to fading, occasionally result in excessive demands on radio bandwidth. The radio resource scheduler of simulated nodes is currently extremely simple, and will simply disconnect any UE with demand exceeding the remaining capacity. This is, of course, highly unrealistic, and will have to be improved in further developments of this work, but for what it is worth, the results so far, are summarized below, and appear quite promising.

5.6.1.3.2 Evaluation/Simulation/analytical results

5.6.1.3.2.1 Qualitative results

Based on experiments with the new mechanism we have verified that in general, balancing the overload risk reduces the negative impact on QoS since the mechanism tends to use the full range of CRE only in situations when the risk of overload would also have significant QoS impact. This means that we may be able to allow a wider range of CRE values. The standard allows values of ±10 dB while our colleagues in the industry claim that more than ±3 dB could have unacceptable QoS impact in practice.

The first qualitative observation is that while the older mechanism tends to always regulate at least two nodes to the extremes of the range, regardless of the load level and regulation latitude, the new mechanism tends to center all involved nodes around the mid of the CRE range, i.e. close to no regulation. In principle this could make it possible to tune the mechanism to operate within a large range but using large offsets from the mid, only in extreme situations. For such tuning we would probably need access to far more accurate, and ideally live data, than we currently have access to.

Figure 5-24 shows time series of load, risk, and CRE parameter settings generated in simulations for the two variants of the method and a scenario with two macro nodes and three smaller nodes. The spatial configuration for this experiment is shown in Figure 5-25.

Note that for this experiment, we relaxed the constraints on the range of CRE values, so that the load balancer uses expanded range ±10 dB almost fully, even at the expense of reduced QoS for many UEs, while the risk balancer spontaneously (for this experiment) restricts itself to the prescribed ±3 dB. This is not always the case, of course, and the overall load level is a bit lower for this particular example, but the tendency is clear, and representative of what we see in many scenarios.
A second observation is that for small cells, the relative difference between the mean and the risk estimate is larger than for the larger ones. We interpret this as an effect of the impact of individual UE

demands can have on smaller cells, which drives up the coefficient of variance. For our improved mechanism, this means that, in general, we will tolerate higher average load in larger cells before engaging the actuation mechanism.

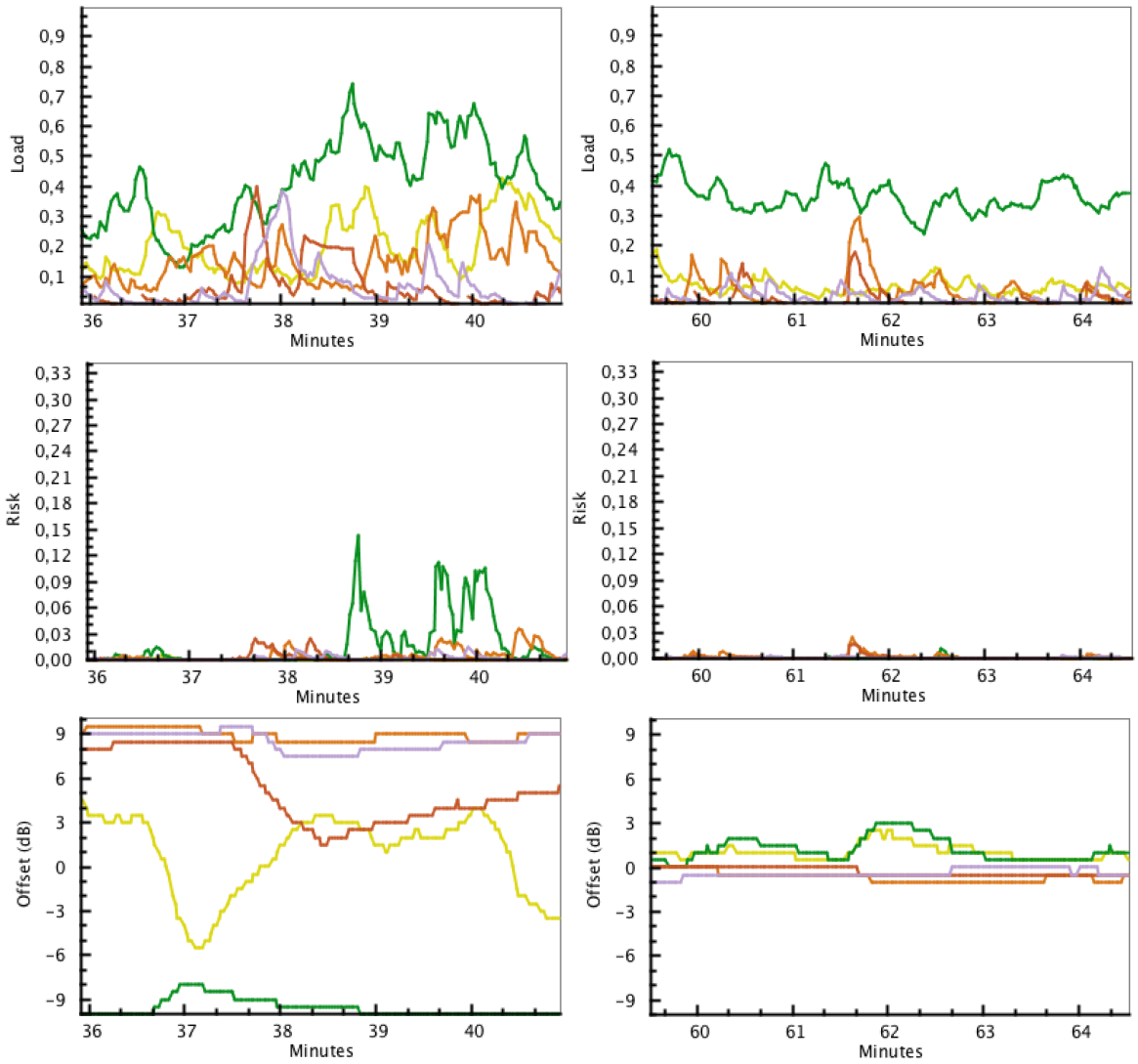


Figure 5-24 Load, risk and CRE value time series for comparable high risk scenarios. Left column balances load, right balances risk.

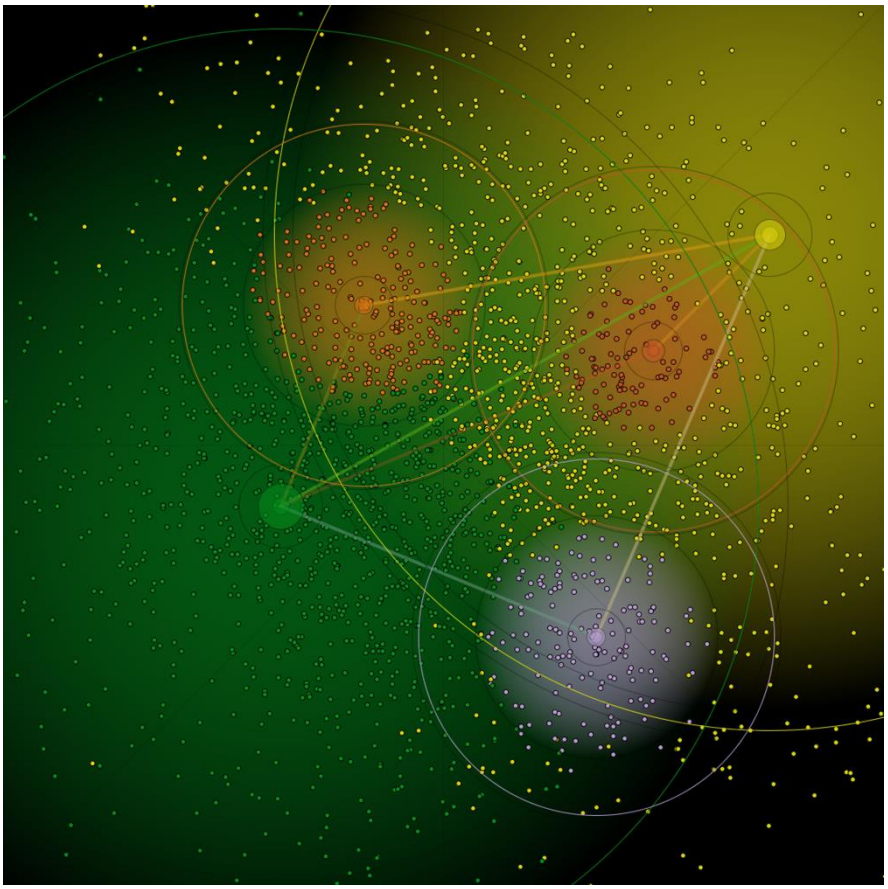


Figure 5-25 Spatial configuration used in qualitative examples

5.6.1.3.2.2 Preliminary quantitative results

At this point, quantitative results are still sparse, but in order to verify that the proposed improvement of the method does indeed improve service quality we have run a small set of experiments with a type of scenario where one macro cell is placed close to a densely populated area and two smaller cells are placed to overlap their coverage and expected to off-load the macro. Some 1250 UEs move around the area according to a type of bounded Levy-walk model, and sample traffic demands and idle and active (connected) periods, normally according to parametric distributions fitted to WCDMA data, but for this particular experiment the throughput samples were drawn from a distribution with increased mean, in order to provoke more congestion. The results for three experiment cell and UE location distributions appear in Table 5-3. Figure 5-26 shows a screen shot of the cell and momentary UE distributions for Scenario 3 of Table 5-3.

The tendency indicated by these preliminary results is clear: The risk balancing mechanism delivers an improvement in terms of overall performance of the system. These numbers are averages over long runs, but for shorter episodes of high risk, the results are even more convincing. We've seen examples the unbalanced case having a service level of less the 20%, while the risk balancer manages to maintain over 90% over similar situations. Such cases are difficult to capture with our current simulator setup, and anyway at best indicative, so further work is need to investigate how typical such cases are. For low risk scenarios, an improvement is also expected, due to minimal impact on QoS from the actuation mechanism for the risk balancing case.

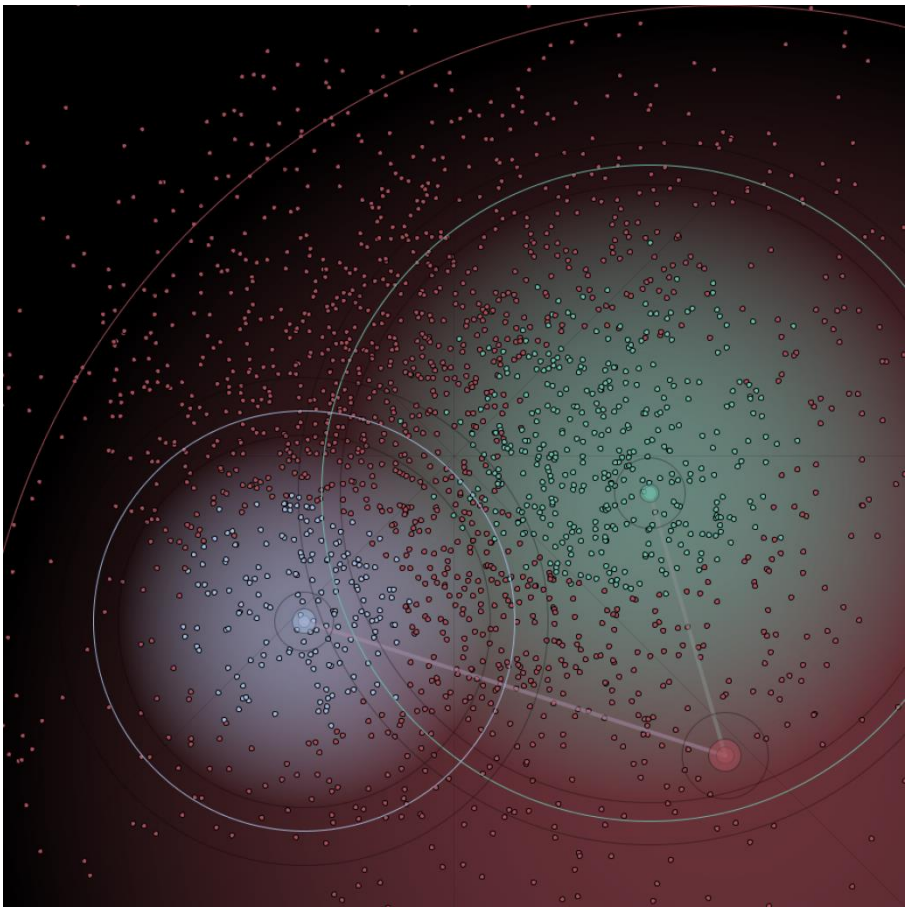


Figure 5-26 Spatial configuration of scenario 3 in Table 5-3

Table 5-3 Proportion of requested throughput served for 3 random placements of one macro and two smaller cells, and ~1250 active/passive and mobile UEs and each of three balancing methods

Balancing	No balancing	Load balancing	Risk balancing
Scenario 1 (30 min.)	83.7%	95.1%	97.5%
Scenario 2 (30 min.)	90.8%	95.1%	99.6%
Scenario 3 (60 min.)	92.6%	95.3%	97.2%

5.6.1.4 Final discussion and summary

These are the first results on a promising improvement of an earlier result, relevant to the COHERENT objectives. We plan to continue work in several directions, as indicated by the WP objectives, in the short term primarily, more extensive quantitative evaluation of the described results. This will involve constructing new and complementary metrics, e.g. handover frequency, or the number active and passive UEs served. It could also include statistics on additional resources. A straightforward first approach to multi objective load balancing would be to minimize a weighted sum of comparable resource usage metrics.

The limitations of the current radio resource scheduler probably also have some impact on the observed results. We will assess the feasibility of adding a slightly fairer scheduler in order investigate more realistic overload scenarios. Verification and/or modification of the model assumptions is also high in our list of priorities, but dependent on an access to better load statistics.

Finally, we will complete the list of metrics we expect to deliver to the network graph, including suitable data structures and abstractions, and outline mechanisms to deploy the network graph to configure our algorithms.

5.6.2 Load balancing in energy-limited HetNets

5.6.2.1 Problem description

Due to limited resources and increasing amount of traffic, offloading or traffic steering for improved load balance between the cells is one of the most significant RRM tasks in heterogeneous networks (HetNets) [112]. Due to expected further densification of upcoming 5G networks, more advanced and accurate load adaptation is needed in order to enable QoS-guaranteed services. A basic feature of 5G HetNets is to allow offloading of traffic from macro base station (MBS) to small base stations (SBSs) for more aggressive spatial reuse of bandwidth or savings in energy consumption. On the other hand, if the resources of an SBS are exhausted, the traffic load can be balanced between SBSs or could be also steered back towards MBS. Typically, a communication system must guarantee minimum performance of users while maximizing network performance. Depending on the primary design objective, there are a number of metrics that can be used to balance the load, including maximum received signal power, minimum interference, maximum user fairness index, maximum network sum capacity, minimum network sum energy, etc.

When the resources of the network are limited, efficient resource usage is of paramount importance to avoid the cells to be congested, i.e. the resource of interest is depleted. If a cell becomes congested, the grade of service in terms of connection blocking and dropping probability will be degraded. In the most general form, load balancing is understood as a way to associate users to desired BSs to avoid congested cells [184]. The congested resource under investigation is most typically the available frequency spectrum. However, the criterion for load balancing decision can also be the available power or available energy of the network node [179].

The load-balancing schemes can be classified based on the balancing cooperation mode, decision making criterion, and user set participating the balancing. The cooperation mode can be based on either resource borrowing from lightly-loaded cells or traffic transferring to lightly-loaded cells [179]. The load-balancing decision criterion can be based on simple threshold triggering or alternatively maximizing selected network utility function. There are a number of possibilities to apply both methods. For instance, the utility function may aim at having the same load in each cell, provided all cells have the same resources and processing capabilities. Since this is typically not the case in HetNets, more sophisticated utility functions are needed instead of aiming at having the same amount of users in each cell. Load balancing method can be *direct*, i.e. only new connection requests participate in load balancing, or *indirect*, i.e. all users participate in load balancing. Obviously, the latter is typically more complicated.

The HetNet system of interest is illustrated in Figure 5-27. Therein, the heterogeneity is revealed by four different aspects: traffic characteristics (M2M, H2H), radio access (SBS, MBS), energy access (on-grid, off-grid), and backhaul link (wired, wireless). The previous studies in HetNets have focused on intercell interference mitigation between SBSs and MBSs [112]. On the other hand, the energy consumption and flexible network deployment problem of very dense HetNets have received increasing attention in 5G community. In particular, the interest in using renewable energy has recently increased especially for dense small cells such as femtocells whose energy consumption is feasible to operate with reasonable size renewable energy sources. The feasibility of using renewable energy with small base station is studied in [178]. In these scenarios, the main expected benefits include easier and faster small-cell network deployments without the need to install on-grid energy and data wires, as well as operational on-grid energy cost savings and reduced carbon emissions, see also [179][181][182]. In practice, not all the nodes can be harnessed solely with renewable energy, leading to a concept of energy-limited partially off-grid HetNet.

In this work, we focus on direct load balancing approach and on heterogeneous traffic-dependent energy consumption of BSs of the partially off-grid multihop SBS backhaul topology. Instead of using a

traditional load criterion to avoid spectrum congestion, we study a provisioned probabilistic criterion which steers the traffic towards the closest SBS, provided the energy depletion probability of the involved SBSs is less than a given threshold. In this perspective, the approach is not aiming at balancing the load evenly between cells, but rather adapting the load of each SBS to avoid energy congestion or depletion of off-grid SBSs having different energy harvesting capabilities. In other words, an adaptive RRM approach is evaluated to balance traffic loads of SBSs based on their respective energy availability. To achieve some way of controlling the connection dropping probability due to potential energy depletions, joint load balancing via multihop routing decisions and sophisticated admission control with a possibility to steer the blocked traffic to MBS are of paramount interest. More specifically, the objective in the traffic load distribution is to minimize the energy depletion probability (EDP) of involved SBSs. Most of the BS energy in cellular networks is typically consumed at PHY layer, as evaluated in [29]. Therefore, energy consumption modelling of PHY is important. At this stage, we focus solely on energy aspects and the detailed evaluation of effects e.g. from limited bandwidth and interference are left for future work.

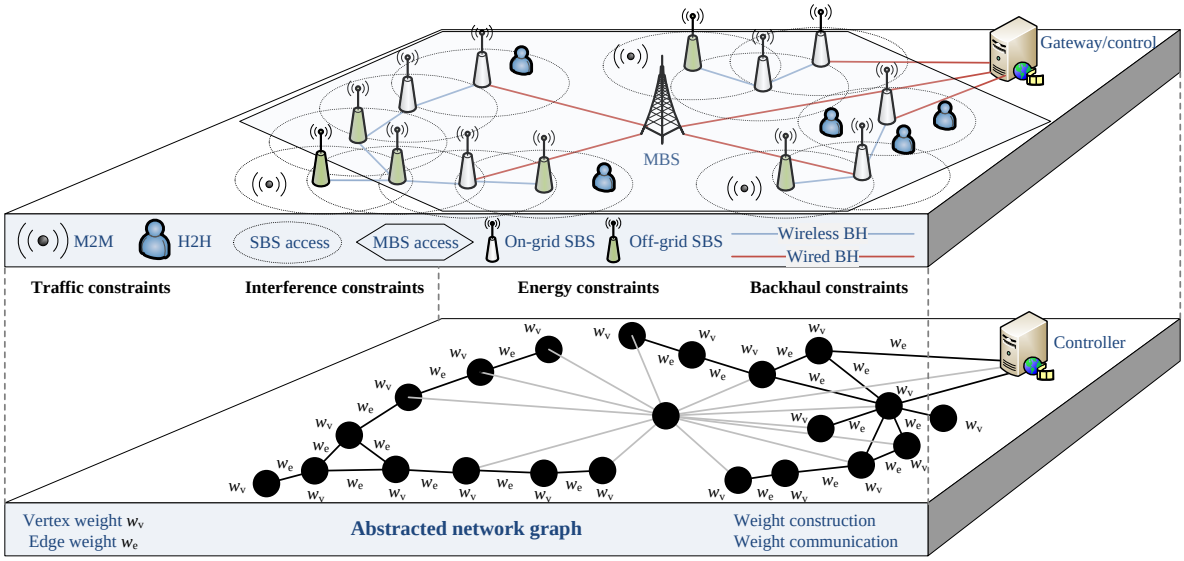


Figure 5-27 Abstracted network graph concept in a HetNet with diversified design constraints

5.6.2.2 Studied solution

Figure 5-28 shows a simplified example of target abstracted energy graph principle. The target SBS multihop backhaul topology is described with a weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ denotes the set of all vertices (BS nodes) and $\mathcal{E}$ is the set of edges (links) between the vertices. Parameter $\mathcal{W} \in \{w_j\}$ represents the set of weights associated with the corresponding vertices. The required control channel can be placed in a separate control plane communicated via MBS or alternatively embedded in the SBS backhaul data links. An edge between two vertices exists if they are within a communication range that supports the target quality of service requirements. If a direct edge does not exist, multihop edges must be used where SBSs act additionally as relay stations.

The SBSs can be either wired on-grid nodes charging energy from a wired electric socket or wireless off-grid nodes charged by a local renewable energy source which can harvest energy provided by sun, wind, etc. To enable a cost-efficient high-density mesh network deployment, only a subset of the SBSs acts as a direct Internet gateway node in a given area. The vertex weights should describe the capability of HetNet nodes to support a target traffic demand for a target time period, given that the available energy of off-grid nodes is limited. Compared to traditional cellular networks, where all BSs are on-grid, this setup leads to an untypical situation where the service availability of BS can vary based on the energy availability. Fortunately, this disruption can be overcome with sophisticated RRM algorithms as will be shown later.

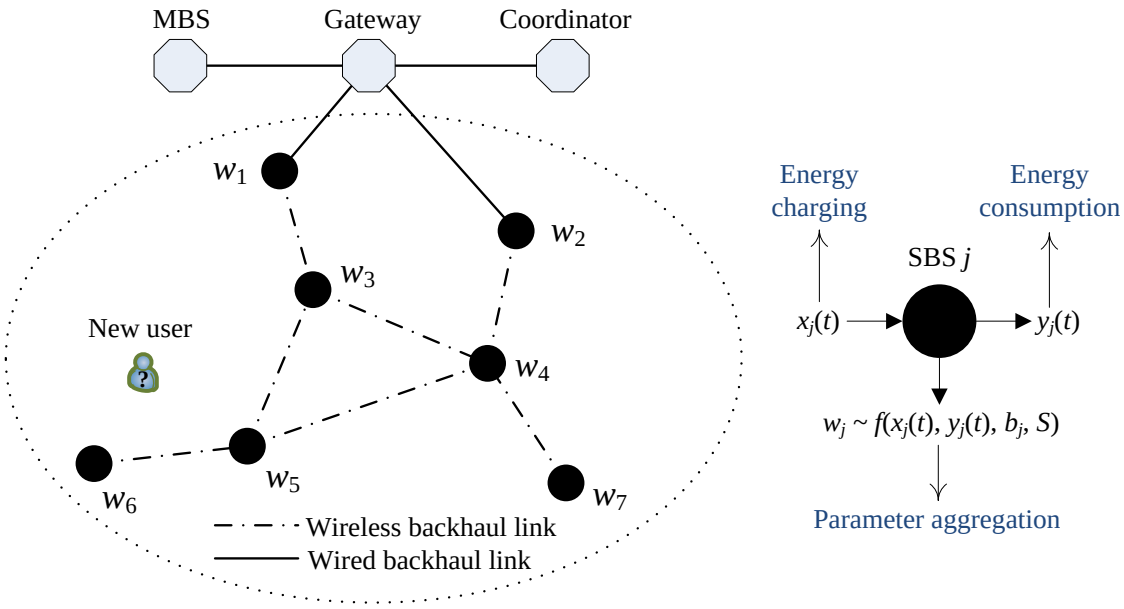


Figure 5-28 Illustration of probabilistic energy graph of a centrally coordinated energy-limited SBS access tier. Energy consumption is heavily affected by the operation of the PHY layer.

5.6.2.3 Abstraction principle and control protocol

The abstraction principle at hand is based on a probabilistic measure of energy status of each SBS. The key idea of the applied abstraction principle is to use a queue-theoretic approach to predict the prevailing status of EDP. The approach allows a high degree of aggregation of different parameters that affect the energy status of an SBS. Furthermore, being a moment-based long term measure, it requires sufficiently low amount of control overhead. Therefore, it is appealing for the use of a central coordination framework.

Each of the energy buffers embedded in SBSs are modelled as energy queueing systems where the random arrival and departure of a single energy charge unit depends on the energy charging process $x_j(t)$ and energy consumption process $y_j(t)$, respectively. Specifically, the number of energy units in the *j*th SBS at the *t* is represented as

$$q_j(t) = \min\{\max\{b_j + x_j(t) - y_j(t), 0\}, B_j\}$$

where b_j is the initial number of energy units, and B_j is the maximum energy level of the energy buffer. An energy depletion event is declared if $q_j(t)$ becomes zero. Provided that B_j is large enough, its effect on energy depletion probability becomes negligible. Given the energy buffer status at $t = t_0$, the main objective is to be able to predict the energy state at time $t = t_0 + S$ where S is the target prediction horizon in time units.

In general, there are a number of alternative system-specific sources which affect the stochastic behaviour of energy arrivals and departures. Instead of studying very specific energy models separately, it is more insightful to understand the relative behaviour of the energy arrival and departure processes. Therefore, without loss of generality, we apply a G/G/1 queueing model where the distributions of the energy interarrival and interdeparture times can be general, and do not need to be known by the system nodes. Yet, the model does require the knowledge of their first- and second-order statistical moments which can be measured.

Using above modelling approach, the EDP at *j*th SBS can be obtained using an integral-based method (cf. [179][180])

$$P_j = \int_0^S \frac{b_j}{\sqrt{2\pi\alpha_j t^3}} \exp\left[-\frac{(\beta_j t + b_j)^2}{2\alpha_j t}\right] dt$$

where $\beta_j = m_{x_j}^{-1} - m_{y_j}^{-1}$ denotes the energy rate difference of random energy arrival and departure processes, $\alpha_j = \sigma_{x_j}^2 m_{x_j}^{-3} + \sigma_{y_j}^2 m_{y_j}^{-3}$ denotes the total variance of the queue process, and m_{x_j} and m_{y_j} ($\sigma_{x_j}^2, \sigma_{y_j}^2$) are the means (variances) of the energy interarrival and interdeparture times, respectively. For accurate results, the above integration requires many integrand evaluations. To simplify the calculation in a closed form for online processing, it is shown in [180] that the EDP with finite prediction horizon *S* can be well-approximated by an upper bound

$$P_j \leq \omega(\beta_j) \exp\left(-\frac{2\beta_j b_j + b_j^2}{2\alpha_j S}\right)$$

where $\omega(\beta_j)$ is a coefficient given in [180]. If we further relax the abstraction approach, so that the prediction horizon *S* is set to be infinite, the asymptotic approximation gives

$$P_j \leq \exp\left(-\frac{|\beta_j| b_j + \beta_j b_j}{\alpha_j}\right)$$

providing even a simpler equation but at the cost of losing accuracy for finite prediction horizons. In above equations, it can be seen that the abstraction approach conveniently combines a number of different parameters, which affect the energy status of SBS, into a single EDP parameter which can used to provision the grade of service via sophisticated RRM protocol to be discussed next.

The applied probabilistic RRM protocol is described as follows. A new randomly arriving user is requesting a data service from the closest SBS. The random state of the SBS tier is defined by the vertex weights $w_j = P_j$ denoting the predicted EDP of the *j*th SBS during a given prediction horizon. The central coordinator selects a route that minimizes the maximum EDP among the involved SBSs. The admission for a new user is granted if the measured EDP is below a target threshold which sets the grade of service in terms of connection dropping probability. The centralized control approach is summarized in Figure 5-29.

1: **if** an admission request arrives to an SBS **then**
2: **for** *j* = 1,2,…**do** *(Loop for the jth candidate route)*
3: **for** *i* = 1,2,… **do** *(Loop for the ith SBS in the jth route)*
4: Measure vertex weights (EDPs) at each SBS
5: Send EDP via control plane or SBS backhaul to coordinator
6: **end for**
7: **end for**
8: Coordinator balances load by selecting the optimal path that minimizes maximum EDP
9: **if** measured EDP is less than set EDP threshold **then**
10: Coordinator informs the SBS to admit new user
11: **else**
12: Coordinator informs the SBS to block the new user and connect via MBS
13: **end if**
14: **end if**

Figure 5-29 Control procedure for centralized coordination of new user request leading to optimized load balance with respect to applied probabilistic energy graph principle

5.6.2.4 Evaluation

The presented energy graph –based RRM approach is evaluated in this section. We first demonstrate the accuracy of the analytical abstraction method and compare it to corresponding Monte Carlo

simulations. Then, we apply the abstraction approach to the target control application and compare the centralized approach with a fully distributed approach regarding the grade of service performance in terms of connection blocking probability and dropping probability. Note that service *availability* is proportional to (1 - blocking probability) whereas service *reliability* is proportional to (1 - dropping probability).

5.6.2.4.1 Abstraction accuracy

In order to demonstrate the accuracy of the energy graph approach, we assume that the energy interarrival times associated with $x_j(t)$ follow a discrete probability mass function with two-state harvesting source. The high-energy state represents direct sunlight or turbulent wind whereas the low-energy state represents indirect sunlight or light wind. The state is randomly selected reflecting the environmental condition of a particular SBS. The energy consumption process $y_j(t)$ is characterized by converting the random traffic arrivals into random energy consumption events, each consuming one energy unit. The interdeparture times are assumed to follow Weibull distribution which is characterized specific shape and scale parameters that are specified by a particular energy consumption process using required transmission and processing energies [180]. If energy consumption is much larger than energy given by harvesting or vice versa, the problem becomes trivial. The most interesting case from efficient resource usage point of view is when the mean energy rate difference of arrival and departure processes is in the similar order (i.e., β_j is close to zero). Then the effect of variability or second order statistics of the random processes really kicks in.

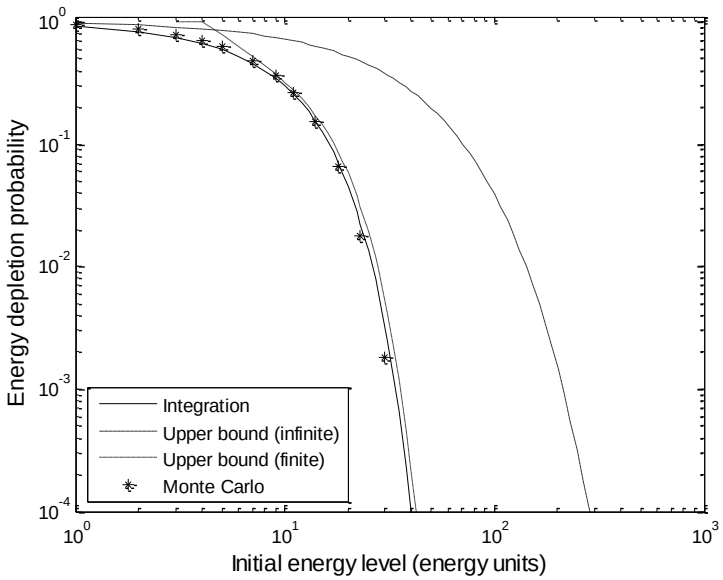


Figure 5-30 Comparison of accuracy of energy depletion abstraction methods with integral-based approach and upper bound

In Figure 5-30, the accuracy aforementioned three EDP methods are demonstrated with comparison to Monte Carlo simulations for S = 100. While the accuracy of the infinite-horizon upper bound is quite loose, the finite-horizon bound can approximate the EDP rather well. The proposed analytical upper bound can also follow the simulation results quite closely while providing simpler way to calculate the EDP online as it is integral-free method. We remark that underestimation of the EDP, which may happen with the imperfect integration, is much more harmful for than overestimation of the EDP because the necessary condition for the connection admission is to ensure that the EDP is below a target threshold. Obviously, underestimation of the EDP from the integral- based approach is avoided by using the upper bound.

5.6.2.4.2 Control application

The main objective in this section is to demonstrate the benefit of having a centralized RRM approach over a distributed approach with the studied energy-limited HetNet concept. In the distributed approach,

all the required RRM decisions are made based on local observed EDP value without sharing the EDP values of involved SBSs. The obvious strength of a centralized approach is to have a global view of the energy status of the network for making provisioned RRM decisions that allow adjustments regarding the grade of service parameters, such as connection dropping probability. This is achieved at the cost of increasing the control overhead. However, since the control information needs to be updated only when the statistical moments varies, the overhead is not extensive if the moments vary slowly.

In Figure 5-31, the dropping and blocking probabilities of the centralized and distributed approaches are compared for the example network configuration shown in Figure 5-28. The results are shown as function of mean energy rate difference of arrival and departure processes $E(\beta_j)$ for the target EDP threshold of 0.05. It is seen that unlike the distributed approach, the centralized approach is able to guarantee that the dropping probability is below a target threshold by balancing the load between SBSs. This is achieved by blocking SBS connection attempts with insufficient energy capacity and possibly steering that traffic towards MBS.

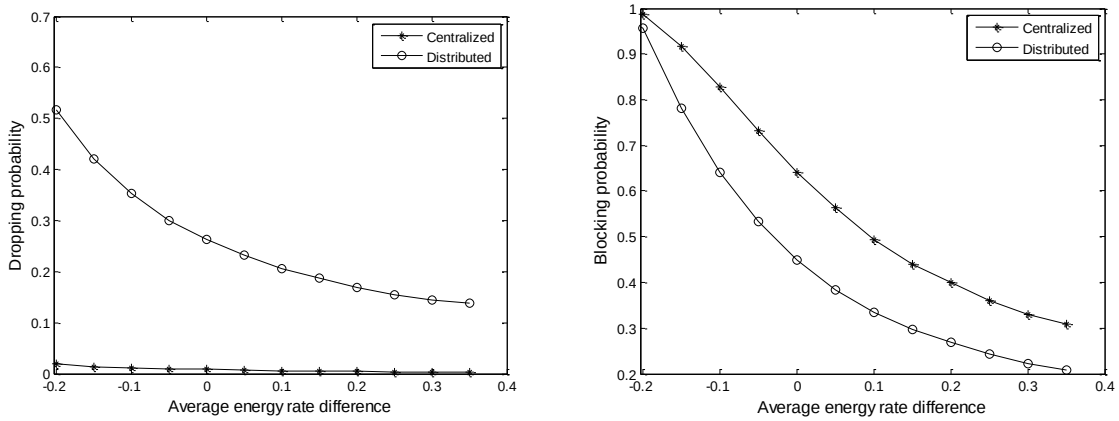


Figure 5-31 Comparison of energy-limited SBS connection dropping and blocking probabilities of centralized and distributed approaches for different average energy rate difference $E(\beta_j)$ of random energy arrival and departure processes

5.6.2.5 Summary

In this study, abstraction of energy status of small base stations harnessed with specific energy charging capability from the surrounding environment is investigated. In this foreseen energy-limited scenario of future dense 5G HetNets, energy sustainability, i.e., the prevention of dropping users during communication, is emphasized. It is known that most energy of base stations is typically consumed by lower layers. Maximizing the use of off-grid energy indirectly affects the energy efficiency of the network while reducing the deployment complexity as it becomes a possibility to use completely wireless base stations. Obviously, this has some distinctive characteristics compared to on-grid wired base stations where energy is not limited.

The used abstraction approach, which aggregates a number of different parameters into a single variable, is used as a predictive method to provision the probability that users experience connection dropping during communication. The main benefit of the approach is that it allows the evaluation of transient effects which lead to more accurate quality provisioning than traditional steady-state approaches. The most interesting case from efficient resource usage point of view is when the mean energy rate difference of arrival and departure processes is in a similar order of magnitude. Then the effect of variability or second order statistics of the random processes really kicks in, and must be considered as is done in the applied abstraction approach. The approach leads to inherent load adaptation between SBSs and, in case of infeasible solution for user admission to any of the SBSs, traffic steering towards MBS to avoid droppings due to energy depletions. It implies the known fact that load balancing does not necessarily mean equalization of the loads of base stations but adapting the relative loads based on the target criterion. In such scenarios, it is therefore important to recognize renewable energy utilization as one of the metrics for balancing the load of base stations.

This study is based on our recent manuscript [180]. In this deliverable, we have further extended the results by comparing the centralized control approach with the corresponding distributed approach to illustrate the potential benefit that could arise from that of centralized control. To increase the reliability of the networks, it is important to make decisions proactively and be able to predict the consequences of the RRM decisions. It is obvious that the applied predictive energy graph principle allows such a capability but only gathers a specific characteristic of future dense HetNets. Therefore, in future work, the energy graph concept is to be combined e.g. with interference network graphs. More detailed analysis is required regarding the control protocol, effects of various system nonidealities, different energy models, and how the current cellular standards would need to be configured to support such RRM framework.

5.7 Cognitive radio in HetNets

Cognitive technologies tend to be a significant tool for wireless and mobile systems. They are applied to sense environmental parameters and reconfigure the network parameters to improve throughput, coverage and QoS provision. For the interest of COHERENT, there are two modes in which these cognitive networks (CN) can operate and provide benefits. The first one is where secondary users (SU), are trying to access the spectrum of cognitive networks which is occupied mostly by primary users (PU). These PUs are the ones who primarily pay for accessing the spectrum of CNs, while the SUs are dynamically searching for spectrum availability. The CN adapts dynamically in order to provide with the best network capabilities for both users. The second case is where secondary small cells, (femtocell, picocell), networks dynamically share the licensed spectrum of primary networks. Usually the secondary users or networks have lower priority in accessing the available spectrum than the PUs or PNs. These are shown in Figure 5-32.

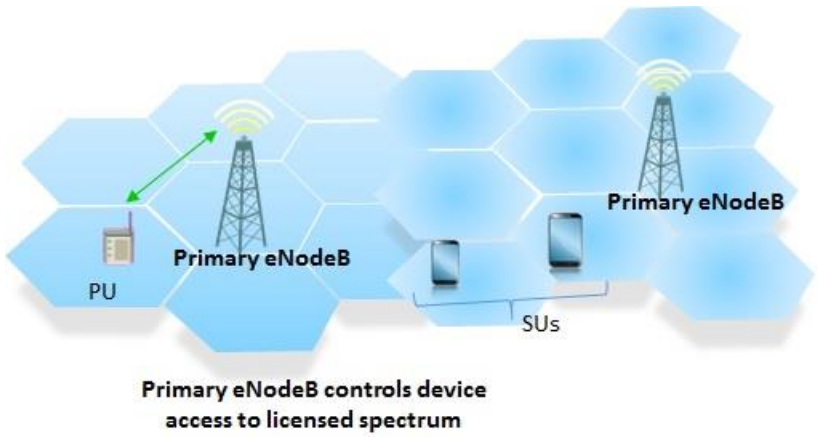


Figure 5-32 Access approach of a secondary user to available channels of an unused spectrum

SU may use machine learning techniques to predict the channel vacancies. Spectrum sensing is based on learning and prediction techniques where a learning model is provided by exploiting historical data, then a new state is observed and from these the next state of spectrum availability can be predicted. The secondary user will be able to sense the spectrum according to the channel status kept in a database and to the channel status which is detected. This work is further to be analysed in deliverable D3.2 of COHERENT project. The network graphs for the CNs consist of all the network parameters which make up the CNs such as spectrum, interference, power, sensitivity, etc. which will be studied further in the upcoming work.

5.8 Device to device communications

5.8.1 Problem description

Device-to-Device Communication is going to be a key enabler for future 5G heterogeneous network. It provides a new kind of connectivity, which can help to boost the network capability, reduce the end-to-end delay for local sharing applications and reduce the device energy consumption. However, the deployment of D2D communication will change the design logic of the traditional cellular network. The interference management, resource allocation and power control should be carefully considered for

this new communication paradigm. COHERENT D2.1 [15] has defined a use case call UC3.CE which is directly related to D2D communication and a set of requirements which are needed to be fulfilled by the COHERENT architecture. In this section, we propose to use the local information drawn from the network graph and interference graph under the framework of COHERENT control system to address the cell association, resource allocation and D2D relaying management problem. Traditionally, the D2D resource allocation and interference coordination problems are considered only inside one cell, either using an underlay or overlay scheme. However, in a dense HMN system, the D2D communication, D2D-relaying can happen across multiple cells, and each network node will see multiple co-channel interference victims and interference generators. Moreover, the interference due to common uplink and downlink transmission to the device to device communication and vice versa should also be taken into account. As a result of this complex interference interaction, the cell boundaries become vague compared to traditional cellular systems for D2D. To tackle the resource allocation and interference coordination for such a complicated network, graph-based control scheme is a natural way. Graph theory is an efficient tool for modelling various types of interactions in cellular network, and it has been widely used to solve the D2D resource management problems in a single cell [126, 127, 128, 129, 130]. However, no modelling and analysis for multi-cell scenarios is presented in these previous works.

The management of the D2D network in HMN includes D2D discovery, cell association, D2D relay selection, resource allocation and interference coordination etc. Before D2D transmission, BS discovery and device discovery are necessary procedures. BS discovery can be done using the traditional measurements as described in current LTE standard. In 3GPP, two types of D2D discovery are provided: Model A ("I am here") and Model B ("Who is there" /"Are you there"). The model of D2D discovery is standardized in 3GPP Rel-12 in which each involved UE is categorized as announcing UE and monitoring UE. The monitoring UE will monitor the discovery signaling sent from the announcing UE. After successful monitoring, the monitoring UE will know the announcing UE is in proximity. Further, 3GPP defines two types of D2D discovery: open discovery and restricted discovery depends on whether extra permission is required to do the D2D discovery or not. For the restricted discovery, the discovery shall be permitted in application layer and use the restricted code in the discovery signal transmission. The Model A can be open or restricted; however, the Model B can be only restricted. Two things that are worth to note are that the D2D discovery is not necessarily coupled with the following D2D communication and that the D2D discovery operation can be done in different carrier frequencies via requesting the transmission/reception gap for inter-frequency D2D discovery. The inter-frequency D2D discovery resource can be allocated by original frequency, common resource or dedicated resource allocation by the concerned frequency. For the public safety (PS), the D2D discovery shall be restricted in either model (model A or B) and can be used in following use cases: (i) Group member discovery, (ii) UE-to-network relay discovery and (iii) Relay discovery addition information transmission. The resource for PS related discovery can be different from the non-PS related discovery via common allocation or dedicated allocation.

On top of direct D2D communication, 3GPP has defined a UE relaying feature which is used for public safety scenarios [132]. This feature is applicable in situations such as UE-to-Network relaying or UE-to-UE relaying. The UE-to-network relay acts as the relay node between the out-of-coverage remote UE and the radio access network. It shall relay the unicast UL/DL traffic via one-to-one D2D communication link. For the uplink traffic, the relay applies its UL traffic flow template (TFT) for the packet prioritization which may be independent of the packet prioritization applied for the D2D link between the relay node and the remote UE. The downlink traffic priority is mapped from QoS class identifier (QCI) of the EPS bearer provisioned by the relay. Moreover, the relay node can also relay the downlink multicasting traffic (e.g., MBMS) via one-to-many D2D communication link based on the request from the remote UE. Further, the UE-to-UE relay is still under standardization which can be applied for the group communication of mission critical push to talk (MCPTT) application. This requires each involved MCPTT UEs to have knowledge about the nearby UEs for the signal routing from the source UE to the target UE.

5.8.2 COHERENT system model for D2D network

As described in COHERENT D2.2 [140], most of management for D2D network (e.g. interference coordination, D2D relay selection etc.) are typically real-time, and these management tasks should be

performed by COHERENT local real-time controller (RTC). RTC is responsible for the construction of local D2D network graph. In addition to RTC, C3 controller is responsible for tasks such as load-balancing by D2D relaying. Our system model is depicted in Figure 5-33. In the D2D network graph, there are two kinds of edges representing the interactions between nodes. For some close device-to-BS edge, we assume that perfect CSI (including fast fading related to mobility) is available at RTC either by timely channel measurements or feedback from devices. For the device-to-device or harmful interference edges, only average CSI (distance-dependent path loss plus shadowing) is available. Please note that both instantaneous and average CSI reporting are currently not supported by the standard (see Section 5.9.2.2), it is a new added feature required by our system. When scheduling a D2D transmission, the information of interference edges is considered when performing resource allocation. The potential interference (interference spillage) is modeled as a function of interference edge weights, traffic load of transmission link. Using this function, we can attach a price for each D2D link, with higher price meaning larger interference spillage. When doing D2D communication scheduling, only local graph information related to the D2D pair is needed, which reduces the controlling overhead. It should be noted that the depicted locations in Figure 5-33 are not the same as the actual geographic locations, the weight of each edge is a result of lots factors including location, shadow fading etc.

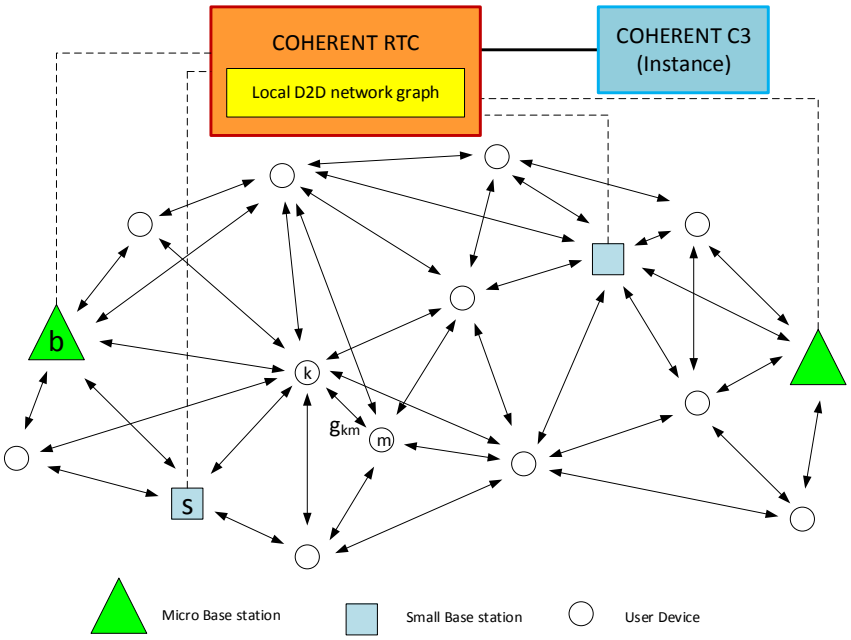


Figure 5-33 Interference graph for D2D-assisted HMN

5.8.3 Abstraction approach for D2D management

In order to abstract the D2D links of various features (communication, discovery, relay) in different purposes (one-to-one, one-to-many, public safety, etc.), the low-level parameters are necessary to be provided and abstracted for the controller, for instance, D2D link functionality (D2D relay, communication, discovery), usage of the D2D link (PS-related, Non-PS related), grouping information (one-to-many, one-to-one), available resource in time and frequency of each user, intended frequency information, destination user list, link quality (sidelink RSRP and sidelink discovery RSRP defined by 3GPP), and UE geo-location. Based on this information, the network graph for various D2D links is formed and used by the higher-layer D2D management application to allocate available resources, carriers, transmission power for the D2D links. These D2D management applications can have different D2D link priorities, adjust D2D link performance, manage D2D interference levels and modify the priority of common downlink and uplink and the sidelink. One shall notice that the impact to the normal downlink or uplink communication shall also be taken into account when forming the D2D links. The impact can be reduced via adjusting the allocated time and frequency resource, the transmitted power on D2D links and by allocating a suitable UE as a D2D relay. After forming the D2D network via D2D management application, the underlying network can be formed via allocating the resource, carrier to each D2D links and the D2D relays. These allocations are done by the D2D network management.

However, the common downlink and uplink information can not only be used as the input to the D2D management applications but also as the output from the applications. The priority setting of different links might be used to prioritize some D2D links (e.g., for public safety purpose) when using the underlay resource allocation scheme or some common uplink and downlink setting are modified in order to support D2D links (e.g., modify the gap setting for inter-frequency D2D discovery transmission and reception). In the current 3GPP standard, the legacy uplink and downlink operation is prioritized over sidelink operation used for D2D transmission and reception.

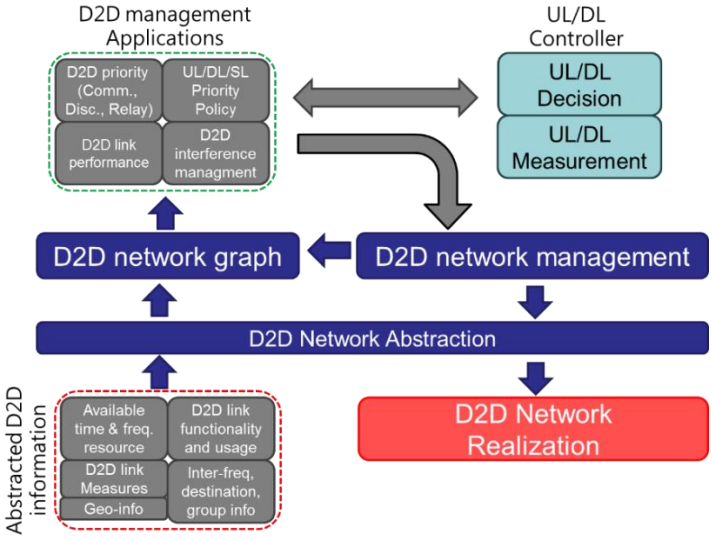


Figure 5-34 D2D management application and underlying D2D network management

5.8.4 Proposed graph-based solution for D2D network management

In this section, we describe how to apply the COHERENT network graph to address the problem of D2D network management. To achieve this goal, we need to consider the construction of D2D network graph based on D2D discovery and channel measurements, graph-based resource allocation for D2D network in HMN and the D2D relaying management problems.

5.8.4.1 D2D discovery and cell association

In [131], D2D measurement and discovery is done by measuring the preamble signal in PRACH, other researchers also propose to use dedicated D2D discovery beacon. In 3GPP Rel-13, it is proposed to apply periodic dedicated D2D discovery beacon for D2D discovery. The periodicity of D2D discovery beacon is called the discovery period and it ranges from 16ms up to 10.24s. The sub-frame allocation of the beacon is also provided via the sub-frame bitmap within the discovery period. Using these kinds of measurements over D2D links and also the legacy downlink and uplink links measurements, RTC can construct the network graph including D2D edges. Both the device and the RTC are aware of the local graph information. By collecting all the graph information coming from each RTC, the C3 can possess a central view of the graph.

5.8.4.2 Resource allocation for D2D network

Traditional D2D communication in cellular network has two resource allocation schemes: overlay D2D (orthogonal spectrum between D2D and cellular UEs) and underlay D2D (non-orthogonal). However, in the multiple-cell HMN, where interferences occur among multiple D2D links and multiple BSs, the simplified division by underlay and overlay cannot capture the complicated interference interactions. We propose that the resource allocation for D2D-enabling HMN should be addressed in a unified framework by using the graph-based scheme. By using this graph-based method, whether two links can use non-orthogonal resource or not depends on QoS constraints on the multiple-cell network scale. For such a HMN, we can define a resource pool which consists of a list of spectral resources. It is possible to have different component carriers in the resource pool. For some component carriers which use frequency-orthogonal transmission methods (such as SC-FDMA), frequency domain allocation inside

a component carrier is also supported. In LTE-A, the minimum scheduling resource unit is one Physical Resource Block (PRB). For 802.11 networks, only time domain scheduling inside one component is possible. The transmission bandwidth for D2D can be fixed or flexible depending on the capabilities of devices and on the standard specifications. The spectrum resources in the resource pool are labelled in the frequency domain. The capacities (i.e. max Tx power, receiver sensitivity, max Tx/Rx bandwidths, component carriers supported, etc.) of each node (Devices, Micro BS, Small BS) inside the HMN should be considered when doing the graph-based resource scheduling on both time and frequency domains. To make the scheduling easy to be implemented, time domain scheduling happens after frequency domain resource allocation has been done.

Moreover, the allocation of the resource and Rx power shall also consider different communication types (e.g., one-to-one relay, one-to-one non-relay, one-to-many relay, one-to-many non-relay) for better resource utilization and less interference. For the one-to-many communication, the members of a group can share common group context, e.g., a layer-2 group ID, a common group IP multicast address and encrypt information via a group security key that are pre-configured. For the one-to-one relay-related communication, the resource allocation is under the UE-pair basis: the pair of remote UE and relay UE for each D2D communication. Further, some resources are used by the UE to transmit communication in exceptional conditions (e.g., RRC connection reestablishment, too many out-of-sync indicators in a period) or when no normal resources are allocated.

The goal of the graph-based resource allocation algorithms deployed in the RTC or C3 is to meet various performance targets (such as received SINR, link bandwidth, reliability, etc.) for each link, with some links having priority. Figure 5-35 depicts a simple graph-based resource allocation for D2D enabling network by the RTC. In this example, there are two resources available, we assume that the RTC knows the average CSIs for each link, and the Tx power of each D2D device is fixed. The target is to maximize the average SINRs for all links, and with the assumptions mentioned above, the RTC needs to utilize the graph information to decide which resource should be used by each link. The picture illustrates (non-trivial) interference by dashed lines. In this example, because the D2D links *ab* and *ef* are most separated, they can reuse the same resource.

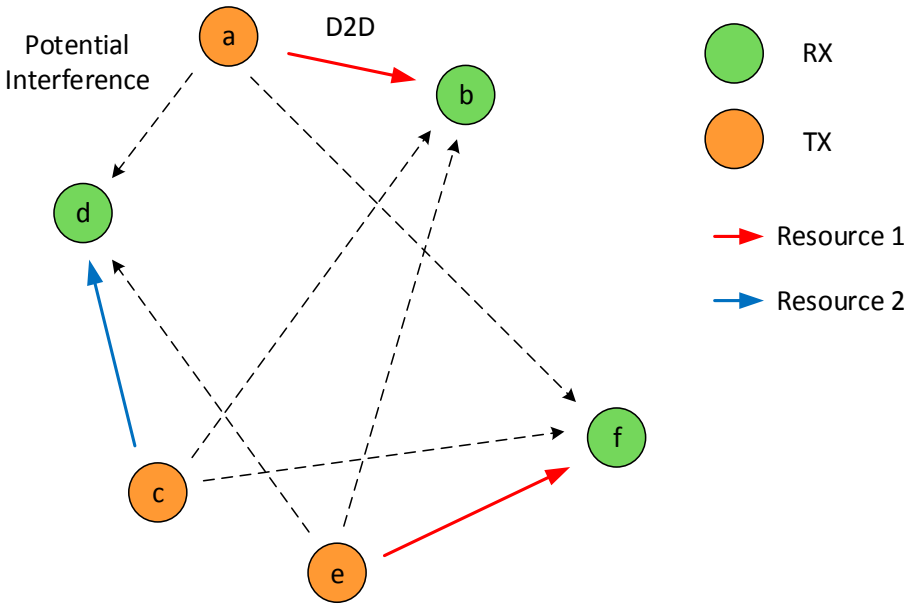


Figure 5-35 A simple example of graph-based resource allocation for D2D

The example presented here is rather trivial. However, in a complicated network with large resource pool and more links, the graph-based method would become very hard since the general graph-colouring problem is NP complete.

5.8.4.3 Relaying management in D2D network

Since most of the user traffic goes to or comes from the network side, D2D relaying is going to be an important feature for D2D-enabling cellular network. It helps to extend cell coverage and increase end-to-end throughput for users which have no good direct link to BSs. To support the relay, both D2D discovery and one-to-one D2D communication functions are needed. The underlying idea is that there is a large number of devices that may act as relays, such as user-deployed devices, nomadic nodes, or mobile stations, which may act as relay stations to help to convey user traffic to or from the network. In those two-hop cases, the access hop is based on D2D communication, whereas the relay hop is a self-backhauling connection based on traditional cellular technologies. To implement D2D relaying, the resource allocation and relay selection should be addressed together. If there are multiple relay candidates, one should select the relay that can maximize the end-to-end performance and at the same time minimize the interference spillage to other links. Using interference coordination method, graph-based channel allocation can be used to mitigate the interference.

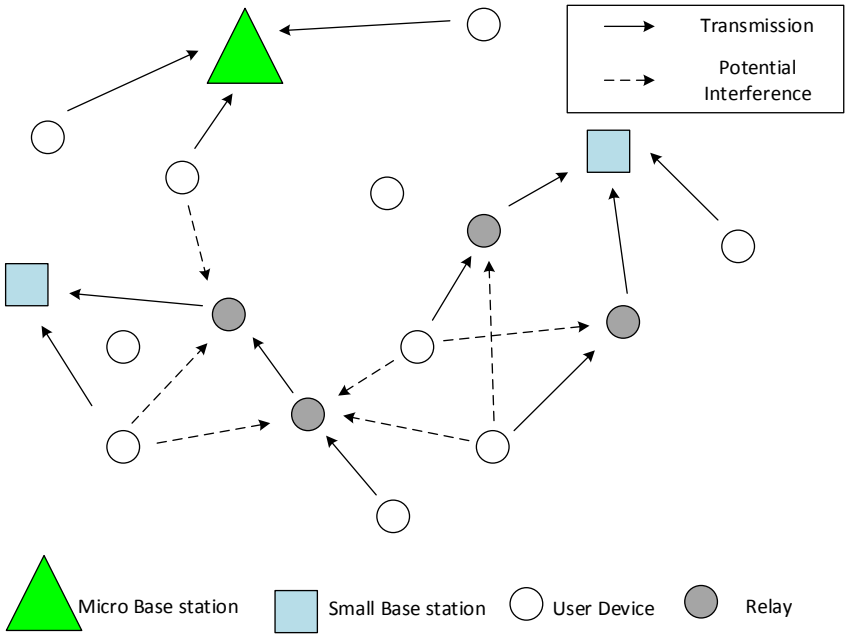


Figure 5-36 HMN with D2D-relaying feature

5.8.5 Evaluation of graph-based interference coordination in D2D network

We study the performance gains achieved by using D2D relaying and graph-based interference coordination in a Manhattan scenario for downlink transmissions. The UE and relay candidate are both vehicles equipped with a sub-6GHz omnidirectional antenna. Vehicles are uniformly distributed along the two-lane streets, and the BSs are placed at corners of the streets with inter site distances (ISD) as 283m. A cell-edge UE can use sub-6GHz D2D relaying depending on its channel qualities. Interference coordination (IC) is performed based on the distributed graph coloring algorithm. We define a neighbor link L_n for a receiver UE_x if (1) L_n transmitter causes interference power larger than $I_{th} = 0.1$ of the desired signal received power for UE_x from relay; (2) the interference caused by the transmitting relay for UE_x to L_n receiver is larger than I_{th} of desired signal received power for L_n. Under above assumptions, the interference graph is asymmetric directed graph. When the links are dense in space, usually we cannot solve all conflicts by simple coloring method, what we can do is to avoid strong interference for poor UEs. An interference conflict happens if a neighbor uses the same link color. In the simulation, an event array is used to record the message exchanges between BSs, relay candidates and destination UEs. Controlling decisions for relay selection, resource allocation and IC are made based on the event array. The event array consists of measurement reports, user/link identifications and interference avoidance requests based on I_{th}. The network graph is constructed by RTC using the information in this event array. Links inside a local cell are colored using random coloring (RC) in initialization of the simulation. After that, the controlling functions are performed in an iterative way. The RTC is responsible for the update of the D2D link colors. The update of the link color is based on

a local proportional fairness utility function involving all the neighbor users. We compare three different network operating methods for the network. The three methods are: (1) without D2D relaying; (2) D2D relaying with RC; (3) D2D relaying with IC. The simulation is conducted using Matlab and we collected the data for 100 runs. The simulation parameters are listed in Table 5-4. The results are shown in Figure 5-37. We define a variable called User Experience Consistence (UEC) as the ratio of the 5% throughput and the mean throughput, to evaluate the cell-edge performance. As shown in Figure 5-37, with D2D relaying and IC, the UEC value is increased almost 70% and 46% compared to (1) without D2D relaying and (2) D2D relaying with RC. The D2D relaying and graph-based interference coordination can help to improve cell-edge UE performance, which shows the advantage of the COHERENT network graph application in the D2D networks.

Table 5-4 Simulation parameters for D2D network

Scenario	Manhattan grid, 1200m × 1200m
Boundary conditions	Wrap-around in XY-direction
ISD	283m
sub-6GHz carrier frequency	5.9GHz
sub-6GHz bandwidth	40MHz
PL model for sub-6GHz	WINNER II B1
LOS probability model	$\left(\min\left(\frac{27}{\mathrm{d}},1\right)\left(1-\exp\left(-\frac{\mathrm{d}}{27}\right)\right)-\exp\left(-\frac{d}{63}\right)\right)^2$
LOS/NLOS correlation distance	10 m
Number of sub-6GHz antennas	BS: one omnidirectional antenna Vehicle UE: one omnidirectional antenna
Maximum TX power	BS: 24 dBm Vehicle UE: 21 dBm
Resource Scheduling	Proportional Fairness (PF)

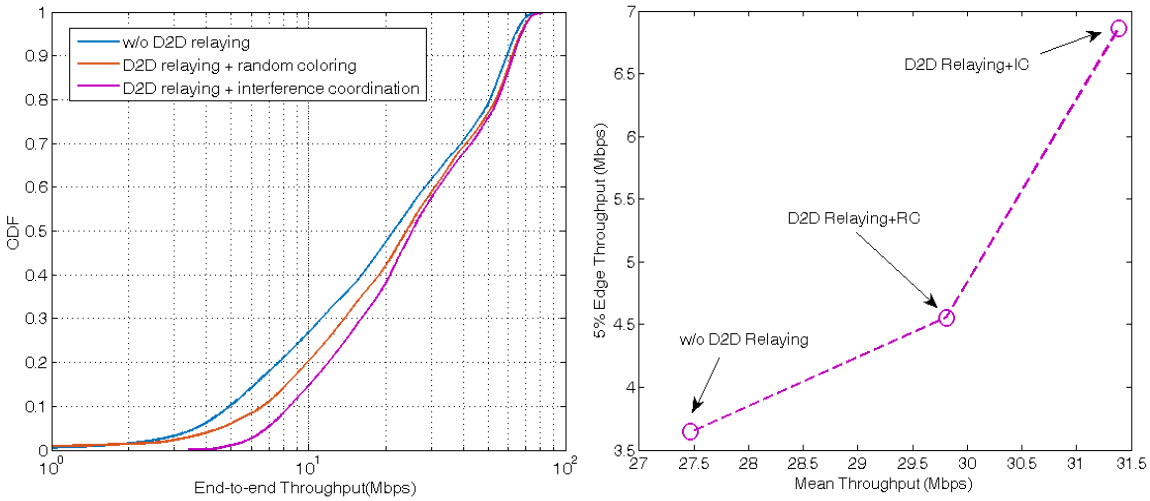


Figure 5-37 Simulation result of graph-based interference coordination for D2D-relaying

5.9 D2D relaying for coverage extension

5.9.1 Problem description

In 2008 the FCC in the USA decided to assign 10 MHz in the 700 MHz band for future broadband PMR. Almost at the same moment the main Public Safety (PS) actors in the USA proposed LTE as the reference technology for the future broadband PMR system. Since then, the idea to use LTE for future broadband PMR systems continued to grow, and today it is widely accepted. Standardization efforts in order to progressively introduce in 3GPP LTE the features required for supporting PMR services started with the creation of the 3GPP SA6 working group on mission critical applications. At the radio access network (RAN) level, the work mainly started with the introduction of initial D2D capability and group call system enablers.

In this work we focus mainly on the coverage extension capabilities provided by devices, meaning handheld or vehicular equipment, in general not fixed, motivated by the use case 3.CE "Coverage extension and out-of-coverage communications" described in COHERENT D2.1 [15]. Briefly, it is fundamental, especially for PS operators, to be able to communicate out of coverage and when possible, to be relayed towards the command center. This situation is illustrated in Figure 5-38, see also Section 3.3.3 in [15].

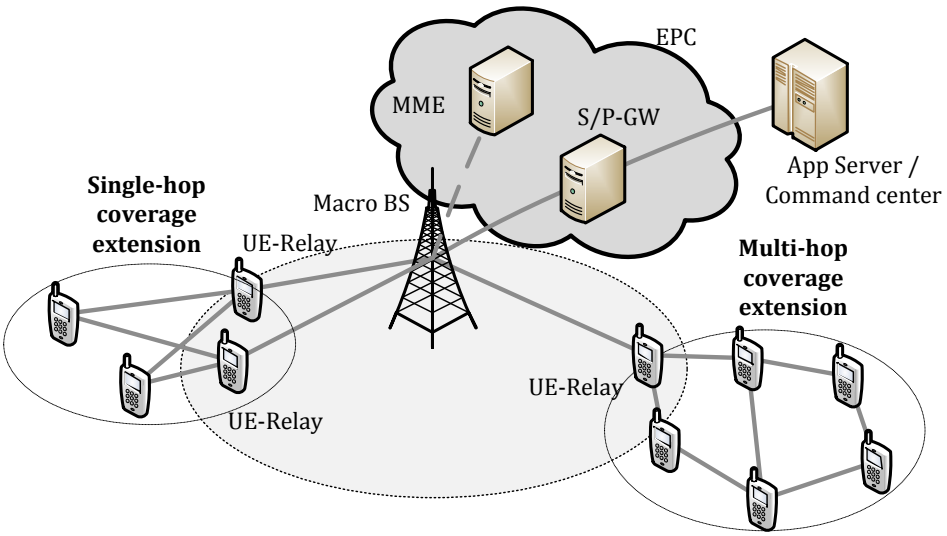


Figure 5-38 Coverage extension with UE-relay

3GPP LTE Rel 12 and Rel 13 have introduced in the standard Proximity Services (ProSe), i.e. services enabled and supported by the capacity of devices to discover each other and to directly communicate, without going through an eNodeB (base station). Currently, LTE Rel 13 supports direct communication between UE in coverage, out of coverage and in partial coverage. The last case corresponds to a communication between a UE in coverage and another UE out of coverage, or in an adjacent cell. Building on top of this capability, LTE has introduced the so called UE-to-network relay, under network control, which is able to extend the coverage of the network (see Section 5.9.2 and Section 5.8).

Notice that in LTE, coverage extension can be implemented also through the use of fixed relays, called Relay Nodes in LTE starting from Rel 10 [155]. However, this specification, which has not evolved since Rel 10, does not match exactly the needs of PS users and will not be considered here. . Note also that Relay Nodes must be established at deployment, this lack of flexibility is matched only with a quite reduced set of operational cases (for instance when there is a hole in the coverage but it is too costly to add a standard eNodeB). Moving relays [156], i.e. moving Relay Nodes to be used in applications like public transportation, is a technology which can also potentially improve coverage. It has been studied since Rel 11, but has finally never been standardized. This technology, too, does not match exactly to the PS user feature of the UE-relay node.

The focus here will be on UE relaying intended in a broader sense than the UE-to-network relay. Multi-hop relaying in out-of-coverage situation will also be studied as shown in Figure 5-38. The issue we deal with is the one of the control of links for coverage extension, inside the control framework proposed by the COHERENT project. Current 3GPP LTE Rel 13 specification does not foresee a control in the interface between two UEs, as detailed in Section 5.9.2. Here we are talking about the high-level control, which allows configuring HARQ and QoS parameters of the communication (LTE Rel 13 already foresee some limited low-layer control). The question is if introducing a certain degree of control can help and to which extent, to satisfy the QoS constraints related to PMR services, and how it can be implemented in case of D2D. Also the kind of relaying or forwarding strategy can have a strong impact on the performance. Finally, which kind of network vision (network graph) can be realistically provided to a central coordinator in these situations of coverage extension, or out-of-coverage communications?

For developing our investigation, the selected general framework of D2D communications is the one developed in 3GPP, even if a big amount of work has been done in general on D2D and in specific situations like wireless sensor networks and ad hoc networks. For instance, reference [157] reviews D2D communications in several areas of applications. As remarked by the authors, there is a fundamental difference between D2D communications overseen or controlled by a central entity (e.g. a BS in cellular networks, a cluster head in mesh networks, etc.) and D2D communications without any centralized control. The difference is about how the control plane is managed. When a centralized control plane exists, many difficult tasks like synchronization of nodes, resource allocation, routing, can be substantially simplified, see also Section 5.8. On the other hand, when all the tasks are implemented in a flat architecture, considerable control overhead circulates on the network and, for instance, decreases available data rate. We selected the 3GPP framework because, as pointed out by [157], too, D2D for PS applications is in big part standardized in 3GPP, which is moreover on track for specifying future 5G systems. Of course, this fact will not prevent us to study options outside the current 3GPP implementations.

The next section presents the current situation about D2D communications in 3GPP, as a practical starting point for seeing how coverage extension can be implemented with this approach and what kind of metrics are available today for monitoring the links and building a network graph.

5.9.2 State-of-the-art of D2D communications in 3GPP

The current simplified architecture for D2D in 3GPP is depicted in Figure 5-39. It is called Proximity Service (ProSe) architecture, the term ProSe being used when talking about D2D communications from the perspective of service. When addressing D2D communication at a lower layer, typically L2, the term SideLink (SL) communication is used in order to distinguish the direction of the communication from the traditional UpLink (UL) and DownLink (DL).

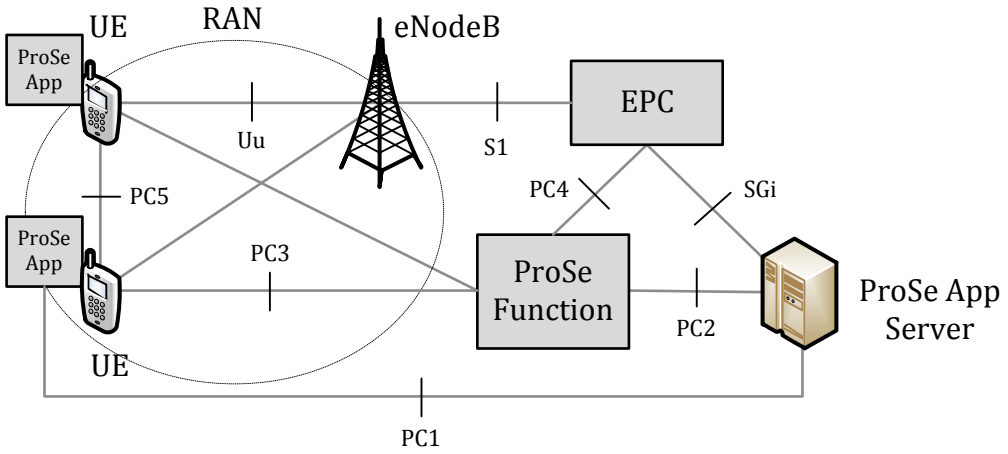


Figure 5-39 Simplified 3GPP architecture for ProSe (from 3GPP)

In 3GPP, UEs with D2D communications capability can discover each other, which does not necessarily imply the exchange of a big amount of data, and, for PS applications, can send and receive data [158]. All these D2D communications happen on the PC5 interface, which has been defined as connection-less, meaning that a UE does not establish and maintain a logical connection to receiving UEs prior to a sidelink communication [85]. This is substantially different from an UL (or DL) connection between a UE and the eNodeB in which a logical connection is maintained between the two entities and the UE can be in RRC_IDLE or RRC_CONNECTED state and can send data only in the latter. Only a limited form of control plane with an RRC protocol associated to the communication on the user plane over PC5 is available, and it is used to send fundamental parameters related to the SL communication [85]. Some form of control in the lower layers also exists, as it will be shown afterwards. Moreover, except for the case of UE-to-network relay, the standard says nothing about the relaying in coverage or out of coverage of the UEs involved in direct communications. In fact, in the current 3GPP vision, multi-hop

relaying could be handled at the IP level (which boils down to routing) and one can imagine that it could be managed at the application level.

Interface PC5 was first defined as a one-to-many interface, in order to stress the multicast and broadcast nature of the interface, motivated by the group dimension of the communications for PS. 3GPP LTE Rel 13 extended also PC5 to the case one-to-one.

An important building block for the ProSe architecture is the ProSe Function entity which has the following main functionalities:

- Direct Provisioning Function: it provides to the UE all the parameters necessary for the configuration of direct discovery and direct communications.
- Direct Discovery Name Management Function: it allocates and processes ProSe Application (App) ID's, which represent the applications that the users want to run and which have communication and/or discovery needs, and ProSe Application Codes, which are assigned by the network after a Discovery Request procedure. This function is responsible also for managing permissions concerning the discovery and ProSe applications running in the UE. It manages also the authorizations for direct discovery and the integrity of the discovery procedure.
- EPC-level Discovery Function: it is responsible of discovery procedures done at EPC level; in particular it deals with the authorization and configuration of the UEs for this kind of procedure.

In the following, the review will be focused mainly on L2 and L1. The aim is to extract the main parameters that can help in the modelling of the low layers of D2D communications and the measurements that could be used for feeding a network graph, or to identify new measurements if needed. For more details on the protocol stacks for D2D in 3GPP LTE, the reader can refer to COHERENT D2.2 [140] as well as [85] and [205].

5.9.2.1 General view on low layers for SL communications

At low levels, 3GPP decided to use UL modulation, SC-FDMA, and subframes both in TDD and FDD configurations for SL communications [163]. The frame and sub-frame definition is exactly the one of UL and DL. As for UL and DL there are logical Traffic and Broadcast channels (STCH, SBCH) with the corresponding transport and physical channels which follows the usual naming with the UL or DL part replaced by SideLink, see Table 5-5. The STCH is mapped on the corresponding transport and physical channels SL-SCH and PSSCH, which may experience collisions, depending on the allocation of resources done by the eNB or UEs. The SBCCH broadcasts control information, in particular synchronization signals for out-of-coverage UEs or for UEs being in cells different from the one of the UE generating the synchronization signal, in order to enable direct communications between UEs in different cells.

As for UL and DL, there is also a Transport and Physical SideLink Control Channel carrying the SL Control Information (SCI). There is a new transport and physical channel for direct discovery: Sidelink Discovery Channel (SL-DCH) and the Physical Sidelink Discovery Channel (PSDCH). Discovery channels are defined only starting from transport level and PHY level because protocol at L2 is directly connected with application layers which give the messages to be sent.

Table 5-5 Mapping of channels for Sidelink in 3GPP LTE

Logical channels	SL Traffic Channel (STCH)	SL Broadcast Channel (SBCH)		
Transport channels	SL-Shared Channel (SL-SCH)	SL-Broadcast Channel (SL-BCH)		SL-Discovery Channel (SL-DCH)
Physical channels	Physical SL Shared Channel (PSSCH)	Physical SL Broadcast Channel (PSBCH)	Physical SL Control Channel (PSCCH)	Physical SL Discovery Channel (PSDCH)
Use	Traffic channel used to send data	Control channel used for sending the MIB-SL	Control channel at PHY for sending SCI for PSSCH	Discovery messages

The structure of the MAC for Sidelink in 3GPP Rel 13 is given in Fig. 4.2.1-3 in TS 36.321 [167] and reported below for reference.

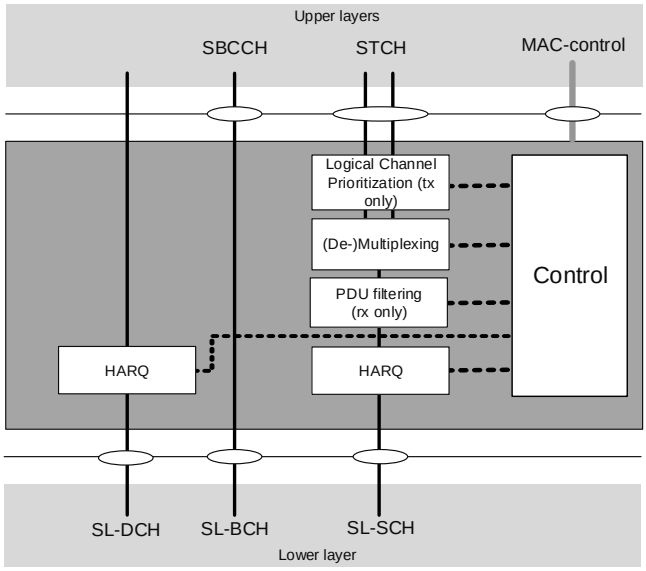


Figure 5-40 MAC structure for SL communications (Fig. 4.2.1-3 in TS 36.321 [167]])

In the upper part of the MAC, Logical Channels are organized in Logical Channel Group (LCG) depending on the priority of the sidelink logical channel and the mapping between LCG ID and priority which is provided by upper layers in logicalChGroupInfoList. Each LCG (and hence each LC) is defined per destination.

Current MAC structure for SL is relatively simple with respect to the standard LTE MAC structure for UL and DL. This fact implies also that there are no complex scheduling algorithms inside the UE. Evolved scheduling algorithms in the UE require evolutions of the MAC structure, or they could be implemented at application level. Scheduling algorithms for SL are possibly implemented in the eNB for scheduling SL communications under network control (in-coverage), but this is out of the scope of the 3GPP standard.

As can be seen from Figure 5-40, SL MAC has traditional functionalities, like channel mapping, multiplexing and demultiplexing of service data units (SDU), hybrid automatic repeat request (HARQ) management, transport format and radio resource selection. However there are functionalities specific to SL like:

- LC prioritization (tx): the standard specifies a particular policy of prioritization. The MAC, when filling the resources, shall serve first the destination having LC associated to the highest priority (with traffic to send) and not served before. For that destination, LC with highest

priority shall be sent and if there is room, the resources are used to send packets from LC of decreasing priority. LC with equal priority shall be equally served.
- Protocol data unit (PDU) filtering: these filters are configured by higher layers and allow ignoring packets related to applications that the UE is not allowed to receive.

In Rel 13 at MAC level each packet contains MAC-SL header carrying:

- A version field (4 bits) specifying the type of destination identifier (group-cast or unicast).
- 4 reserved bits for future use.
- The identifier of the source (ProSe UE ID): this uniquely identifies the sender
- The 16 most significant bits of the identifier of the destination (ProSe L2 Group ID) which can be interpreted as unicast or group-cast. Note that the 8 remaining least significant bits of the destination ID are sent in the SCI (control information) in the PSCCH.

These IDs are provided by the network or preconfigured in the devices (for out-of-coverage operation for instance), thanks to network and security procedures. The 8 Least Significant Bits (LSB) of the ProSe L2 Group ID are used in the PSCCH for filtering data packets at the reception. For PDU(s) associated with one SCI, MAC shall consider only logical channels with the same Source Layer-2 ID-Destination Layer-2 ID pair.

MAC header information is minimal: source, destination and if the communication is to be considered unicast or multicast. Nevertheless, some bits are reserved for future use and can be used for evolutions.

Concerning PHY layer, SL communications can use either normal or extended CP. The resource block definition is the same of the one of DL and UL PHY channels 12 subcarriers and 7 or 6 time symbols depending respectively on the normal or extended prefix. SL communications uses exactly the same bandwidth of UL communications in coverage, which means the same number of resource blocks. However, in out of coverage, where there is access to network configurations, the number of resource blocks is preconfigured to {6, 15, 25, 50, 75, 100} [166]. Even if the definition of resource block is the same, the total theoretical rate achievable by SL communications is smaller than the UL one because the last SC-FDMA symbol is left unused and acts as a guard band.

For SL operation only one spatial layer is allowed in 3GPP LTE Rel 13.

Table 5-6 illustrates the modulation and coding parameters allowed for SL communication in Rel 13 [163], [164].

Table 5-6 Modulations and coding parameters of SL PHY channels

Physical channels	Physical SL Shared Channel (PSSCH)	Physical SL Broadcast Channel (PSBCH)	Physical SL Control Channel (PSCCH)	Physical SL Discovery Channel (PSDCH)
Modulations	QPSK, 16-QAM	QPSK	QPSK	QPSK
Code	Turbo code	Convolutional	Convolutional	Turbo code
Rate Matching	Yes	Yes	Yes	Yes
Interleaving	PUSCH	PUSCH	PUSCH	PUSCH

For PSSCH and PSDCH, the processing chain is like the one of the PDSCH. However, for all channels, interleaving is done as in Physical Uplink Shared CHannel (PUSCH), but without any control information multiplexed in the subframe resources. This interleaving spreads symbols first in time in order to send contiguous coded bits in separated OFDM-symbols, which will increase the robustness of the communication especially if frequency hopping is applied.

5.9.2.2 SL PHY measurements

Two main measurements are defined at PHY layer for SL communications, i.e. on the PC5 interface [166]:

- Sidelink Reference Signal Received Power (S-RSRP): it is the linear average over the power contributions in [W] of the resource elements that carry demodulation reference signals associated with PSBCH, within the central 6 PRBs of the applicable subframes.
- Sidelink Discovery Reference Signal Received Power (SD-RSRP): the linear average over the power contributions in [W] of the resource elements that carry demodulation reference signals associated with PSDCH for which CRC has been validated.

The previous measurements can be configured by the network through the configuration of the SL-BCH and the SL-DCH, hence the frequency of those measurement is 40 ms for S-RSRP and according to the periodicity of the discovery messages for the SD-RSRP. Those measurements are in particular used when the network control in-coverage SL communications. Those measurements are also available at the UE when out of coverage. In the standard, out of coverage is defined with respect to a carrier. When the UE does not find any cell of any operator for all the available RATs, it is said to be in Any Cell Selection state, which is a strong version of "out of coverage".

Nothing is explicitly written about measurements on the traffic channel, this is because currently in 3GPP fine link adaptation is not managed even for in coverage direct communication.

The network uses also standard RSRP measurements between the UE and eNB on the DL in order to understand if the UE is at cell edge or not. UE-to-network relaying for coverage extension is in fact activated only when the UE is at the cell edge.

Notice that sounding reference signals (SRS) in UL operation are used for sounding UL bandwidth on different resource blocks from the ones the communications has used, thus enabling frequency-selective scheduling. SRSs are not available for SL communications. When the UEs are in coverage, [168] presents a possible new method for exploiting existent SRS: UEs interested in acquiring wideband channel knowledge of a SL channel overhear the SRS sent to the eNB by the target UE source. This method however cannot be directly applied for out-of-coverage UEs, the SRS configuration being done by the eNB, while it can provide more information for channel quality estimation for in-coverage direct communications.

5.9.2.3 SL synchronization and common broadcast messages

3GPP LTE Rel 13 specifies two signals for SL synchronization:

- SL synchronization signals (SLSS): they are organized like DL synchronization signals into a primary sidelink synchronization signal (PSSS) and a secondary sidelink synchronization signal (SSSS). Generating parameters are however different (see [163]).
- Master information block – SL (MIB-SL): additional information for synchronization and configuration.

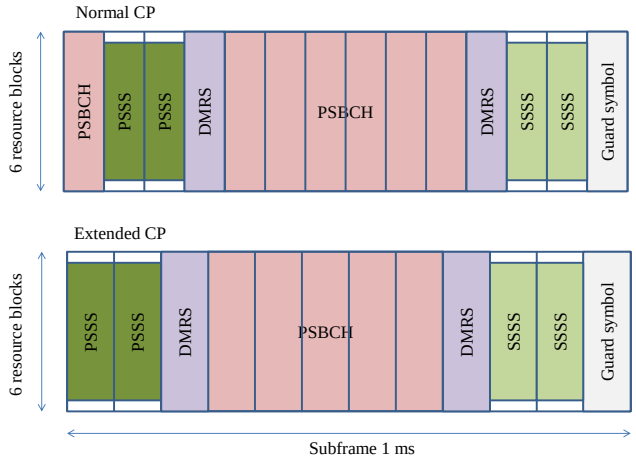


Figure 5-41 SL subframe structure with resources filled with synchronization and broadcast signals

Figure 5-41 illustrates the structure of the subframe for SL synchronization, for normal and extended CP. Demodulation Reference Signals (DMRS) follow the definition and location defined in UL. Like in DL, PSSS and SSSS are sent in two consecutive time slots of the same subframe, 62 subcarriers of the 6 central RBs are used onto two contiguous SC-FDMA symbols. PSSS and SSSS encode a Sidelink ID (SID): SID = 0, …, 167 (similar to cell ID) for in-coverage (transmit UE is in coverage or it gets the synchronization from a UE in coverage) , SID = 168, …, 335 for out-of-coverage (with no connection to a UE in coverage). This method allows keeping track of the situation of the UE, with respect to the synchronization signal.

The same subframe is used also for sending PSBCH. This subframe with synchronization and broadcast information is configured by higher layer and never used for traffic and control or discovery channels.

The MIB-SL information carries sidelink common configuration information:

- a flag for signaling whether the sender is in coverage or not,
- if the sender is in coverage, the MIB-SL carries the system frame number (SFN) and subframe number, otherwise a subframe internal counter for UE out-of-coverage.
- the SL bandwidth,
- the TDD/FDD configuration.

The MIB-SL may change at any transmission i.e., unlikely the MIB from the eNodeB neither a modification period nor a change notification mechanism is used.

Synchronization signals may be sent either under instruction of the network if the UEs are in coverage, or by decision of the UE both for the in-coverage and out-of-coverage cases.

In case of network-instructed synchronization, the UE always sends the SLSS and MIB-SL every 40 ms either if it has no data to send or if it has data to send (MIB-SL has priority higher than data). The UE constructs the MIB-SL content starting from information received in System Information Block (SIB) Type 1, 2 and 18. Note that SIB Type 18 contains also the information of when sending the synchronization signals and PSBCH. Even if in coverage, sometimes the receiving UE must synchronize to the sending UE rather than the network, an example could be the communication in between adjacent cells which are not completely synchronized. 3GPP has specified messaging in order to communicate to the receiver the fact that it must synchronize to the sender.

In case of coverage extension the UE recovers from SIB Type 18 a synchronization configuration element which gives RSRP thresholds with respect to the eNB defining the conditions for the UE to be at cell edge. If the UE satisfies those conditions, it can decide autonomously to send the SLSS if there is no synchronization aided procedure going on. The UE reconstructs the signal to send by looking at the SL-SyncConfig inside SIB Type 18, and sends it only during the resources allocated by the network for the transmission of SL data.

In case of out-of-coverage situations there are two preconfigured subframes each 40 ms, one is used for reception of SSLS and PSBCH and the other is used for their transmission in case the UE becomes a source for synchronization [85]. The following procedure for sending the synchronization signals is used: first the UE checks if it is in coverage, otherwise it searches for SLSSs. The UE performs signal quality measurements (S-RSRP) on the reference signals of the PSBCH sent in the same subframe of the PSSS/SSSS. If the S-RSRP is below a certain threshold it starts sending out the synchronization signal. This is done in order to limit the number of UE sending out the synchronization signal. Moreover, when the UE becomes a synchronization source, it changes the SFN and subframe number (and hence the moment in which it sends the SLSS) by choosing them in a set of preconfigured values, and trying to avoid the values sent by the other UEs. This is done in order to diminish the interference.

While searching for a UE sending the SLSS, the receiving UE scans all the SLSS IDs. If the receiving UE finds a UE, then the sender becomes a synchronization reference UE (SyncRef UE) and the parameters in the MIB-SL are used, i.e. time, frequency, bandwidth, and frame structure. In order to synchronize, the S-RSRP must be over a certain threshold and the MIB-SL must be correctly decoded. If multiple SLSS IDs are selected, the UE prioritizes them according to the situation (in-coverage w.r.t. MIB-SL flag, in-coverage w.r.t. SLSS ID range, out-of-coverage) and inside each class according to the S-RSRP. When a UE receives the SSLS and PSBCH from multiple senders, the SyncRef UE can be updated according to S-RSRP changes measured on the links. The SyncRef UE change is not instructed by the network, but decided by the UE on its own measurements.

5.9.2.4 SL discovery

SL discovery is in part described in Section 5.8.1 and Section 5.8.3. Here we focus on parameters related to lower layers.

In 3GPP LTE Rel 13, discovery is extended to Restricted Discovery (social-type model) but always in-coverage, and multicarrier operation [160]. For Public Safety (PS) users, discovery service authorization may be saved inside the UE device or in the USIM. For non-PS users, network shall authorize the service. Once the authorization and provision is granted, UE wanting to use D2D sends a discovery request with an Application ID, either in monitoring or announcing mode for Model A; for Model B it sends in discoveree or discoverer mode. The ProSe function responds with a Discovery Response with parameters about this request and in particular the ProSe Application Code (which is a sort of grant authorizing the UE to use the corresponding ProSe Application) plus a validity timer for this specific code. This timer controls the capability of the UE to use the related ProSe Application. In case this timer expires, the device needs to request a new code from the network (see [158] for more details).

The UE transmit power is indirectly set by the message describing the announcing policy, and three ranges are possible: short, medium, long. A maximum transmit power level is associated to each range and it is sent in the SIB Type 19. The network indicates also if the UE is authorized to discover other UEs out of network coverage and with which policy.

Discovery resources are configured by the network or could be pre-configured in the device. If configured by the network, the configuration is carried in SIB Type 19 which defines the Resource Pool (RP) and the length of the discovery period during which the resources are valid. Discovery period can be configured with the following values: {4, 6, 7, 8, 12, 14, 16, 24, 28, 32, 64, 128, 256, 512, 1024} in frames of 10 ms, i.e. from 40 ms to 1024 ms [166] and it is defined with respect to the System Frame

Number (SFN) = 0 when in coverage or with respect to the Direct Frame Number[4] (DFN) = 0 when out of coverage, with a configurable offset. Resource pool is defined temporally by a bitmap saying which subframes inside a discovery period can be used and with which repetition, and the PRB to be used for discovery. The bitmap has different lengths (from 4 to 42 bits), and the length to be used depends on the frame configuration. For instance, for FDD, the length of the bitmap is 40 (see Section 6.3.8 in [168]). As shown in Figure 5-42, the configuration parameter numRepetition is the number of times which the subframe bitmap is repeated; it is in between 1 to 5 for FDD configuration (see Section 6.3.8 in [168]). The PRB allocated to the PSDCH have always the form shown in Figure 5-42, two subbands which can be configured as one.

The selection of resources to be used to send the Discovery Message is done according to the state of the UE (RRC_IDLE or RRC_CONNECTED). In the idle state, the UE monitors resources pools indicated by SIB Type 19. In the connected mode the device can request from the network an exact allocation or the network can tell the UE to autonomously select the resource in a resource pool. Moreover, radio resource allocation depends on the type of discovery: Type 1 is non-UE specific, while Type 2B is UE-specific and semi-persistent. The two types of discovery impact on how the resource of PSDCH are identified and used.

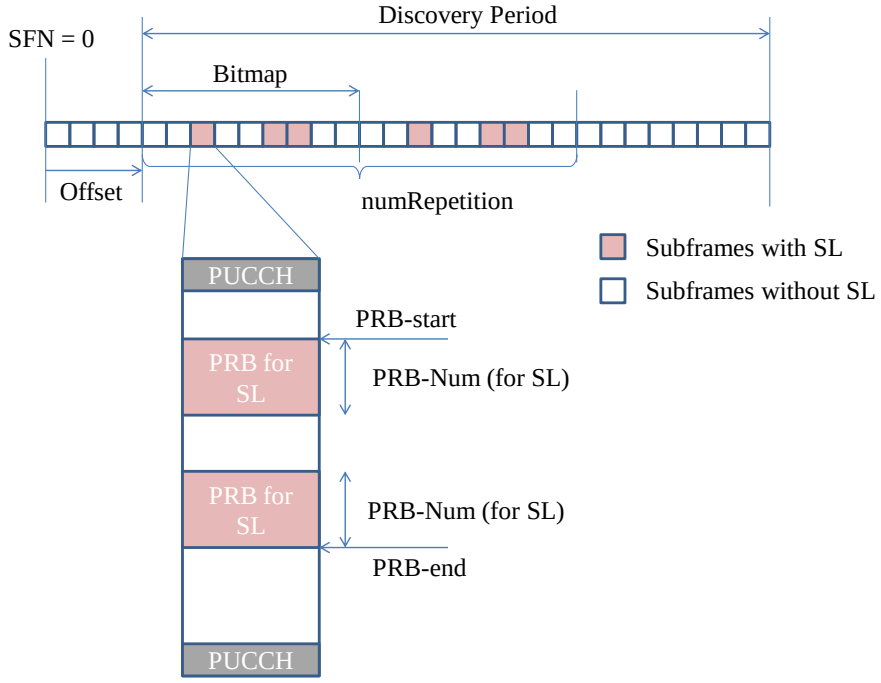


Figure 5-42 Resource pool definition for SL discovery. In the figure, only subframes for UL (eligible to SL) are shown

The direct discovery message sent on PC5 contains the information necessary to identify the ProSe Application the receivers may be interested in, as well as the Prose Application Code which contains also the mobile network identifier (PMLN ID). The discovery message is 232 bits long and contains an integrity check (CRC), and it is sent using QPSK (see Table 5-6) on 2 adjacent PRB per time slot. Hence, in a single transmission, the discovery message is coded with a code rate equal to 0.439 and 0.537 respectively for normal and extended CP. In order to improve robustness, the discovery message is associated to a HARQ process at the MAC level, whose number of retransmissions can be configured inside the RP to 0, 1, 2, 3. At the PHY layer different incremental redundancy versions are used with the fixed order 0, 2, 3, 1. Also, from one repetition to another, the couple of PRBs carrying the message hops inside the RP resources according to a hopping pattern defined in Section 14.3.3 in [165]. This

[4] DFN is a frame number which is established by the UE out of coverage with respect to its internal time reference, since there is no synchronization with the network.

parameter can be configured by RRCReconfigurationMessage, meaning that it can be set by the network, for UEs which are in RRC_CONNECTED state. For UE out-of-coverage that cannot go into that state, the parameter is preconfigured [167].

5.9.2.5 SL communication

In 3GPP direct communication is defined in-coverage, out-of-coverage or for coverage extension, but only for public safety devices. Direct communications uses the STCH with the associated transport and PHY channels, and it uses the PSCCH in order to control link parameters at MAC/PHY levels see Table 5-5. On top of that, on the user plane (see COHERENT D2.2 [140]) over the PC5 interface there is an instance of Radio Link Control (RLC) and Packet Data Convergence Protocol (PDCP) (two L2 sublayers) for each direct communication which is defined by a couple of ProSe UE ID and L2 group ID. RLC/PDCP are instantiated when possible but they are configured only on reception of the first RLC PDU and their lifetime is not specified by the network (usually those instances last the duration of the connection, but in this case there is no connection). RLC is always in Unacknowledged Mode meaning that there is no ARQ. However HARQ at MAC/PHY exists but without feedback, as will be detailed below.

Provisioning is a pre-requisite for communication. A certain number of parameters (PLMN for in-coverage communications, authorization of out-of-coverage communication, radio parameters for out-of-coverage communication, L2 group ID, IP addresses, security parameters for group communications) must be transmitted to the UEs interested in D2D by the ProSe Function or must be pre-configured. Frequency carriers of the system can support or not direct communication, so UEs wanting to use direct communications must search for good carriers. For out-of-coverage, UE uses carriers which are stored in the device or Universal Subscriber Identity Module (USIM). For coverage extension (partial coverage), the in coverage UE (i.e. the UE-to-network relay) can use either pre-configured or network-controlled carriers. Note however that the out-of-coverage UE (the remote UE) will always use preconfigured resource. Hence carrier coordination is necessary to enable communications between an in-coverage UE and out-of-coverage UE and to control interference.

Selection of resources can be done autonomously, by looking at the Resource Pool (RP) in SIB Type 18, or can be scheduled by the network. There are transmission and reception Resource Pools (RP), provided by SIB Type 18 information or preconfigured in the device. Of course, a transmission RP must be equal to a receive RP for enabling communication inside a cell. However, multiple receive RP can be defined for covering communications with out-of-coverage UE or for communication with UEs in adjacent cells. Nothing is said about the coordination of these RP.

Resource allocation is associated to the PSCCH and it is done in 2 modes:

- Mode 1, the eNodeB schedules the resources for direct communication inside the RP given by the SIB Type 18. UEs must be in RRC_CONNECTED mode to use this method.
- Mode 2, the UEs independently choose the resources inside the RP (pre-configured or given by the network) for sending scheduling information (PSCCH) and data (PSSCH). UE can be out-of-coverage or in RRC_IDLE state.

The main structure of the RP for communication is the same as for discovery. The PRB allocated for communication inside an authorized subframe are a pair of subbands as in Figure 5-42. The structure and signaling of the subframes dedicated for control and data transmission in time is different, even if the principle of bitmap for indicated subframes allocated to communication is maintained. An RP is temporally described by an SL Control (SC) period of configurable length {40, 60, 70, 80, 120, 140, 160, 240, 280, 320} subframes, hence ranging from 40 ms to 320 ms [168]. The SC period is divided into the first part dedicated to the transmission of the control (PSCCH) and the second part for data (PSSCH). The SC period is repeated thus providing a periodic structure, a sort of frame, for SL communications. The repetitions of the SC period are limited and their end must fall before 10240 subframes (10.24 s) after the initial reference (SFN or DFN equal to 0), see Section 14.2.3 in [165].

Concerning the transmission of PSCCH, it depends on the allocations mode (scheduled – mode 1, or autonomous – mode 2):

- In mode 1, the RB dedicated to the transmission of the SCI is signalled by the eNB through a new format 5 Downlink Control Information (DCI) scrambled with a Sidelink Radio Network Temporary Identifier (SL-RNTI) of the concerned D2D link with the traditional DL procedure (see [170] for more details). This resource must be inside the SC periods on the resources allocated to direct communication.
- In mode 2, the UE autonomously uses the resources of the RP dedicated to the control part.
- For both modes, these resources are described with the offset and bitmap method used for discovery, see Section 5.9.2.4. For instance, for FDD, the bitmap is 40 bit long.

The transmission of the data over PSSCH follows the control part of the SC period. The allocation depends on the allocation mode:

- Mode 1: the resources of the data part starts immediately after the last subframe dedicated to control (PSCCH) having the bitmap at 1. In fact, all the subframes in the data part of the SC period are allocable. The scheduled subframes are signalled through a Time Resource Pattern of Transmission (T-RPT) which is a bitmap, whose length is of 6, 7, or 8 bits, depending on the FDD/TDD configuration, e.g. for FDD it is 8 bits long. The T-RPT is signaled into the SCI Format 0 (see [165] for their definition) but this information is coming directly from the network through the DCI Format 5. These bitmaps, which are short, are repeated inside the SC period until its end (truncation possible at the boundary of the SC period). They are shown in light red in Figure 5-43, Allocation Mode 1 (see [165] for more details).
- Mode 2: in this case the data part starts after a configurable offset (in subframes) from the start of the SC period. The subframes of the resource pool are defined and configured through a second bitmap which has the same structure of the one of the control part (e.g. 40 bits for FDD) which is repeated and possibly truncated at the end of the SC period (see Figure 5-43, Allocation Mode 2). The scheduled subframes are selected on top of this pattern, thanks to the T-RPT (defined slightly differently with respect to mode 1 procedure, see [165] for more details).

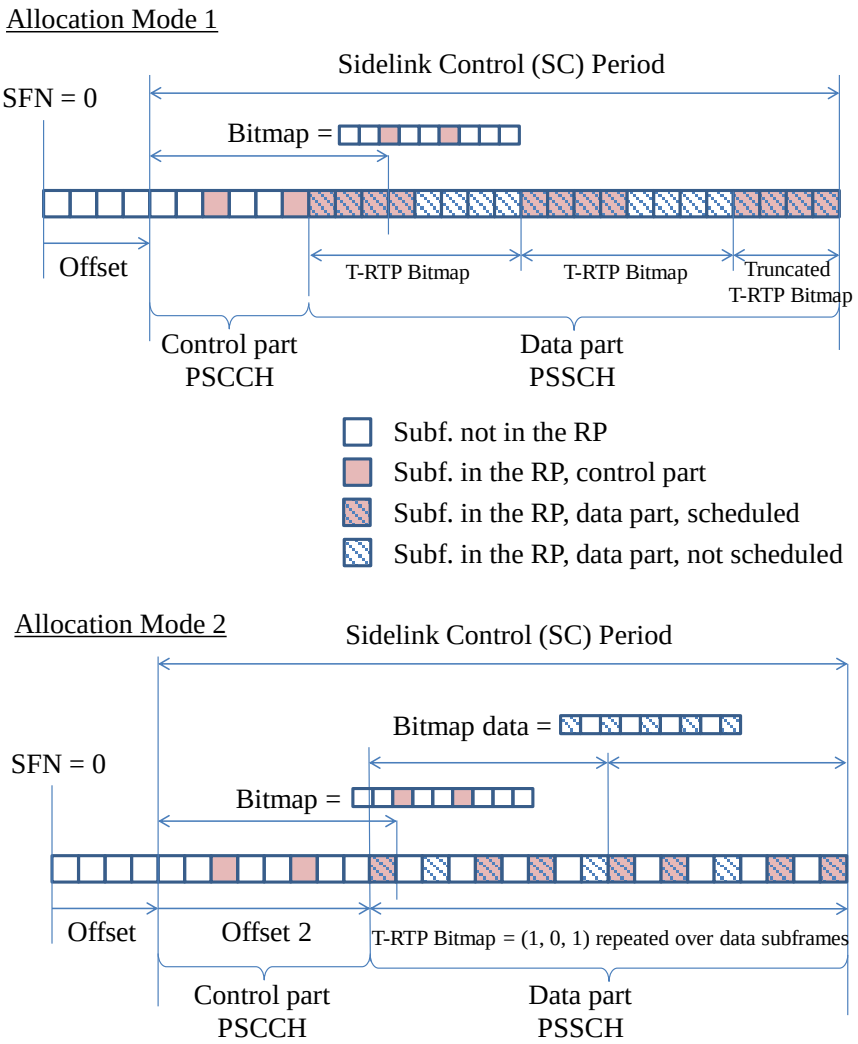


Figure 5-43 Resource pool definition for SL communication, temporal aspect. In the figure, only subframes for UL (eligible to SL) are shown

The difference in the definitions of the RPs for allocation mode 1 and 2 can be explained as follows. In mode 1 the network has full control, hence, even if all UL subframes are in the resource pool for SL communications, the subframes which are not explicitly scheduled by the eNodeB are guaranteed without interference from SL communications. This is no longer true in allocation mode 2, where autonomous resource selection is done by the UEs. Hence, the standard specifies a method through the bitmap for data in order to limit the subframes in the RP. In this way, even without knowing the scheduling decisions of the UEs, it can guarantee the absence of SL interference on the UL subframes out of the RP.

Let us go in more details on the transmission procedure:

- In mode 1, the UE is aware that the eNB supports SL communication when it sends SIB Type 18 signal. Then, when it has something to send, the UE sends an information message to eNB, containing frequency of interest for SL and a list of destinations. In this way it asks eNB for resources for transmission. The eNB sends a RRCConnectionReconfiguration message to the UE with MCS and the RPs, the SL-RNTI and also the Buffer Status Reports (BSR) configuration of the UE. In this way the UE is configured by the eNB to send BSRs about the amount of SL data available for transmission. RRC controls BSR reporting for the SL by configuring the two timers periodic-BSR-TimerSL (default infinity, possible values {5, 10, 16, 20, 32, 40, 64, 80, 128, 160, 320, 640, 1280, 2560, infinity} subframes) and retx-BSR-TimerSL

(default sf2560, possible values {320, 640, 1280, 2560, 5120, 10240} subframes) [167]. Of course BSR is deactivated when autonomous resource allocation is active (mode 2). When the UE has data to send it will send a BSR (or a scheduling requests if it has no resources already scheduled in the UL). Finally, in response to the BSR the eNB sends the allocations for the SL through a newly defined DCI format 5 and it signals the SL link through the SL-RNTI. Note that the MCS is not present in the DCI as for normal DL operation, but it is set directly as a configuration by the eNB RRC inside the RRCConnectionReconfiguration message. The SL Control period must start at least 4 subframes after the reception of the grant (see Section 5.14.1 in [167]). This may be helpful to calculate latency of control. A DCI Format 5 is valid until the expiry of the SC Period which it refers to. By using the DCI Format 5 and other configuration information, the sender UE can build the SCI and send it for data allocation.

- Mode 2 works out of coverage or in-coverage while the UE is in RRC_IDLE state (no control from the RRC layer). In this case the SCI is directly filled by the UE either using information coming from SIB Type 18, or, when out of coverage, from preconfigured resources. When there are multiple RPs that can be used, it is possible to preconfigure a list of priority of RPs, the UE shall select the ones associated to the Logical Channels with the highest priority, having traffic to send (see Section 5.14.1 in [167]). If two RPs have the same priority, it is up to the UE to decide which one to use. Then, specific resources for sending SCI and data must be selected randomly according to a uniform distribution over all the possibilities. Concerning the content of the SCI for the allocation of the data part, the sender UE can put any choice but at the current level of specification there is no measurement or report for understanding the quality of the channel, so the allocation is done at random and uniformly over the available choices.

In the control part of the RP, the SCI is sent over the PSCCH and it is repeated twice, over two subframes of the specified resource pool and using only 2 RB, so one PRB per slot in each of the two subframes. While in mode 1 the eNodeB can control the position of the SCI inside the RP such that collisions can be avoided between multiple SL senders using the same RP, in mode 2 each UE selects randomly and uniformly the position of the SCI. A frequency hop is used for the second subframe, see the detailed procedure in TS 36.213 [165]. The structure of SCI Format 0 is as follows:

- MCS: 5 bits, determined by RRC or by pre-configuration.
- Time resource pattern (T-RPT) 7 bits.
- Timing advance indication: 11 bits: in mode 2 it is always set to 0, while in mode 1 the network can specify a non-zero timing advance.
- Group destination ID: 8 bits: 8 LSBs of the ProSe L2 Group ID: this field specifies the destination (it substitutes in part the RNTI of the DCI message). This information must be combined with the destination identity of the MAC header to locate the exact identity of the receiver or of the group.
- Resource block assignment and hopping flag: 5-13 bits.
- Frequency hopping flag: 1 bit.

The last two fields determine the resource block assignment for the PSSCH and if there is or not frequency hopping and which kind of frequency hopping.

In both modes, PSSCH is scheduled by using resource allocation Type 0. In resource allocation of Type 0, a bitmap indicates the resource block groups (RBGs) which are allocated to the scheduled UE, where an RBG is a set of consecutive PRBs. The RBG size is a function of the system bandwidth [170]. PSSCH transmission supports (like PUSCH) no frequency-hopping or two types of frequency hopping: inter-subframe frequency hopping (for improving diversity for HARQ) and inter+intra-subframe frequency hopping (for improving diversity even inside one code block in addition to diversity for HARQ).

For both modes, transmission of the data over the PSSCH works as follows. The received SCI, together with the definition of the RP allows identifying the RB allocated for the first packet. Then, HARQ is always activated with a total of 4 transmissions for each packet. The retransmissions are always sent in consecutive subframes of the RP, but the RBs are defined in a deterministic way according to the

frequency hopping scheme and their position can be calculated by the receiver. The sequence of transmissions is fixed as well as the sequence of incremental redundancy versions, which is 0, 2, 3, 1. Once the first packet is sent, if the UE has more traffic it forms another MAC SDU and sends for transmission during the next set of four consecutive subframes in the SC period. Packets with the same pair of source and destination L2-indentities can be sent inside the same SC period. However, it is possible to receive from multiple sources and it is also possible that the same UE receives multiple transmissions if it belongs to different groups.

The maximum number of transmitting SL processes associated with a HARQ entity is currently 8 see TR 36.331 [168], the UE must declare if it supports it or not. Rel 13 states that the RB dedicated to SL transmission in a subframe cannot exceed 50 total RBs, see Section 14 in [165]. Moreover, SL transmission must happen only in contiguous RBs and there cannot be concurrent SL receptions with different CP in the same subframe. In case of interference in time with UL transmission allocations, the priority is given to the UL. A UE using a subframe in a carrier for direct communication cannot use the same subframe for normal cellular traffic.

Multiple transmissions within overlapping SC periods to different ProSe Destinations are allowed subject to single-cluster SC-FDM constraint.

Maximum powers for the PSCCH and PSSCH can be tuned separately but never go over the maximum power of the UE class. They are cell-specific and configured by RRC messaging or preconfigured in the devices.

5.9.2.6 SL UE-to-network relay

Some work was done in Rel 13 on the introduction of UE-to-Network Relays, i.e. a relay in between another UE (called remote UE) and the network, in order to provide network coverage extension [161]. Currently in the standard an L3 solution is foreseen, meaning that the relaying is done at IP level, the relay acts as a router. In general it seems that relaying could be managed at application level since there is no direct control for SL channels.

A UE-to-network relay must be by definition in the RRC_CONNECTED state, which means that UE-relay must be in coverage and able to receive configurations from RRC (see Section 5.10.10 in [168]). This UE can then be used for coverage extension, which is an important use case for public safety scenarios. The UE-to-network relay can transmit to UEs which are in-coverage (both states) or out-of-coverage, using SIB Type 18 configuration or using predefined values (see Section 5.10.4 in [168]). In any case, the resources for relaying are controlled by the network.

On the other hand, a UE can be a remote UE (i.e. a UE wanting to communicate with a UE relay) in coverage (both RRC states) or out-of-coverage. If it is in coverage and in RRC_IDLE state, then the UE can be a remote UE only if it is at the cell edge (see Section 5.10.11 in [168]).

A remote UE in RRC_IDLE state can transmit to a UE-relay only if it has selected a UE relay and it must also receive a RSRP metric from its primary cell which is below a certain threshold specified in SIB Type 19 (see Section 5.10.4 in [168]), i.e. it is at cell edge. In the other cases, the corresponding resource pools must be opportunely configured.

The UE-to-network relay can send a PS discovery message either in RRC_CONNECTED state with the right configuration or in RRC_IDLE state according to SIB Type 19 parameters. Notice that certain power thresholds conditions must be satisfied (see Section 5.10.10.3 in [168]). They are used to identify if the UE-to-network relay is at the cell edge (the RSRP must be below a certain high value minus a hysteresis value and above the minimum value plus hysteresis). Hence, also for discovery, the UE relay can be substantially used to send a discovery announcement in order to search UEs which are out of coverage or at cell edge. On the other hand, a remote UE can also send a PS discovery message either in coverage or out of coverage, but in RRC_IDLE mode it must satisfy the usual power conditions equivalent to be at the cell edge. For a remote UE in out of coverage, the announcement message works

with a random choice of the discovery configuration from the preconfigured ones and following the timing reference of a SynchRef UE if it exists or its own timing.

From the receiver perspective, it is possible to reselect the UE-to-network relay. This can be done out-of-coverage and in-coverage (both RRC states). Initial UE-relay selection is based on a received Sidelink Discovery Reference Signal Received Power (SD-RSRP). In case of reselection of UE-Relay, the reselection can be triggered by higher layers, or because SD-RSRP of the currently selected sidelink relay UE is below q-RxLevMin and in that case the UE selects the UE-relay which is in the list reselectionInfoIC (in coverage) or reselectionInfoOoC (out of coverage) and whose power is above the minimum level by an hysteresis value [168]. The standard also says that reselection can be influenced also by metrics at higher layers which are UE implementation dependent.

Finally notice that the definition of cell edge we have used in this section is based on thresholds which have configurable values. The network can hence control, at least to a certain extent, the geographical location in which D2D is allowed.

5.9.3 Proposed solution

The state of the art work on D2D systems very often covers only particular aspects or functionality of the system or of the architecture (see also the review in Section 5.8). The 3GPP proposal of D2D described in Section 5.9.2 has the advantage of describing a complete system which could possibly evolve in the next generation (5G). Hence it is our baseline for improvements.

In this section we are particularly interested in the relaying aspect for coverage extension and out of coverage communications. From Section 5.9.2 we have the following observations:

- Lack of control from the network of out-of-coverage UE at any distance (in hops) from the UE-to-network relay or from the eNodeB
- Only UE-to-network relay is defined for the moment, at L3 level. Inter-UE relaying is possible at L3 but out of the scope of the 3GPP definition.
- For out-of-coverage communications, everything, ranging from security and identification, to resource pool definition and parametrization of lower layer, is managed through preconfigured values. When an out-of-coverage UE interacts with a fixed network, it is up to the person doing the pre-configuration to handle the definition of the resource pools so that a communication is possible.
- Nothing is said about how to manage resource pools when interacting at cell edge between different cells in order to limit interference and satisfy requirements for public safety services.
- Selection of resources at L2 for out of coverage users is done randomly with uniform distribution on the set of possible choices, which may introduce high level of interference depending on the available resources.
- PHY layer coding and procedures are very robust (multiple repetitions with HARQ), which is in line with the previous vision of an autonomous system generating unmastered interference and collisions.

The solution we want to investigate for the COHERENT D3.2 is to introduce a form of control for coverage extension and relaying, in the framework of the control architecture defined in COHERENT D2.2 [140]. The aim is to see to what extent introducing more control than what is currently defined will improve the support of PS requirements, which may be stringent. For instance, for traditional voice application like group calls, which are of paramount importance for PS users, the latency in multi-hop coverage extension scenarios may be too high. Different possibilities are open, ranging from the introduction of minimal additional control of the existing configurable parameters, up to full control. Indeed, optimal relaying strategies depend also on the degree of control and on the available measurements and metrics. In the proposed solution it will then be important to evaluate the feasibility of gathering sufficient useful information on the state of the D2D links and nodes for building the network graph. The impact on system performance of the measurements and signalling procedures

related to the construction and sharing of the network graph should be evaluated as a function of the different control strategies.

5.10 Mobility management in HetNets

In a broad sense, mobility management in wireless networks encompasses handovers, location management and roaming. We consider the first type of mobility, with the focus on LTE(-A) HetNets, and give our insights into the required functionality and possibilities for abstractions of operating parameters that are included into COHERENT control framework.

Mobility procedures depend on the UE's RRC state: "Idle" or "Connected". Recently, also "Connected Inactive" state has been proposed. It enables fast and lightweight transition to the Connected state, by keeping the UE context alive in the network during UE's inactivity periods. That is, from the core network's perspective the UE stays in the connected state. The need for the new intermediate state stems from the control plane signalling and latency reduction requirements in the support for cellular IoT. The related suspend and resume procedures are being defined in [152].

In higher frequency bands, cell search and mobility support become challenging because the coverage narrows down from cells to beams. Then the network has to actively track the UE in its movement by using beam steering techniques. Here, the control plane at the lower-frequency layers may assist in the beam detection and direction synchronization.

With the focused support for the verticals, like various forms of MTC, mobility support will need to be separately optimized for (semi-)static devices on one hand and for various applications in fast moving vehicles, on the other hand. This means that a UE may perform measurements on reference signals only on demand, and that the signals may be transmitted along with data and only when needed.

A further research issue for 5G is seamless integration of mobility management across heterogeneous radio technologies. In this case, access network discovery and selection function (ANDSF) was supposed to facilitate access network selection between LTE and WiFi, but this feature was not attractive for the operators. Instead, UE makes an autonomous decision on WiFi usage. The tighter inter-RAT integration, e.g. LTE and WiFi link aggregation (LWA) entails new requirements for mobility management, the fulfilment of which may benefit from the centralised network view provided by C3. Load-aware access network selection among multi-radio heterogeneous networks, and especially in cellular and WiFi interworking, has been examined in [153].

5.10.1 Problem description

Mobility in the connected mode typically refers to UE's handovers between eNBs. Briefly, the procedure involves the phases of cell detection by UE, measurements on the detected cells according to eNB's configuration, and reporting of filtered measurements to eNB at certain conditions (e.g. at A3 or A5 events). Finally, the serving eNB negotiates on the handover with the target eNB. Here, the configuration of the measurements has performance implications in terms of handover failures, unnecessary handovers (e.g. ping-pongs), throughput, signalling overhead, power consumption, and interference, among others. Therefore, it is important to carefully plan the mobility parameters like signal strength thresholds, time-to-trigger (TTT) hysteresis, measurement filtering coefficients, reporting interval and the maximum number of reported cells [135]. The measurement parameterization needs also to take into account contextual information related to cells sizes, UE mobility, traffic types, and energy constraints, for example [136].

Along with the HetNet developments, including carrier aggregation and CoMP, that distribute the data plane among multiple small cells, handover has become a means for maintaining a robust control channel between UE and a macro cell. Indeed, UE's movements among densely deployed small Secondary Cells are not signalled as handovers but they appear as additions, removals, and changes of cells. Further evolution towards 5G with ultra-dense mmWave beam cells may allow nearly borderless mobility (without a need for paging), which makes use of uplink beaconing [137]. Dense mobile network deployment with multi-connectivity makes it difficult to decide on handover parameters like

the cell offset values. A centralized approach with network view could help to control those parameters dynamically. The problem that is aimed to be solved here fulfil the COHERENT requirement #54 stated in D2.1[15].

5.10.2 Solution approach

The centralized COHERENT C3 controller for mobility management applies the abstracted measurement results such as RSRP/RSRQ from the network graph as the input to the mobility management application. The optimal mobility management parameters can be determined and set by the mobility management application; after all, the handover will be triggered by the mobility management inside the COHERENT controller and enacted to the network. Figure 5-44 depicts the interaction between the mobility management application and the COHERENT control framework based on the X2 handover in LTE. The same methodology is applicable to the S1 handover.

Specifically, the mobility management parameters are described as follows:

- Time to trigger (TTT): Time-to-trigger parameter
- Hysteresis: the hysteresis parameter for this event.
- OFN: the frequency specific offset of the frequency of the neighbour cell
- OCN: the cell specific offset of the neighbour cell, and set to zero if not configured for the neighbour cell.
- OFS: the frequency specific offset of the serving frequency.
- OCS: the cell specific offset of the serving cell, and is set to zero if not configured for the serving cell.
- OFF: the offset parameter for this event. In LTE, several different events (A1~A6) [90] can be used to trigger the intra-LTE handover process with different configuration on each individual offsets (e.g., a3_offset).
- Threshold: The threshold parameters used for different events. For some events, there are two thresholds used for the different cells.
- L3 Filtering coefficient RSRP/RSRQ: Parameter for exponential moving average (EMA) filter for smoothing any abrupt measurements variations.

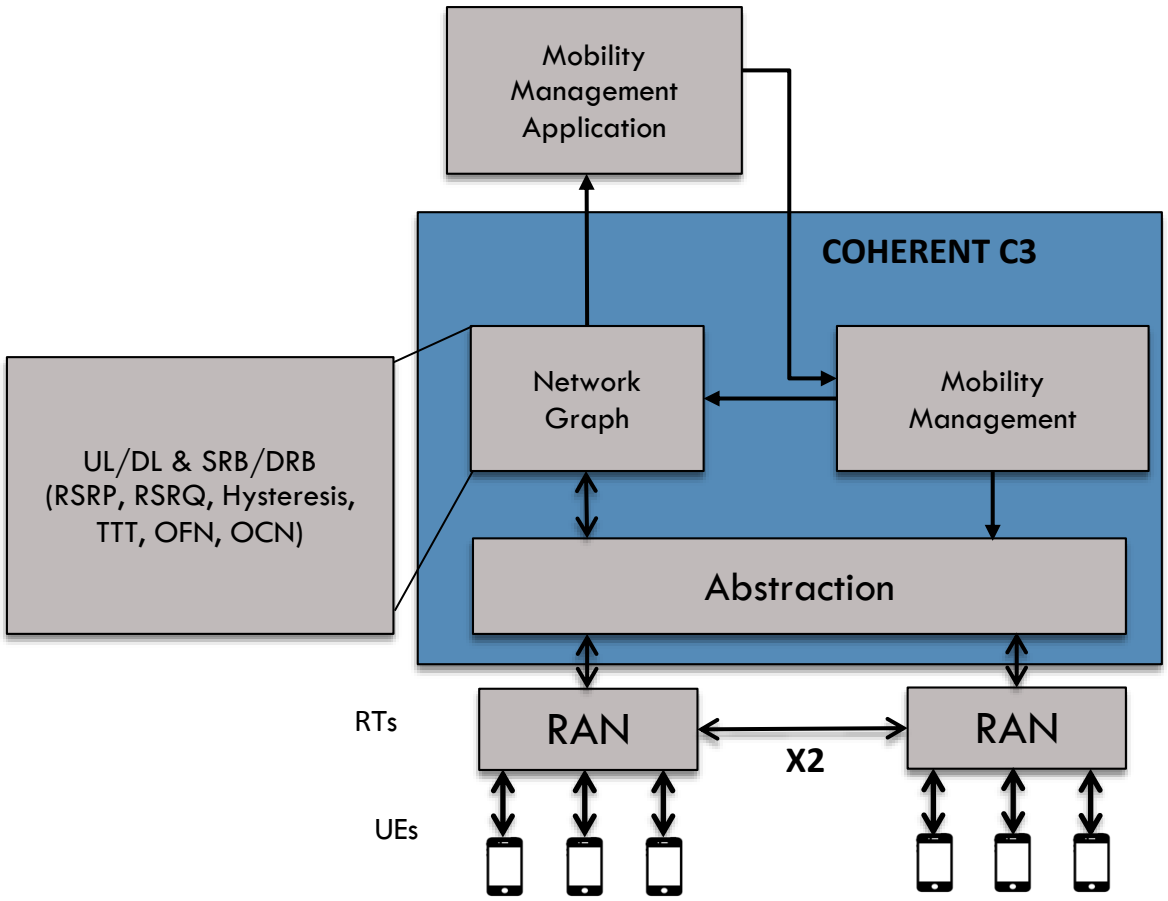


Figure 5-44 Mobility management application interaction with SDN controller

The mobility management application communicates with the mobility management entity located in the SDN controller side via a northbound API. The mobility management entity enables the abstraction entity to acquire the intended RAN raw data information, e.g., RSRP, RSRQ, SINR, CQI, PRBs from the underlying network via a southbound API and perform abstraction methodology, e.g., BLER or capacity estimation. To that end, using the abstracted information other performance parameters such as average delay, cell load, data rate, KPIs for UL or DL and other QoS related parameters can be computed and stored in the network graph supported by the proposed framework. Then, this information is sent to the application that can run a background algorithm based on a combination of load, traffic, power and other handover criteria for optimizing the aforementioned mobility management parameters. The output values of the parameters of the application are sent back to the mobility management entity in order to be applied to the RAN.

Mobility across diverse radio interface technologies requires technology-agnostic abstractions of the mobility control functionality. One effort towards that direction is IEEE 802.21 Media Independent Handover (MIH) [207]. An SDN-based framework for centralised RAN control, including programming abstractions for mobility management in WiFi access networks, has been proposed in [139].

The high-level research objective is the integration of LTE's mobility control functionality into the COHERENT control framework. This may entail:

- Abstractions of LTE's mobility-related parameters and measured entities that are to be included in the COHERENT graphs.
- Centralisation of the currently decentralised handover decision algorithms and related signalling

Identification of the basic biasing features in current LTE SON (e.g., mobility load balancing and mobility robustness optimization) and search for optimal biasing factors with the aid of the network graph approach.

5.11 Mobility study for high speed platforms

5.11.1 Problem description

As a result of low latency needs and reliable communications with low CAPEX (installation costs) and low OPEX (operation & exploitation costs), both high speed train operators and civil flight operators are looking toward the use of 4G & 5G technologies as opposed to the more-expensive and higher-latency option which is currently represented by satellite communications. This would allow for example higher throughput and lower latencies to the users located in the airplanes at a much lower cost, as shown in the comparison table in Figure 5-45, for the comparison between DA2GC (Broadband Direct Air to Ground Communications) systems and SATCOM (Satellite Communications).

TCS/OPS/HTE/STR /WFD **Source**	Intended Coverage	Lifetime	CAPEX	Expected Efficiency	Cell Range	Preferred Usage	Potential Date of Service	Remarks
DA2GC (5-10$/ month/user)	**500 eNBs** (EUROPE only but can be extended to other regions, with extra cost)	**> 25 years,** with possible upgrade (e.g. new release) **Low OPEX**	**At least 15 Millions €** (if 30k per eNB)	**Smaller cells, higher order modulation, better spectral efficiency**	**50-70 km** (eNB in the middle)	**Continent** Main advantage: Capacity & Cost	eNBs on the shelf, end 2017 for tests **Estimated service time: 2018**	Possibility to add extra cells, use of trisectorized cells, use of smaller cells. This could increase system capacity by a factor of at least 10. Moreover, DA2GC system is to be used only for airplanes (dedicated system), meaning a system not shared with normal users, and therefore sufficient for this type of use.
SAT COM (>60$/ month/user) https://www.gogoair.com/	**640 satellites** **(worldwide)** http://spacenews.com/competition-to-build-oneweb-constellation-draws-2-u-s-3-european-companies/	**< 25 years,** without possible upgrade **Very high OPEX**	**At least 1800 Millions €** (2 billion $) https://forum.nasaspaceflight.com/index.php?topic=36590.0 (500k per satellite, with high expenses to launch and put into orbit) http://spacenews.com/competition-to-build-oneweb-constellation-draws-2-u-s-3-european-companies/	**Larger "cells", lower order modulation, lower spectral efficiency**	**500-700 km** (satellite in the middle of the cone)	**Sea/ Ocean** Main advantage: coverage **System extension, not high capacity network:** http://dcinno.streetwise.co/2016/01/26/va-based-internet-satellite-co-oneweb-plans-2016-raise-hiring/	640 satellites to be manufactured Estimated first launch 2017-2019, during 3 years time (18/Month) **Estimated service time: 2020 (or after)**	Once deployed, nothing to do, since the system maintenance is complex and system upgrade is almost impossible. Moreover, satellite system can be used for providing any type of connectivity: 5 billion users around the world. Since is not a dedicated system, the system may be shared by multiple operators/ and for multiple usages and therefore it may not be sufficient for aircraft use. Also, due to higher Doppler, tracking antennas, handovers & higher gains, the receivers will be more expensive and adapted for satellite bandwidths such as 6GHz Ku bandwidth.

Figure 5-45 Comparison between broadband direct air to ground communications satellite communications

However, a new design with innovative improvements of the current technology (i.e. LTE-Advanced and LTE-Advanced Pro up to Release-14) or new technology robust to new service requirements (i.e. Release-15, Release-16 with 5G) is required. This new network design is needed in order to be able to provide necessary communication capabilities for products for High Speed Platforms such as Airplanes and High-Speed-Trains.

For trains and airplanes the propagation is different from classical ground-to-ground propagation due to reflections and very particular channel models, high Doppler shift and high Doppler spreads, interference, service interruptions and handover issues which make LTE system unusable in the above mentioned conditions. As a result of these previous mentioned issues, 3GPP decided to initiate work for 5G systems in order to cope with High-Speed broadband access & Vehicle-to-Vehicle communications Use Cases (and more generally, Vehicle-to-X communications or V2X – for one-to-one or one-to-many) in the standardization group called SA1 (for details see 3GPP SA1 SMARTER TR 22.891 [171]). More precisely and as also cited in COHERENT deliverable D2.2 [140], SA1 SMARTER TR 22.891 and SA1 TR 22.863 [172] refer to improvements related to high user mobility related to "enhanced Mobile Broadband" (eMBB) & "enhanced Vehicle-to-X communication" (eV2X):

- SMARTER UC10 Mobile Broadband Services with Seamless Wide-Area Coverage (duplicate within eMBB & eV2X);
- SMARTER UC29 Higher User Mobility (for eMBB);
- SMARTER UC66 Broadband Direct Air to Ground Communications (i.e. DA2GC, for eMBB).

5.11.2 Proposed system

With respect to DA2GC use case, Thales has already made an estimation of the number of eNBs needed to cover the entire Europe and reached the conclusion that around 470 eNBs may be sufficient with respect to the current traffic, as represented in Figure 5-46.

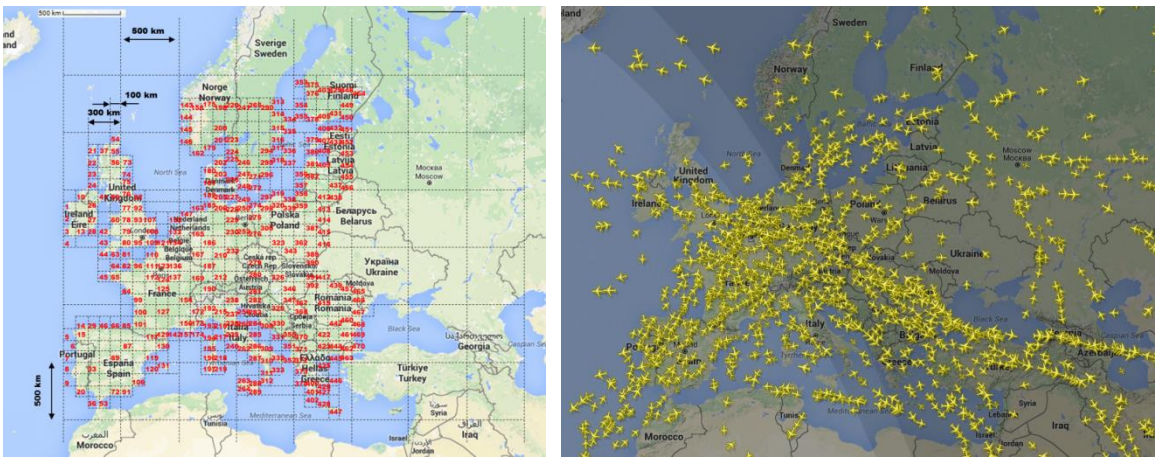


Figure 5-46 Estimation of needed eNodeB for European coverage for DA2GC

The estimation was based on the current traffic model for business and non-business needs:

Table 5-7 Traffic model for business and non-business needs for DA2GC

User profile	Business (no video)	Common (with video)
Data rate average by user (Kb/s) DL	51	5 778
Data rate average by user (Kb/s) UL	13	620

If we consider a passenger average capacity per airplane of 180-200 passengers and an internet user estimation per airplane of 25%, and considering e.g. business user case estimated needs of 51 kb/s downlink (DL) data rate and of 13 kb/s uplink (UL) data rate, we will find that the more stringent requirement for a business user and airplane may be respected, but the more stringent requirement for a non-business user and airplane is more difficult. However, for normal non-business user access, premium user strategy might be applied, which will decrease the total required traffic in DL and increase the service quality per (premium) user. These results can be resumed in Table 5-8.

Table 5-8 Throughput requirements for DA2GC

Traffic Type: UL (UpLink) or DL (DownLink)	Service Type (examples)*		Usage / 100x100 km² (for business access only)	
	On-board Internet Access (business access only, during peak periods)	On-board Internet Access (normal user access only, during peak periods)	50-80 airplanes - maximum density (e.g. near airports)	15 airplanes - typical density
UL Minimum Required Traffic	0.7 Mbits/sec	30 Mbits/sec	Maximum 35-56 Mbits/sec/10000km²	Maximum 10 Mbits/sec/10000km²
DL Minimum Required Traffic	2.5 Mbits/sec	300 Mbits/sec	Maximum 125-200 Mbits/sec/10000km²	Maximum 37,5 Mbits/sec/10000km²

Of course, other additional services might exist, e.g. for operating the airplane – but this would require very low data rate with very high priority. Anyway, it is clearly shown that current minimum required throughput needs are not satisfied by satellite communications technology. However, LTE might be the

answer, and therefore further deployment 5G versions offering backward compatibility but also new services and waveforms. This is theoretically possible as shown in Table 5-9, Table 5-10, and Table 5-11 for FDD mode with 1 RF chain, 2 RF chains and 4 RF chains, but LTE has not been designed for very high speed platforms (e.g. above 500 km/h at 2 GHz) and has to be evaluated and modified accordingly.

Table 5-9 Capacity [Mbps] for FDD, 1 RF chain

Mod. order	MCS/TBS Index	5 MHz (25 RBs)	10 MHz (50 RBs)	15 MHz (75 RBs)	20 MHz (100 RBs)
2	0/0	0,68	1,38	2,08	2,79
	4/4	1,8	3,62	5,35	7,22
	9/9	4,00	7,99	11,83	15,84
4	10/9	4,00	7,99	11,83	15,84
	13/12	5,73	11,44	16,99	22,92
	16/15	7,73	15,26	22,92	30,57
6	17/15	7,73	15,26	22,92	30,57
	23/21	12,57	25,45	37,88	51,02
	28/26	18,336	36,69	55,05	75,37

Table 5-10 Capacity [Mbps] for FDD, 2 RF chain

Mod. order	MCS/TBS Index	5 MHz (25 RBs)	10 MHz (50 RBs)	15 MHz (75 RBs)	20 MHz (100 RBs)
2	0/0	1,36	2,76	4,17	5,58
	4/4	3,6	7,24	10,70	14,44
	9/9	8,01	15,98	23,66	31,68
4	10/9	8,01	15,98	23,66	31,68
	13/12	11,47	22,89	33,98	45,84
	16/15	15,47	30,52	45,84	61,15
6	17/15	15,47	30,52	45,84	61,15
	23/21	25,15	50,91	75,77	102,05
	28/26	36,67	73,39	110,11	150,75

Table 5-11 Capacity [Mbps] for FDD, 4 RF chain

Mod. order	MCS/TBS Index	5 MHz (25 RBs)	10 MHz (50 RBs)	15 MHz (75 RBs)	20 MHz (100RBs)
2	0/0	2,72	5,536	8,352	11,16
	4/4	7,2	14,49	21,40	28,89
	9/9	16,03	31,96	47,32	63,36
4	10/9	16,03	31,96	47,32	63,36
	13/12	22,94	45,79	67,96	91,68
	16/15	30,94	61,05	91,68	122,3
6	17/15	30,94	61,05	91,68	122,3
	23/21	50,30	101,82	151,55	204,1
	28/26	73,34	146,78	220,22	301,5

In order to obtain LTE's maximum achievable capacity (here expressed in Mbps) in FDD mode for different RF chains, we have used MCS and TBS indexes taken from Table 7.1.7.1-1 from TS 26.213 [173]. Moreover, it has been considered that approximately 25% of overhead is used for control & signalling, which leads to a very good approximation of useful expected throughput.

Table 5-12 resumes the QoS requirements/service. The first connection establishment time does not need to be very low for normal Internet access but the delay variation has to be decreased, especially if we want to include voice communication and other real-time communications requiring low interruption times, low latencies, and more robustness.

Table 5-12 QoS requirements per service for DA2GC

QoS Requirements	Service Type (examples)**
	On-board Internet Access (business access only)
First Connection Establishment Time	< 5 sec.
Connection Delay	< 5 sec.
Delay Variation	< 1 sec.
Availability of Data Rate	In-flight & Approach Area (without On-Ground Airplanes that already landed or are preparing to take-off)
Packet Loss	Yes (Allowed)
Required Speed	Up to 1000 km/h
# of Users Sharing the Same Connection	180-200 users/airplane

Other additional services might exist (see Table 5-12), e.g. for CCTV applications, for control purposes, or for location purposes – but this would require very low data rate with very high priority, so throughput is not a problem for this type of service. However, the prioritization and the differentiation of data flow should be reconsidered.

5.11.3 Proposed solution

The SA group (and more precisely SA1) therefore started to define NEW requirements for NEW services resulting from the above mentioned use cases. These requirements will be transformed in technical and system specifications for future releases (starting from Release-14) in different standardization groups such as RAN (Radio Access Network, and more specifically RAN1 to RAN5), CT (Core Network and Terminals) and SA2. Here are a few requirements as found in SMARTER TR 22.891 [171]:

- Services: voice and video call, internet access and mobile multimedia services, during flights (see [174]).
- Pre-conditions: A service access network infrastructure, e.g. eNB and WiFi AP (both already certified for on-board implementation) is provided in an airplane to offer passengers in-flight mobile voice and broadband data communication services. During a flight, the speed of an aircraft ranges between 500 km/h and up to [900 km/h] at different altitudes between 4000 meters and 10000 meters.
- Service flows and post-conditions: During the flight, web surfing, phone calls, video conference meetings full HD streaming live video, full HD streaming movie and video calls have to be available.

Due to bandwidth limitation and very low waveform robustness of DA2GC system, streaming live video and other services may deteriorate and my encounter reduced quality such as lower resolution, high latency or service reduction to voice-only and even non-real-time transmissions.

5.11.3.1 Solution description

We believe that COHERENT architecture is very suitable to these new cases and would allow for a more flexible RAN to be used. This will improve the overall system performance without significant system modifications.

The entire LTE communication system has therefore to be properly evaluated with respect to new usage constraints:

1) usage for high-speed platforms

2) usage for air-to-ground propagation scenarios.

New handover algorithms and system robustness to offsets and high mobility have to be properly evaluated. Through network graphs information and new COHERENT controller, the system will permanently adapt providing robust services for a very high mobility. Some of the encountered possible issues are described in Figure 5-47 and Figure 5-48. For example, a UE mounted on an airplane is connected to eNB A, but it receives the signal from eNB A with an initial offset (see step 1), which is able to compensate at PHY level during the cell search process. The serving cell (which is eNB A in this case) may configure UE on airplane with the cell list to be measured (e.g. eNB B), with the proper triggers, reporting configuration and measurement ID (as seen in step 2). However, as represented in step 3, the UE "sees" the pilots from eNB B with another frequency offset, offset that it is not able to compensate. For this reason, at step 4 UE measures eNB B and underestimates received power from eNB. This will result to incorrect and unreliable measurement values, which will be anyway probably reported to eNB A too late during the step 5. At step 6, eNB A performs handover, but a huge interruption time is expected. This interruption should be normally avoided with the help of new architecture thanks to COHERENT controller and/or waveform robustification.

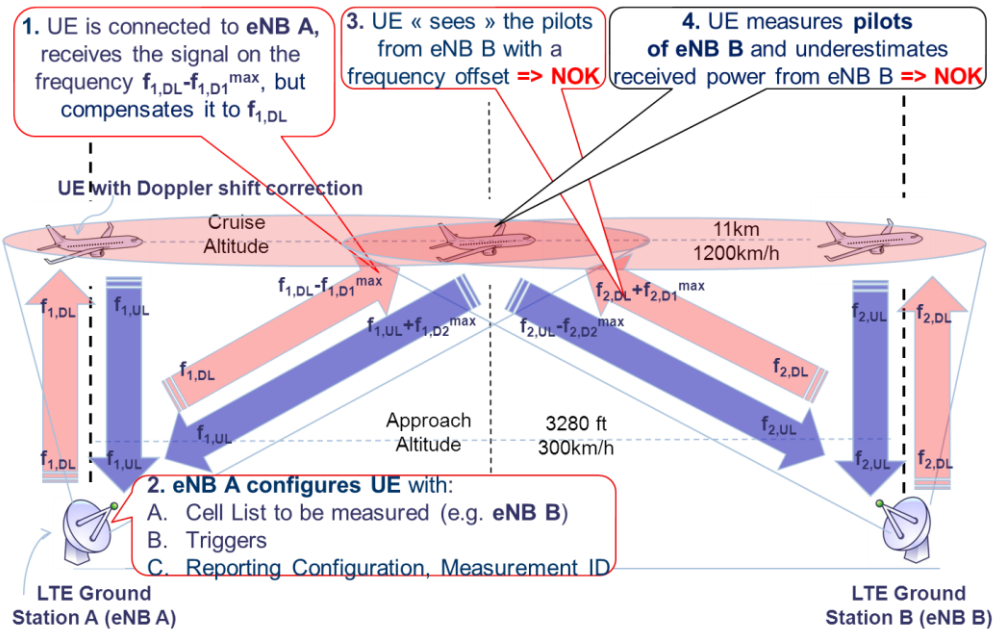


Figure 5-47 LTE measurement issues related to high Speed (part I)

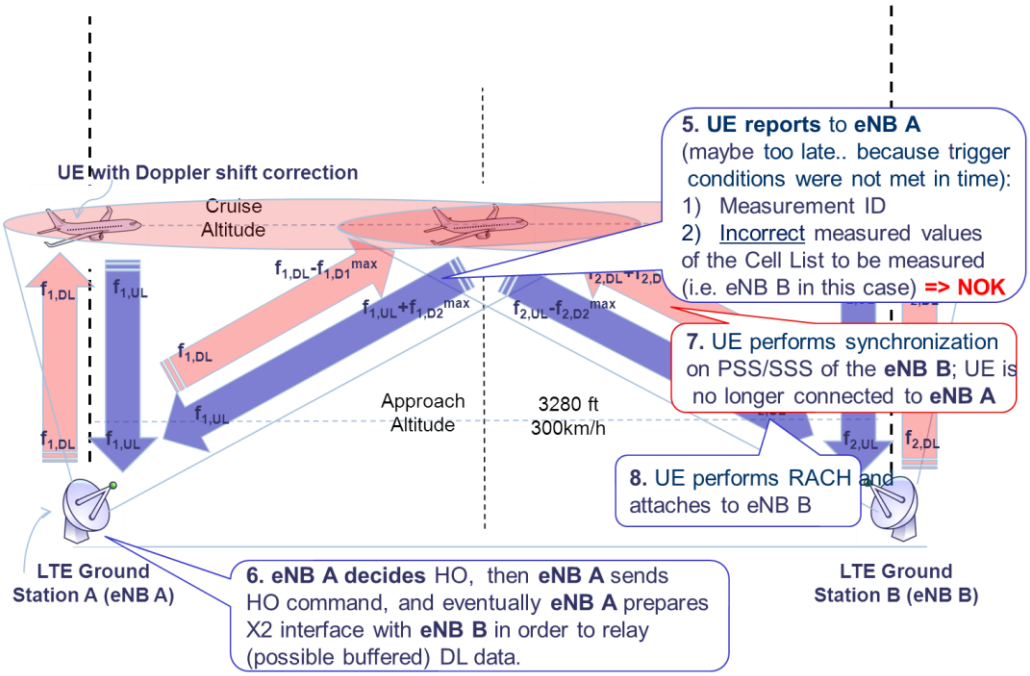


Figure 5-48 LTE measurement issues related to high Speed (part II)

5.11.3.2 Network graphs and abstraction aspects

Final results, methods and/or evaluations will be provided in COHERENT deliverable D3.2. For example, an evaluation of PSS and SSS and/or CRS signals will be considered. Abstraction metrics will further be investigated in D3.2, together with potential different implementation schemes with network side impact and/or User Equipment impact. Periodicity of measurements, dedicated system configuration, and other system parameters will be discussed, with potential impact at different layers.

6. Conclusions

The key innovations of COHERENT lie in the proper abstraction of physical and MAC layer in 5G heterogeneous network and the development of a unified and programmable control framework. In essence, our main objective is to investigate logically centralized and physically distributed approaches for producing computationally light-weight models of the network state which can be represented in a compact form as part of the network graph.

6.1 Summary of main outcome

In this deliverable, an overview is first given in Section 2 on network graphs, control paradigms, abstraction principles, and current standardization activities. Some key challenges are provided to which the developed network graph concept should respond. Next, in Section 3, the COHERENT system model is outlined and the target abstraction framework is introduced. The work of WP2 is closely taken into account to the extent allowed by planning of the work which scheduled the parallel work on architecture in WP2 and on PHY/MAC layer modelling and abstractions in WP3. In this deliverable, key performance indicators of D2.1 [15], network architecture developed in D2.2 [140], and high level requirements referred to the abstractions for the PHY and MAC layers of COHERENT architecture were taken into consideration. Section 4 presents the main relevant metrics and measurements already available in the state of the art of the systems targeted by COHERENT (namely WiFi and 3GPP LTE) that can potentially be used to generate the COHERENT network graph. Some new ideas that are also relevant to WP4-WP7 in the form of utilizing network graphs are introduced. Moreover, in Section 5, selected interesting control applications are presented, and evaluated how the network graph concept can be applied. Important remarks and conclusions are made on the applicability of the network graphs in different use cases, along with indications of how the proposed network graph concept can be used to address the defined problems. Some of the key highlights and observations are as follows.

6.1.1 Network graphs

Concerning the network graph concept, at this stage of the work, the following conclusions can be drawn. The network graph is considered as an abstracted data structure, with nodes describing network elements, their attributes and sets of such, and edges describing their interactions, labelled by tensors of measures such as connectivity state, inter-node resource constraints, and interference coupling. The measures in question are differentiated in several dimensions. On one hand, there are distinctions between LTE and WiFi measures (generally among different radio access technology measurements), and between PHY and MAC layer measures (generally among measurements coming from any protocol stack layer or sub-layer). On the other hand, there are distinctions between different time scales, where local information is updated in the network nodes dynamically, whereas global information is aggregated on a semi-static time scale, and accessed by an entity of the COHERENT control framework.

A network graph is any aggregation of the idealized data structure that contains all network elements, and all possible metrics between any pair of network elements. The most important general aggregation principles are the following:

1) **Clustering**, where network elements with similar qualitative or quantitative properties are considered as a single node in the graph, and metrics are replaced by annealed metrics over the whole cluster. An example is a cellular network graph, where interactions are not considered between primary network elements but between cells.
2) **Edge selection**, where a certain metric is stored only between edges where this or some related metric exceeds a certain threshold value. An example is how spectrum assignment is performed by coloring a conflict graph, in which the only edges considered are between cells where the co-channel interference is non-negligible with respect to the targeted application.

3) **Statistical revision**, where metrics are replaced by their average over the dimension(s) of interest (e.g. time, frequency, space, number of users, etc.), or by estimates of their probability distributions.
4) **Functional revision**, where collections of parameters are replaced by functions of these, such as CQI, Shannon capacity, or user experience.

The network graph concept is already present in the literature and implicitly or explicitly used in many works for modelling the metrics of interest in the scope of those works in a concise but effective way. However, the COHERENT project inserts the concept of network graph inside the COHERENT architecture and, by means of the previous aggregation principles, tries to offer a general framework and systematic approach to the implementation of the network graph concept at the different functional entities on the said architecture. Each of the previous aggregation principles can be performed whenever information is passed between two levels in the hierarchical COHERENT architecture.

Network nodes and C3 entities may build local graphs, depending on the use of their information, according to the SLA and QoS requirements. In most cases, measurements and processing can be done locally at the nodes as a trade-off between network observability and scalability. The channel state parameters are allowed to take values of individual measurements as well as parameter estimates for probability distributions. In discussions between WP2 and WP3, we see that the COHERENT controller framework consists of a hierarchical structure, where the C3 layer provides a logically centralized network view, based on information provided by COHERENT controllers administrating network equipment within certain regions and RATs. For scalability and time-critical applications, real-time control actions are carried out by local control functions (e.g. scheduling) at relevant nodes. The hierarchical C3 framework enables flexible creation and aggregation of network graphs as needed for different levels of control.

A COHERENT C3 gathers and operates on network graph information based on: i) abstractions of low-layers procedures, protocols, techniques, etc.; ii) detailed local low-layer network graphs of the RAN links, at PHY and MAC layers. In C3, the network graph should be, as far as possible, technology agnostic in order to be exposed at the northbound interface for generic control and management applications, while network graphs in the RTC or at lower layers are RAT dependent. The south-bound interface is between a COHERENT controller and the nodes of the RAN of different RATs. It is used to communicate parameters coming from the RAT that will be used by a COHERENT controller inside low-layer abstractions of the RAN (implemented in a COHERENT controller), so that a COHERENT controller can build local low-layer network graphs. The south-bound interface is technology dependent.

6.1.2 Control applications

Section 5 presents a number of control applications in which the COHERENT control framework, including network graphs as seen in this project, is used in different forms. The following paragraphs summarize the main contributions in Section 5.

In Section 5.1, the LTE downlink PHY transmission model is described and link-level performance is evaluated. Specifically, cell-specific reference signals and SNR abstractions are emphasised. A sophisticated control model to handle the overall channel estimation process is described as well. A test bench for the studied PHY link-level model is generated and the resulting link throughput is evaluated.

Section 5.2 provides the motivation to use a probabilistic and distributed aggregation model to build a CQM. This model enhances the network abstraction by capturing the underlying distributions of time varying links. Bayesian estimation over the joint distribution of RSSI and PDR in time windows is used to effectively capture fading as well as interference effects and can be used as link deterioration triggers based on prior information. Our current work involves the evaluation of the metrics and models proposed for the CQM in the context of different control and management functions that they enable.

In Sections 5.3 and 5.4, distributed antenna system and CoMP technologies are, respectively, studied. Specifically, a distributed antenna system (DAS) architecture, which consists of multiple remote radio heads (RRHs) connected to a centralized base station via optical fiber network, is considered for modelling and evaluation of the system at PHY level based on the 3GPP LTE processing blocks. DAS promises to deliver viable solutions using its distributed RRHs, thereby supporting rapid frequency re-use, minimizing interference due to very low power RRHs and increased coverage. PHY layer processing blocks for incorporating the DAS functionality to 3GPP LTE have been analysed and implemented on a digital signal processor (DSP) based software defined radio (SDR) platform. A DAS technique has been tested using a SDR platform and LTE analyser to measure the performance of DAS focussed on error vector magnitude observations. Usability of network graphs has been highlighted. Complete analysis on abstractions, metrics and methodology of using network graphs for DAS will be done in D3.2. The CoMP technology is shortly outlined with future plans to be studied in D3.2.

In Section 5.5, dynamic inter-node interference modelling study is presented. Specifically, the power spectral densities of different transmission technologies are provided. Furthermore, adjacent channel interference ratio is evaluated between GSM carrier and LTE carrier. The existing standards have static requirements on the required adjacent carrier leakage ratio. However, as the number of systems operating in adjacent bands increases, it is foreseen that a more dynamic approach to interference modelling and protection is required.

Section 5.6 presents the background and progress of a distributed load balancing mechanism for RAN in which the cell range expansion parameter is dynamically adjusted based on a locally derived target value, in essence a neighbourhood average. The progress achieved within COHERENT is the definition and prototype implementation of a new target metric, balancing the risk of overload rather than the relative radio resource load directly. This reduces the QoS impact of the earlier mechanism, and takes the variability of the load metric into account. Preliminary quantitative analysis of simulator results w.r.t. system efficiency is also described. Furthermore, load balancing is studied in energy-limited HetNets. Specifically, an abstraction method of energy status of small base stations harnessed with specific energy charging capability from the surrounding environment is proposed. We compare the centralized control approach with the corresponding distributed approach to illustrate the potential benefit that could arise from that of centralized control in this scenario.

Section 5.7 provides a short introduction to cognitive radio and presents the idea that certain cognitive radio capabilities may bring a benefit to 5G networks. The functionality of these networks is centered in the lower layers of the architecture, therefore it is of interest to examine such networks since their management and configuration is very relevant to the ideas and concepts of COHERENT project. These networks can be very efficient in terms of satisfying the requirements imposed by 5G PPP. In particular sensing and all the tools developed for monitoring the electromagnetic environment should be considered as an input for giving information to feed the COHERENT network graphs.

Section 5.8 considers the D2D network management problems (e.g. D2D discovery, cell association, D2D relay management, resource allocation, and interference coordination) in HetNets. In dense HetNets, cell dimension diminishes and cell boundaries get blurred at the same time. The traditional assumption that D2D resource allocation and interference management can be done as if one single cell has all D2D communications under its control is less valid. Moreover, the management of the D2D network in HetNets shall also take the legacy uplink or downlink decision and measurements into account in order to have better resource utilization and interference coordination. We propose to utilize D2D network graph under the COHERENT control framework to address these D2D network management problems. After presenting a state of the art in the open literature and inside 3GPP, as far as high layers and discovery functionality is concerned (which complements with the review in Section 5.9), the COHERENT control model for D2D is introduced. Then the first detailed proposal of the abstraction approach for D2D management is defined, including how measurements could be used to populate a network graph for D2D at the level of the local real time controller and of the C3 control instance. This work can be further refined in the second deliverable if necessary and can also be used as an input for the work described in Section 5.9 which is also based on D2D. Then, graph-based interference coordination for D2D relaying is studied further by simulation on a Manhattan grid

scenario. D2D relaying with graph-based interference coordination improves cell-edge performance considerably, which shows the advantage of the COHERENT network graph application in the D2D networks.

Section 5.9 presents the state of the art of D2D inside 3GPP especially at low layers, which is complementary to the work presented in Section 5.8. The state of the art was focused on current 3GPP implementation of D2D because it provides practical values of the measurement reports which may be important information for characterizing how the low-layers may be modeled based on this information and for identifying missing but useful measurements. Moreover, 3GPP is the place where 5G will most probably be defined, and D2D as proposed by 3GPP will be part of it. While the state of the art review of 3GPP was done for both in-coverage, out-of-coverage D2D communications as well as coverage extension based on D2D, future work will focus on whether it is necessary or not to introduce more control (handled inside the COHERENT control framework) especially for supporting PMR services for the coverage extension use case and in the out-of-coverage case. The interesting point will be to understand how much control should be introduced and how much will the overhead be in terms of latency.

Section 5.10 deals with the problem of active state mobility management in 3GPP HetNets. Dense mobile network deployment with multi-connectivity makes it difficult to decide on handover parameters. In this context, a centralized control approach with network view can help to control those parameters dynamically. Based on the abstracted RAN information, the required parameters can be stored in the network graph of C3 architecture and the higher-layer application utilizes these parameters for managing the user mobility and sends the modified parameters based on its algorithm for various purposes, e.g., load, traffic, power, etc. This work, to be finalized in the second part of the project, aims at the integration of LTE's mobility control functionality into the COHERENT control framework. This may include abstraction of LTE's mobility-related parameters and measured entities; centralisation of the currently decentralised handover decision algorithms; and identification of the basic biasing features in current LTE SON features and the search for optimal biasing factors.

Section 5.11 deals with mobility issues for high speed platforms like high speed trains and airplanes. Note that operators of such transportation systems are looking toward the use of 4G and 5G technologies as opposed to the more-expensive and higher-latency option represented by satellite communications. Moreover, direct communication from ground to the airplane has been identified by SA SMARTER as one of the use cases for 5G (see Section 5.11 for other relevant SMARTER use cases). In this deliverable we presented the use cases, gave an estimation of the number of eNodeB needed to cover the Europe, compared the satellite and terrestrial 4G and 5G communication systems at high level, and evaluated the most important KPIs. Future work will deal with abstraction metrics and with implementation schemes in order to solve the mobility issues described in this document inside the COHERENT control framework.

6.2 Future work needed in D3.2

The purpose of this deliverable D3.1 (First report on physical and MAC layer modelling and abstraction) is to present the initial results on modelling and abstraction of HMNs from the studies undertaken thus far. The complete results from all considered use-cases will be presented in a follow-up deliverable D3.2 (Final report on physical and MAC layer modelling and abstraction).

In this stage, the following topics, which need further attention in D3.2, have been recognized (see also Annex):

- Coordinated multipoint techniques
- Network graphs for high level control
- LTE X2 protocol enhancement
- Traffic steering and load balancing
- Common COHERENT interface for D2D
- D2D discovery, clustering, load balancing, and relay selection
- D2D relay standardization

- Cognitive radios
- Fast handover and discovery algorithms
- Cooperative moving relays, coverage extension, and link-to-system mapping
- Interface to high level programmable framework for coverage extension and mobility management

References

1. Alcatel-Lucent. 9900 wireless network guardian. http://www.alcatel-lucent.com/products/9900-wireless-network-guardian, 2013.
2. C. Cranor, T. Johnson, O. Spataschek, and V. Shkapenyuk, Gigascope: a stream database for network applications. In Proceedings of the 2003 ACM SIGMOD international conference on Management of data (New York, NY, USA, 2003), SIGMOD '03, ACM, pp. 647–651.
3. Alcatel-Lucent. Motive big network analytics, 2014 http://networks.nokia.com/portfolio/solutions/telecom-analytics.
4. Verizon adds cloudera's cloud-based big data analytics solution to verizon cloud ecosystem. http://www.verizon.com/about/news/verizon-adds-clouderas-cloudbased-big-data-\analytics-solution-verizon-cloud-ecosystem/, 2013.
5. P. Cardieri, "Modeling interference in wireless ad hoc networks," Communications Surveys Tutorials, IEEE, vol. 12, no. 4, pp. 551–572, April 2010.
6. D. Niculescu, "Interference Map for 802.11 Networks," in Proc. of ACM IMC, 2007.
7. C. Reis, R. Mahajan, M. Rodrig, D. Wetherall, and J. Zahorjan, "Measurement-based models of delivery and interference in static wireless networks," in Proc. of ACM SigComm, 2006.
8. N. Ahmed, U. Ismail, S. Keshav, and K. Papagiannaki, "Online Estimation of RF Interference," in Proc. of ACM CoNEXT, 2008.
9. V. Shrivastava, S. Rayanchu, S. Banerjee, and K. Papagiannaki, "PIE in the Sky: Online Passive Interference Estimation for Enterprise WLANs," in Proc. of USENIX NSDI, 2011.
10. M. Vutukuru, K. Jamieson, and H. Balakrishnan, "Harnessing exposed terminals in wireless networks," in Proc. of USENIX NSDI, 2008.
11. R. Mahajan, M. Rodrig, D. Wetherall, and J. Zahorjan, "Analyzing the MAC-level Behavior of Wireless Networks in the Wild," SIGCOMM Comput. Commun. Rev., vol. 36, no. 4, pp. 75–86, Aug. 2006.
12. Y.-C. Cheng, J. Bellardo, P. Benkö, A. C. Snoeren, G. M. Voelker, and S. Savage, "Jigsaw: Solving the puzzle of enterprise 802.11 analysis," in Proc of ACM SigComm, 2006.
13. D. Giustiniano, D. Malone, D. Leith, and K. Papagiannaki, "Measuring transmission opportunities in 802.11 links," Networking, IEEE/ACM Transactions on, vol. 18, no. 5, pp. 1516–1529, Oct 2010.
14. S. Rayanchu, V. Shrivastava, S. Banerjee, and R. Chandra, "Fluid: Improving throughputs in enterprise wireless lans through flexible channelization," in Proc. ACM MobiCom, 2011.
15. D2.1, Use cases and architecture, COHERENT project deliverable, Jan. 2016.
16. Y. Lee et al., "Recent advances in radio resource management for heterogeneous LTE/LTE-A networks," IEEE Commun. Surveys & Tutorials, vol. 16, no. 4, 2014, pp. 2142-2180.
17. M. Peng et al., "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," IEEE Wireless Commun., Dec. 2014, pp. 126-135.
18. D. Xenakis et al., "Mobility management for femtocells in LTE-Advanced: key aspects and survey of handover decision algorithms," IEEE Commun. Surveys & Tutorials, vol. 16, no. 1, 2014, pp. 64-91.
19. A. Ahmed et al., "Enabling vertical handover decisions in heterogeneous wireless networks: a state-of-the-art and a classification," IEEE Commun. Surveys & Tut., vol. 16, no. 2, 2014, pp. 776-811.
20. F. Gapozzi et al., "Downlink packet scheduling in LTE cellular networks: key design issues and a survey," IEEE Commun. Surveys & Tutorials, 2013, pp. 678-700.
21. G. Ku et al., "Resource allocation and link adaptation in LTE and LTE Advanced: A tutorial," IEEE Commun. Surveys & Tutorials, 2015, pp. 1605-1633.

22. N. Feng et al., "Pricing and power control for joint network-centric and user-centric radio resource management," IEEE Trans. Commun., vol. 52, no. 9, Sep. 2004, pp. 1547-1557.
23. G Song et al., "Utility-based resource allocation and scheduling in OFDM-based wireless broadband network," IEEE Commun. Mag., Dec. 2005, pp. 127-134.
24. L. Tan et al., "Utility maximization resource allocation in wireless networks: methods and algorithms," IEEE Trans. Systems, Man and Cybernetics, vol. 45, no.7, Jul. 2015, pp. 1018-1034.
25. X.-F. Qi, "5G as a user-centric network," IEEE 5G Summit, May 2015.
26. R. Stankiewicz et al., "A survey of QoE assurance in converged networks," Computer Networks, vol. 55, no. 7, May 2011, pp. 1459-1473.
27. K. Zheng et al., "10 Gb/s HetSNets with millimetre-wave communications: access and networking –Challenges and protocols," IEEE Commun. Mag., Jan 2015, pp. 222-231.
28. M. Shariat et al., "Scheduling as an important cross-layer operation for emerging broadband wireless systems," IEEE Commun. Surveys & Tutorials, 2009, pp. 74-86.
29. G. Auer et al., "How much energy is needed to run a wireless network?," IEEE Wireless Commun., vol 18, no. 5, Oct. 2011, pp. 40-49.
30. E. Björnson et al., "Multiobjective signal processing optimization: The way to balance conflicting metrics in 5G systems," IEEE Signal Processing Mag., Nov. 2014., pp. 14-23.
31. A. Hamza et al., "A survey on inter-cell interference coordination techniques in OFDMA-based cellular networks," IEEE Commun. Surveys & Tutorials, vol. 15, no. 4, 2013, pp. 1642-1670.
32. E. Pateromichelakis et al., "On the evolution of multi-cell scheduling in 3GPP LTE/LTE-A," IEEE Commun. Surveys & Tut., vol. 15, no. 2, 2013, pp. 701-717.
33. P. Rost et al., "Cloud technologies for flexible 5G radio access networks," IEEE Commun. Mag., May 2014, pp. 68-76.
34. T. Chen et al., "SoftMobile: Control evolution for future heterogeneous mobile networks," IEEE Wireless Commun., Dec. 2014, pp. 70-78.
35. M. Peng et al., "Recent advances in underlay heterogeneous networks: interference control, resource allocation, and self-organization," IEEE Commun. Surveys & Tutorials, vol. 17, no. 2, 2015, pp. 700-729.
36. A. Liu et al., "Hierarchical radio resource optimization for heterogeneous networks with enhanced inter-cell interference coordination," IEEE Trans. Sig. Process., vol. 62, no. 7, Apr. 2014.
37. N. Wang et al., "Backhauling 5G small cells: a radio resource management perspective," IEEE Wireless Commun., Oct. 2015, pp. 41-49.
38. R.-A. Pitaval et al., "Full-duplex self-backhauling for small-cell 5G networks," IEEE Wireless Commun., Oct. 2015, pp. 83-89.
39. D. Lopez-Perez et al., "Towards 1 Gbps/UE in cellular systems: understanding ultra-dense small cell deployments," IEEE Commun. Surveys & Tutorials, vol. 17, no. 4, 2015, pp. 2078-2101.
40. O. Aliu et al., "A survey of self-organization in future cellular systems," IEEE Commun. Surveys & Tut., vol. 15, no. 1, 2013, pp. 112-119.
41. A. Mohamed et al., "Control-data separation architecture for cellular radio access networks: a survey and outlook, " IEEE Commun. Surveys & Tutorials, to be published.
42. D. Kreutz et al., "Software-defined networking: a comprehensive survey," Proc. IEEE, vol 103, no.1, Jan 2015, pp. 14-76.
43. Y. Kishiyama et al., "Future steps of LTE-A: Evolution toward integration of local area and wide area systems," IEEE Wireless Commun., Feb. 2013, pp. 12-18.
44. S.-Y. Lien et al., "Ultra-low-latency ubiquitous connections in heterogeneous cloud radio access networks," IEEE Wireless Commun., Jun. 2015, pp. 22-31.

45. S. Zhang et al., “How many small cells can be turned off via vertical offloading under a separation architecture,” IEEE Trans. Wireless Commun., vol 14, no. 10, Oct. 2015, pp. 5440-5453.
46. X. Jin et al., “SoftCell: Scalable and Flexible Cellular Core Network Architecture,” 9th ACM Conf. Emerging Networking Experiments and Technologies, 2013, pp. 163–74.
47. A. Gudipati et al., “SoftRAN: Software Defined Radio Access Network,” 2nd ACM SIGCOMM Wksp. Hot Topics in Software Defined Networking, 2013, pp. 25–30.
48. K. Pentikousis et al., “MobileFlow: Toward software-defined mobile networks,” IEEE Commun. Mag., Jul. 2013, pp. 44-53.
49. S. Khan et al., “Application-driven cross-layer optimization for video streaming over wireless networks,” IEEE Commun. Mag., Jan. 2006, pp. 122-130.
50. F. Foukalas et al., “Cross-layer design proposals for wireless mobile networks: a survey and taxonomy,” IEEE Commun. Surveys & Tut., vol. 10, no. 1, 2008, pp. 70-85.
51. R. Riggio et al., “Programming abstractions for software-defined wireless networks,” IEEE Trans. Network Service Manag., vol. 12, no. 2, 2015, pp. 146-162.
52. D. Gajski, Principles of Digital Design, Prentice Hall, 1996.
53. A. Jerraya et al., “Programming models and HW-SW interfaces abstraction for multi-processor SoC,” in Proc. DAC, 2006, pp. 280-285.
54. K. Brueninghaus et al., “Link performance models for system level simulations of broadband radio access,” in Proc. IEEE PIMRC, 2005, pp. 2306-2311.
55. C. Liang et al., “Wireless network virtualization: a survey, some research issues and challenges,” IEEE Commun. Surveys & Tut., vol. 17, no. 1, 2015, 358-380.
56. F. Hu et al., “A survey on software-defined network and OpenFlow: From concept to implementation,” IEEE Commun. Surveys & Tut., vol. 16, no. 4, 2014, pp. 2181-2206.
57. F. Ganz et al., “A practical evaluation of information processing and abstraction techniques for the Internet of things,” IEEE Internet of Things Journal, vol., 2, no. 4, Aug. 2015, pp. 340-354.
58. S. Marvin et al., “A unified approach to the probability of error for noncoherent and differentially coherent modulations over generalized fading channels,” IEEE Trans. Commun., vol. 46, no. 12, Dec. 1998, pp. 1625-1638.
59. H. Bai et al., “Error modeling schemes for fading channels in wireless communications: A survey,” IEEE Commun. Surveys & Tut., vol. 5, no. 2, pp. 2-9.
60. H. ElSawy et al., “Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey,” IEEE Commun. Surveys & Tut., vol. 15, no. 3, 2013, pp. 996-1019.
61. M. Win et al., “A mathematical theory of network interference and its applications,” Proc. IEEE, vol. 97, no. 2, Feb. 2009, pp. 205-230.
62. K. Gulati et al., “Statistics of co-channel interference in a field of Poisson and Poisson-Poisson clustered interferers,” IEEE Trans. Sig. Processing, vol. 58, no. 12, Dec. 2010, pp. 6207-6222.
63. J. Andrews et al., “A tractable approach to coverage and rate in cellular networks,” IEEE Trans. Commun., vol 59, no. 11, Nov. 2011, pp. 3122-3134.
64. A. Iyer et al., “What is the right model for wireless channel interference,” IEEE Trans. Wireless Commun., vol. 8, no. 5, May 2009, pp. 2662-2671.
65. F. Kaltenberger et al., “On scalability, robustness, and accuracy of physical layer abstractions for large-scale system-level evalutions of LTE networks,” in Proc. IEEE ASILOMAR, 2013, pp. 1644-1648.
66. J. Ikuno et al., “System level simulation of LTE networks,” in Proc. IEEE VTC, 2010, pp. 1-5.
67. B. Bjerke et al., “Packet error probability prediction for system level simulations MIMO-OFDM based 802.11n WLANs,” in Proc. IEEE ICC, 2005, pp. 2538-2542.

68. T. Jensen et al., “Fast link adaptation for MIMO OFDM,” IEEE Trans. Veh. Tech., vol. 59, no. 8, Oct. 2010, pp. 3766-3778.
69. Z. Hanzaz et al., “Analysis of effective SINR mapping models for MIMO OFDM in LTE system,” in Proc. IWCMC, 2013, pp. 1509-1515.
70. A. Cipriano et al., “Calibration issues of PHY layer abstractions for wireless broadband systems,” in Proc. IEEE VTC, Sep. 2008, pp. 1-5.
71. P. Mogensen et al., “LTE capacity compared to the Shannon bound,” in Proc. IEEE VTC, Apr. 2007, pp. 1234-1238.
72. P. Munoz et al., “Estimation of link-layer quality parameters in a system-level LTE simulator,” in Proc. IB2Com, 2010.
73. P. Tan et al., “Link adaptation based on adaptive modulation and coding for multiple-antenna OFDM system,” IEEE J. Sel. Areas Commun., vol 26, no. 8, Oct. 2008, pp. 1599-1606.
74. D. Martin-Sacristan et al., “LTE-advanced system level simulation platform for IMT-Advanced evaluation,” 2011.
75. Q. Liu et al., “Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links,” IEEE Trans. Wireless Commun., vol. 3, no. 5, Sep. 2004, pp. 1746-1755.
76. D. Dechene et al., “Energy-aware resource allocation strategies for LTE uplink with synchronous HARQ constraints,” IEEE Trans. Mobile Comp., vol. 13, no. 2, Feb. 2014, pp. 422-433.
77. K. Zheng et al., “Cross-layer queueing analysis on multihop relaying networks with adaptive modulation and coding,” IET Commun., vol. 4, no. 3, 2010, pp. 295-302.
78. Q. Liu et al., “Queueing with adaptive modulation and coding over wireless links: cross-layer analysis and design,” IEEE Trans. Wireless Commun., vol. 4, no. 3, May. 2005, pp. 1142-1153.
79. R.Y. Chang, Z. Tao, Jinyun Zhang, and C.-C.J. Kuo, "Multicell OFDMA Downlink Resource Allocation Using a Graphic Framework," IEEE Transactions on Vehicular Technology,vol.58, no.7, pp.3494-3507, Sept. 2009
80. A. Akhtar et al., “Cooperative ARQ-based energy efficient routing in multihop wireless networks,” IEEE Trans. Veh. Tech., vol. 64, no. 11, Nov. 2015, pp. 5187-5197.
81. H. Yilmaz et al., “Radio environment map as enabler for practical cognitive radio networks,” IEEE Commun. Mag., Dec. 2013, pp. 162-169.
82. J. Perez-Romero et al.,”On the use of radio environment maps for interference management in heterogeneous networks, ” IEEE Commun. Mag., Aug. 2015, pp. 184-191.
83. Wei Ni et al., “Graph theory and its applications to future network planning: software-defined online small cell management,” IEEE Wireless Commun., Feb. 2015, pp. 52-60.
84. H. Bogucka, P. Kryszkiewicz, A. Kliks, "Dynamic spectrum aggregation for future 5G communications," in Communications Magazine, IEEE , vol.53, no.5, pp.35-43, May 2015.
85. 3GPP TS 36.300 v13.1.0; 3rd Generation Partnership Project; 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 13)
86. 3GPP TS 36.423 v13.1.0; 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 application protocol (X2AP) (Release 13)
87. 3GPP TS 36.101 v13.2.0; 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 13)
88. 3GPP TS 36.213 v.13.0.1; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (Release 13)

89. 3GPP TS 36.211 v.12.8.0; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (Release 12)
90. 3GPP TS 36.331 v.12.8.0; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification (Release 12)
91. Ning, Xuan Jie, et al. "A Link Evaluation Method Employing Statistical Means of Received Signal Strength Indicator and Link Quality Indicator for Wireless Sensor Networks." Applied Mechanics and Materials. Vol. 470. 2014.
92. Shang, Fengjun, et al. "A location estimation algorithm based on RSSI vector similarity degree." International Journal of Distributed Sensor Networks 2014 (2014).
93. ZhongPeng, Liu, and Liu Li Juan. "Bayesian Optimization RSSI and Indoor location Algorithm of Iterative Least Square." International Journal of Smart Home 9.6 (2015): 31-42.
94. iJOIN: Interworking and JOINt Design of an Open Access and Backhaul. Network Architecture for Small Cells based on Cloud Networks.
95. Zheng, Guanbo, et al. "A link quality inference model for ieee 802.15. 4 low-rate wpans." Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE. IEEE, 2011.
96. H. Yanikomeroglu and E. S. Sousa, “CDMA distributed antenna system for indoor wireless communications,” in Proc. IEEE Int. Conf. Univ.Pers. Commun., Oct. 1993, vol. 2, pp. 990–994.
97. P. Chow, A. Karim, V. Fung, and C. Dietrich, “Performance advantages of distributed antennas in indoor wireless communication systems,” in IEEE Veh. Technol. Conf. (VTC), Jun. 1994, vol. 3, pp. 1522–1526.
98. K. J. Kerpez and S. Ariyavisitakul, “A radio access system with distributed antennas,” in Proc. IEEE Global Telecommun. Conf.(GLOBECOM), Nov.–2 Dec. 1994, vol. 3, pp. 1696–1700.
99. G.-H. Chen, C.-M. Yu, and C.-C. Huang, “A simulation study of a distributed antenna-based CDMA system,” in Proc. IEEE Symp. Pers., IndoorMobile Radio Commun. (PIMRC), Oct. 1996, vol. 2, pp. 517–521.
100. A. Obaid and H. Yanikomeroglu, “Reverse-link power control in CDMA distributed antenna systems,” in Proc. IEEE Wireless Commun. Netw. Conf. (WCNC), 2000, vol. 2, pp. 608–612.
101. W. Choi and J. G. Andrews, “Downlink performance and capacity of distributed antenna systems in a multicell environment,” IEEE Trans.Wireless Commun., vol. 6, no. 1, pp. 69–73, Jan. 2007.
102. P. Jung, B. Steiner, and B. Stilling, “Exploitation of intracell macro-diversity in mobile radio systems by deployment of remote antennas,” in Proc. IEEE Int. Symp. Spread Spectrum Techn. Appl. (ISSSTA), Sep.1996, vol. 1, pp. 302–307.
103. M. V. Clark, T. M. I. Willis, L. J. Greenstein, A. J. J. Rustako, V. Erceg, and R. S. Roman, “Distributed versus centralized antenna arrays in broadband wireless networks,” in Proc. IEEE Veh. Technol. Conf.(VTC), Rhodes, May 2001, vol. 1, pp. 33–37.
104. I. Toufik and R. Knopp, “Wideband channel allocation in distributed antenna systems,” in Proc. IEEE Veh. Technol. Conf. (VTC), Sep. 2006, pp. 1–5.
105. T. Wu and P. Hosein, “Radio resource management strategies for distributed antenna systems,” in Proc. IEEE Wireless Commun. Netw. Conf. (WCNC), Apr. 2010, pp. 1–6.
106. X. Li, M. Luo, M. Zhao, L. Huang, and Y. Yao, “Downlink performance and capacity of distributed antenna system in multi-user scenario,” in WiCom, 2009, pp. 1–4.
107. S. D. Peethala, T. Kreul and T. Kaiser, “Real-time testbed for validating distributed antenna scenarios,”, 20th International ITG Workshop on Smart Antennas March 9-11, 2016 in Munich, Germany (accepted, to be published)
108. M, Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, M. Tanno, "Coordinated multipoint transmission/reception techniques for LTE-advanced [Coordinated and Distributed MIMO]," in Wireless Communications, IEEE , vol.17, no.3, pp.26-34, June 2010

109. P. Kryszkiewicz, H. Bogucka, "Dynamic determination of spectrum emission masks in the varying cognitive radio environment," in Communications (ICC), 2013 IEEE International Conference on , vol., no., pp.2733-2737, 9-13 June 2013.
110. P. Kryszkiewicz, H. Bogucka, "Advanced interference reduction in NC-OFDM based Cognitive Radio with Cancellation Carriers," in Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European , vol., no., pp.571-575, 1-5 Sept. 2014.
111. J. G. Andrews, "Seven ways that hetnets are a cellular paradigm shift," IEEE Commun. Mag., vol. 51, no. 3, pp. 136–144, 2013.
112. J. Andrews, S. Singh, S, Y. Qiaoyang, L. Xingqin, H. Dhillon, "An overview of load balancing in hetnets: old myths and open problems," IEEE Wireless Commun., vol.21, no.2, pp.18-25, April 2014.
113. S. Singh, H. S. Dhillon, and J. G. Andrews, "Offloading in Heterogeneous Networks: Modeling, Analysis, and Design Insights," IEEE Trans. Wireless Commun., vol.12, no. 5, May 2013, pp. 2484–97.
114. Q. Ye et al., "User Association for Load Balancing in Heterogeneous Cellular Networks," IEEE Trans. Wireless Commun., vol. 12, no. 6, June 2013, pp. 2706–16.
115. G. Boudreau et al., "Interference Coordination and Cancellation for 4G Networks," IEEE Commun. Mag., 2009.
116. vD. Lopez-Perez et al., "Enhanced Intercell Interference Coordination Challenges in Heterogeneous Networks," IEEE Wireless Commun., vol. 18, June 2011, pp. 22–30.
117. A. Damnjanovic et al., "A Survey on 3GPP Heterogeneous Networks," IEEE Wireless Commun., vol. 18, no. 3, June 2012, pp. 10–21.
118. E. Stevens-Navarro, Y. Lin, and V. Wong, "An MDP-Based Vertical Handoff Decision Algorithm for Heterogeneous Wireless Networks," IEEE Trans. Vehic. Tech., vol. 57, Mar. 2008, pp. 1243–54.
119. E. Aryafar et al., "RAT Selection Games in HetNets," Proc. IEEE INFOCOM, Apr. 2013.
120. D. Niyato and E. Hossain, "Dynamics of Network Selection in Heterogeneous Wireless Networks: An Evolutionary Game Approach," IEEE Trans. Vehic. Tech., vol. 58, no. 4, May 2009, pp. 2008–17.
121. P. Kreuger, et al. "Autonomous load balancing of heterogeneous networks." Vehicular Technology Conference (VTC Spring), 2015 IEEE 81st. IEEE, 2015.
122. I. Siomina and D. Yuan, "Load balancing in heterogeneous lte: Range optimization via cell offset and load-coupling characterization," in Communications (ICC), 2012 IEEE International Conference on, June 2012, pp. 1357–1361.
123. H. Wang, L. Ding, P. Wu, Z. Pan, N. Liu, and X. You, "Dynamic load balancing and throughput optimization in 3gpp lte networks," in Proceedings of the 6th International Wireless Communications and Mobile Computing Conference, ser. IWCMC '10. New York, NY, USA: ACM, 2010, pp. 939–943.
124. P. Fotiadis, M. Polignano, D. Laselva, B. Vejlgaard, P. Mogensen, R. Irmer, and N. Scully, "Multi-layer mobility load balancing in a heterogeneous lte network," in Vehicular Technology Conference (VTC Fall), 2012 IEEE, Sept 2012, pp. 1–5.
125. A. Lobinger, S. Stefanski, T. Jansen, and I. Balan, "Load balancing in downlink lte self-optimizing networks," in Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st, May 2010, pp. 1–5.
126. R. Zhang, X. Cheng, L. Yang, and B. Jiao, "Interference-aware graph based resource sharing for device-to-device communications underlaying cellular networks," in Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '13), pp. 140–145, Shanghai, China, April 2013.

127. H. Zhang, T. Wang, L. Song, Z. Han, "Graph-based resource allocation for D2D communications underlaying cellular networks," in Proceedings of 1st IEEE International Workshop on Device-to-Device Communications and Networks, pp.187-192, Aug 2013.
128. B. Guo, S. Sun, and Q. Gao, "Graph-based resource allocation for D2D communications underlying cellular networks in multiuser scenario," International Journal of Antennas and Propagation, vol.2014.
129. M. Mezzavilla, K. Somasundaram, M. Zorzi, "Joint user association and resource allocation in UE-Relay assisted heterogeneous networks," in Proceedings of the IEEE International Conference on Communications Workshops (ICC'14), pp.628-634, June 2014.
130. Z. Uykan, R. Jantti, "Transmission-Order Optimization for Bidirectional Device-to-Device (D2D) Communications Underlaying Cellular TDD Networks—A Graph Theoretic Approach," IEEE Journal on Selected Areas in Communications, vol.34, no.1, pp.1-14, Jan. 2016.
131. K. W. Choi, Z. Han, "Device-to-Device Discovery for Proximity-Based Service in LTE-Advanced System," IEEE Journal on Selected Areas in Communications, vol.33, no.1, pp.55-66, Jan. 2015
132. D. Panaitopol, C. Mouton, B. Lecroart, Y. Lair and P. Delahaye, "Recent Advances in 3GPP Rel-12 Standardization related to D2D and Public Safety Communications," arXiv:1505.07140, May 2015.
133. 3GPP TS 23.122, V13.0.0, "Non-Access-Stratum (NAS) functions related to Mobile Station (MS) in idle mode," Dec. 2014.
134. 3GPP TS 36.304, V13.0.0, "User Equipment (UE) procedures in idle mode," Jan. 2016.
135. J. Salo, Mobility Parameter Planning for 3GPP LTE: Basic Concepts and Intra-Layer Mobility, June 2013.
136. D. Xenakis, N. Passas, L. Merakos, and C. Verikoukis, Mobility Management for Femtocells in LTE-Advanced: Key Aspects and Survey of Handover Decision Algorithms, IEEE Communications Surveys & Tutorials, Vol. 16, No. 1, First Quarter 2014.
137. P. Kela, J. Turkka, M. Costa, Borderless Mobility in 5G Outdoor Ultra-Dense Networks, IEEE Access, vol.3, no., pp.1462-1476, 2015.
138. R. Riggio, M.K. Marina, J. Schulz–Zander, T. Rasheed, S. Kuklinski, Programming Abstractions for Software–Defined Wireless Networks, IEEE Transactions on Network and Service Management, vol.12, no.2, pp.146-162, June 2015.
139. Systems and software engineering – vocabulary. ISO/IEC/IEEE 24765 International standard.
140. D2.2, System architecture and abstractions for mobile networks, COHERENT project deliverable, 2016.
141. Vlavianos, Angelos, et al. "Assessing link quality in IEEE 802.11 wireless networks: which is the right metric?." Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on. IEEE, 2008.
142. Ning, Xuan Jie, et al. "A Link Evaluation Method Employing Statistical Means of Received Signal Strength Indicator and Link Quality Indicator for Wireless Sensor Networks." Applied Mechanics and Materials. Vol. 470. 2014.
143. Shang, Fengjun, et al. "A location estimation algorithm based on RSSI vector similarity degree." International Journal of Distributed Sensor Networks 2014 (2014).
144. ZhongPeng, Liu, and Liu Li Juan. "Bayesian Optimization RSSI and Indoor location Algorithm of Iterative Least Square." International Journal of Smart Home 9.6 (2015): 31-42.
145. Zheng, Guanbo, et al. "A link quality inference model for ieee 802.15. 4 low-rate wpans." Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE. IEEE, 2011.
146. Zhang, Jian, and Ivan Marsic. "Link quality and signal-to-noise ratio in 802.11 WLAN with fading: A time-series analysis." Vehicular Technology Conference, 2006. VTC-2006 Fall. 2006 IEEE 64th. IEEE, 2006.

147. Guha, Ratul K., and Saswati Sarkar. "Characterizing temporal SNR variation in 802.11 networks." Vehicular Technology, IEEE Transactions on 57.4 (2008): 2002-2013.
148. Halperin, Daniel, et al. "Predictable 802.11 packet delivery from wireless channel measurements." ACM SIGCOMM Computer Communication Review 41.4 (2011): 159-170.
149. Luo, Jiayou, and Xingqun Zhan. "Characterization of smart phone received signal strength indication for WLAN indoor positioning accuracy improvement." Journal of Networks 9.3 (2014): 739-746.
150. Kreuger, P. and Steinert, R. "Scalable in-network rate monitoring". Integrated Network Management (IM 2015), 2015 IFIP/IEEE International Symposium on, May 2015.
151. 3GPP R2-104870, "Pull based RAN overload control," August 2010.
152. 3GPP TR 23.720 V13.0.0, "Study on architecture enhancements for Cellular Internet of Things," March 2016.
153. S. Andreev et al., "Intelligent access network selection in converged multi-radio heterogeneous networks," in IEEE Wireless Communications, vol. 21, no. 6, pp. 86-96, December 2014.
154. "Preliminary Views and Initial Considerations on 5G RAN Architecture and Functional Design", METIS II White Paper, March 2016.
155. 3GPP TS 36.216, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer for relaying operation", v10.3.0, September 2011.
156. 3GPP TR 36.836, "Technical specification group radio access network: mobile relay for evolved universal terrestrial radio access (E-UTRA). Technical report. http://www.3gpp.org
157. A. Asadi, Q. Wang, and V. Mancuso, "A survey on Device-to-Device communication in cellular networks", IEEE Communication Surveys & Tutorials, vol. 16, n. 4, p: 1801-1819, fourth quarter 2014.
158. Rohde&Schwarz, "LTE-Advanced (3GPP Rel. 12) Technology Introduction", Whitepaper, August 4th 2015.
159. Rohde&Schwarz, "Device to device communications in LTE", Whitepaper, September 29th 2015.
160. Dino Flore, "LTE Release 13 and road to 5G", presentation at the ATIS 5G Symposium, June 2015.
161. 4G Americas, "Executive Summary: Inside 3GPP Release 13", September 2015.
162. R. Liebhart, D. Chandramouli, C. Wong, J. Merkel, *LTE for public safety*, Wiley, 2015.
163. 3GPP TS 36.211, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation", v13.1.0, March 2013.
164. 3GPP TS 36.212, "Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding", v13.1.0, March 2013.
165. 3GPP TS 36.213, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures", v13.0.0, December 2015.
166. 3GPP TS 36.214, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer; Measurements", v13.1.0, March 2016.
167. 3GPP TS 36.321, "Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification", v13.1.0, March 2016.
168. 3GPP TS 36.331, "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification", v13.1.0, March 2016.
169. SHARING D5.2, "Device-to-device communications innovations', Celtic-plus SHARING deliverable, 12/02/2016.

170. S. Sesia, I. Toufik and M. Baker, *LTE - Long Term Evolution: From Theory to Practice*, Second Edition, John Wiley & Sons, 2011.
171. Technical Report WG SA1, 3GPP TR 22.891, Release-14, "Study on New Services and Markets Technology Enablers", v14.0.0, Mars 2016, available at http://www.3gpp.org/DynaReport/22891.htm.
172. Technical Report WG SA1, 3GPP TR 22.863, Release-14, "FS_SMARTER - enhanced Mobile Broadband", v1.0.0, Mars 2016.
173. Technical Specification WG RAN1, 3GPP TS 26.213, Release-9, « Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures », v9.2.0, June 2010.
174. Report ITU-R M.2282-0, "Systems for public mobile communications with aircraft M Series Mobile, radio determination, amateur and related satellite services", Dec. 2013.
175. ETSI, "Digital cellular telecommunications system (Phase 2+);Radio transmission and reception," TS 5.05, March 1996.
176. CEPT Report 40, "Compatibility study for LTE and WiMAX operating within the bands 880-915 MHz / 925-960 MHz and 1710-1785 MHz /1805-1880 MHz (900/1800 MHz bands)," pp. 1 –81, November 2010.
177. T. Weiss, J. Hillenbrand, A. Krohn, and F. Jondral, "Mutual interference in OFDM-based spectrum pooling systems," in Vehicular Technology Conference, 2004. VTC 2004-Spring. 2004 IEEE 59th, vol. 4, 2004, pp. 1873–1877 Vol.4.
178. G. Piro et al., "HetNets powered by renewable energy sources," IEEE Internet Comp., 2013, pp. 32-39.
179. L. Cai et al., "Dimensioning network deployment and resource management in green mesh networks," IEEE Wireless Commun., pp. 58-65, Oct. 2011.
180. A. Anttonen et al., "Finite-horizon prediction of energy depletions in off-grid wireless networks," IEEE Trans. Veh. Tech., 2016, to be published.
181. J. Andrews et al., "What will 5G be?," IEEE J. Sel. Areas Commun., vol. 32, no. 6, Jun. 2014a, pp. 1065-1082.
182. Z. Hasan et al., "Green cellular networks: A survey, some research issues and challenges," IEEE Commun. Surveys & Tut., vol. 13, no. 4, pp. 524-540, 2011.
183. P. Pathak and R. Dutta, "A survey of network design problems and joint design approaches in wireless mesh networks," IEEE Commun. Surveys & Tut., vol. 13, no. 3, pp. 396–428, Third quarter 2011.
184. O. Tonguz et al., "The mathematical theory of dynamic load balancing in cellular networks," IEEE Trans. Mobile Comp., vol. 7, no. 12, Dec. 2008, pp. 1504-1518.
185. 3GPP TS 36.101, v9.24.0; 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception, Section 5.6.
186. 3GPP TS 36.211, v9.1.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation", Section 6.1.
187. 3GPP TS 36.213, v9.3.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures", Section 7.1.6.
188. Jan-Jaap van de Beek et al., "On Channel Estimation in OFDM Systems", IEEE Veh. Technology Conference, Vol.2, Page(s):815-819, Chicago IL USA, Sep. 1995.
189. CommAgility, Channel estimation and parameter estimation, Internal report.
190. Ericsson, System-level evaluation of OFDM – further considerations, 3GPP TSG-RAN WG 1 35, R1- 031303, November 17-21, 2003.
191. Intel, "Discussion on UE-UE channel model for studies on D2D proximity services," 3GPP R1-130924, Chicage, USA,, Apr. 2013.
192. W. I. Project, "Matlab SW Documentation of WIM 2 Model," Aug. 2008.

193. W. I. Project, “D1.1.1: WINNER II Interim Channel Models,” Nov. 2006.
194. S. Sesia, I. Toufik and B. Matthew, LTE The UMTS Long Term Evolution: From Theory to Practice, Second Edition, Wiley, 2011.
195. W. Jang, L. Goratti, A. Hourani, O. Araft and S. Kandeepan, “Initial Approaches for Opportunistic Relaying for Disasetr Relief and temporary Events,” FP7 ABSOLUTE Project, 2013.
196. S. Cotton, “A statistical characterization of shadowed device-to-device communications in an indoor environment,” in IEEE 8th European Conference on Antennas and Propagation (EuCAP), 2014.
197. S. Cotton, “Channel measurements of device-to-device communications in an urban outdoor environment,” in IEEE XXXIth General Assembly and Scientific Symposium (URSI GASS), 2014.
198. S. Cotton, “Human Body Shadowing in Cellular Device-to-Device Communications: Channel Modeling Using the Shadowed k-mew Fading Model,” IEEE Journal on Selected Areas in Communications, vol. PP, no. 99, pp. 1-9, 2015.
199. A. Al-Hourani, S. Chandrasekharan and S. Kandeepan, “Path Loss Sudy for Millimeter Wave Device-to-Device Communications in Urban Environment,” in IEEE International Conference on Communications Workshops (ICC), Sydney, 2014.
200. T. Rappaport, Wireless Communication: Principles and Practice, New Jersey: Prentice Hall, 1996.
201. M. Gudmundson, “Correlation model for shadow fading in mobile radio systems,” IEEE Electronic Letters, vol. 27, no. 23, pp. 2145-2146, 1991.
202. J. Karedal, F. Tufvesson, N. Czink, A. Paier, C. Dumard, T. Zemen, C. Mecklenbrauker and A. Molisch, “A Geometry-based stochastic MIMO model for Vehicle-to-Vehicle communications,” IEEE Transactions on Wireless Communications, vol. 8, no. 7, pp. 3646-3657, July 2009.
203. P. Kreuger, R. Steinert, S. Liu, J. Ekman. "Scalable high precision rate monitoring and congestion detection". Submitted to CNSM'16.
204. R. Steinert and A. Hess, "Observing Software-defined Networks Using a Decentralized Link Monitoring Approach", Proc. IEEE NetSoft, London, UK, April 2015.
205. 3GPP TS 23.303, "Technical specification group Services and System aspects; Proximity-Based services (ProSe); Stage 2", v13.3.0, March 2016.
206. V. Nguyen et al., “SDN and virtualisation-based LTE mobile network architectures: A comprehensive survey,” Wireless Personal Communications, vol. 86, no. 3, pp 1401-1438, 2016.
207. IEEE Std 802.21-2008, IEEE Standard for Local and Metropolitan Area Networks, Part 21: Media Independent Handover Services, IEEE, January 2009.
208. ETSI TR 136 942 V10.3.0 “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios” July 2012.
209. CEPT Report 40: “Compatibility study for LTE and WiMAX operating within the bands 880-915 MHz / 925-960 MHz and 1710-1785 MHz / 1805-1880 MHz (900/1800 MHz bands)”, November 2010
210. ETSI TR 136 104 V12.9.0 “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception”, October 2015.
211. ETSI TR 136 101 V12.9.0 “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception”, October 2015.
212. Nokia Networks, “Optimizing mobile broadband performance by spectrum refarming”, Nokia white paper, 2014,

http://networks.nokia.com/sites/default/files/document/nokia_optimising_mobile_broadband_performance_by_spectrum_refarming_white_paper.pdf

213. TR-203, “Interworking between next generation fixed and 3GPP wireless networks,” BB forum, technical report, August 2012.

Annex: Objectives mapping of WP3

This annex represents the mapping of objectives of WP3 in each three technical tasks (T3.1, T3.2, and T3.3) as described in description of work (DoW) of COHERENT project. The tables below represent objectives and the topics studied in the respective tasks and where they have been addressed in D3.1. The "comments" column indicates the current plans of WP3 partners of the research work (on those objectives or topics that were not in the focus in D3.1) that will be addressed in the next deliverable D3.2.

Table B-1 DoW objectives to Sections in D3.1 mapping, Task T3.1 view

Objectives from DoW	D3.1 sections	Comments
To construct network graph for high level control	3.2	
To define variety of abstracted network graphs	3.2, 5	
To define edges in the network graphs - state of the link, channel quality (CQI), SINR etc.	3.2, 4.1, 4.2, 4.3, 5.2	
To define an enhancement to LTE X2 protocol for transportation of abstractions to a central coordinator		Will be addressed in D3.2
To identify and abstract key parameters at physical and MAC layer, to be used by high-level programmable control framework	4.1, 4.2, 4.3	
Interaction with other WPs like data base management, critical and non-critical aspects etc.	Partially addressed	Will be addressed in D3.2
Aggressive spatial reuse / Dense small cell deployment	5.3	
Advanced interference mitigation	5.5	
Coordination techniques	5.4 (SOA only)	Will be addressed in D3.2
Cell edge QoS	5.3	
Inter-cell radio resource allocation and ICIC		Will be addressed in D3.2
Traffic steering and load balancing	5.6	
Energy consumption, energy-limited HetNets	5.6	
Connectivity and quality of service	4.2	

Table B-2 DoW objectives to Sections in D3.1 mapping, Task T3.2 view

Objectives from DoW	Sections D3.1	Comments
Model and abstract D2D links to be integrated in abstracted network graphs for resource allocation and interference coordination between D2D links and BSs.	2.4, 4.2, 5.8	
Determine high-level descriptions of results of dynamic device clustering algorithms.		Will be addressed in D3.2
Define a common interface and parameters to integrate the control of D2D communications into high-level programmable framework.	5.8	Additional aspects to appear in D3.2
Develop new methods and algorithms, which utilize abstracted network graphs, to improve the performance of D2D communications and coexistence with BSs in different D2D cases.	5.8, 5.9	Additional aspects to appear in D3.2
Devise a local control scheme for D2D cooperation		Will be addressed in D3.2

Table B-3 DoW objectives to Sections in D3.1 mapping, Task T3.3 view

Objectives from DoW	D3.1 sections	Comments
To identify parameters to be included in abstracted network graph for coverage extension and mobility management support	4, 5.8, 5.9, 5.10	Further developments in D3.2
Investigate the cooperative relaying techniques for network coverage extension and the corresponding link-to-system mapping.	SoA for link-to-system mapping in Section 2.2.3	Will be addressed in D3.2
Develop fast neighbouring discovery mechanisms and fast handover algorithms based on abstracted network graph approach;		Will be addressed in D3.2
Develop mobility prediction algorithms and associated resource allocation scheme to enhance the small cell handover in LTE networks;		Will be addressed in D3.2
Develop the moving cell concept and integrate it with the proposed programmable control framework;	5.11	Further developments in D3.2
Moving relays	5.10	Further developments in D3.2
Coverage extension (D2D relay)	5.8, 5.9	Further developments in D3.2
Link-to-system mapping		Will be addressed in D3.2
Mobility management/ Handover	5.10	Further developments in D3.2
Traffic steering and load balancing	5.6	