Broadcast and Multicast Communication Enablers for the Fifth-Generation of Wireless Systems

Deliverable D3.4

RAT Protocols and Radio Resource Management in 5G-Xcast

Version v3.0

Date: 2019/07/31

Document properties:

Grant Number:	761498
Document Number:	D3.4
Document Title:	RAT Protocols and Radio Resource Management in 5G-Xcast
Editor(s):	Fasil Tesema, Volker Pauli (Nomor)
Authors:	Mikko Säily (NOK); David Vargas (BBC); Heikki Kokkinen (Fairspectrum); Volker Pauli, Fasil Tesema, Christiane Dietrich (Nomor); Peter Sanders (One2Many); Wei Guo, Belkacem Mouhouche (Samsung); Giorgio Calochira, Sergio Barberis, Flavio Muratore (TIM); Tero Jokela (TUAS); Zheng Chu, De Mi (UNIS), Eduardo Garro (UPV)
Reviewers	Gerry Foster (UNIS)
Contractual Date of Delivery:	2019/07/31
Dissemination level:	CO[1]
Status:	final version
Version:	3.0
File Name:	5G-Xcast D3.4_v3.0

Abstract

Evolution of broadcast / multicast vertical market sectors such as multimedia & entertainment, automotive, internet of things and public warning systems is pushing for a rapid growth of the wireless communication sectors that need to meet the technical requirements. While 3GPP's discussion on broadcast / multicast for 5G "New Radio" is at an early stage, the collaborative project 5G-XCast, under H2020 Phase II, has been working towards providing a comprehensive solution for a future generation of broadcast / multicast embedded efficiently into 5G communication networks. Focusing on the Radio Access Technology (RAT) protocol and Radio Resource Management (RRM), this deliverable document presents the 5G-Xcast solution aiming to, first, resolve RAT protocol limitations of the current 3GPP's LTE-A based broadcast / multicast systems that impose constraints on the RAT technical requirements documented in D2.1 [1] and D3.1 [3]; second, provide RRM strategies that are expected to fulfil the functional requirements described in 3GPP's study item TR 38.913 [2]. Furthermore, it includes the performed system-level simulator calibration and evaluations of 3GPP's "New Radio" that has been submitted to ITU as a candidate technology for IMT-2020.

[1] CO = Confidential, only members of the consortium (including the Commission Services)

PU = Public

Keywords

3GPP LTE-Advanced Pro, Point-to-Multipoint, Point-to-Point, Radio Access Network, 5G-Xcast RAN Technical Requirements, Link Level Simulations, System Level Simulations, Coverage Simulations, Media & Entertainment vertical, Public Warning vertical, Automotive vertical, IoT vertical, IMT-2020 evaluation, RAT protocols limitations, RAT protocol and radio resource management

Table of Contents

Executive Summary

3rd Generation Partnership Project (3GPP) has been working on the 5th Generation (5G) radio technology, which is also termed as New Radio (NR), on its release 15 specifications which were completed in March 2019. In the meantime, 5G evolution studies and work packages for release 16 have been approved in June 2018. The considered NR mile stones are identifying and developing technology components that enable fulfilment of the market needs and long-term evolution of the International Mobile Telecommunication for 2020 and beyond (IMT-2020) industries. The design on NR mainly focused on unicast communication; and 3GPP's discussions on broadcast and multicast communication are at an early stage. On the other hand, the 5G-Xcast project, under the umbrella of Horizon 2020 and beyond (H2020) phase II projects, has been working on providing comprehensive broadcast and multicast solutions that accommodate vertical sectors such as multimedia & entertainment, automotive, public warning systems and internet of things.

Two major design components in the 5G-Xcast systems are the radio access protocols and resource management, which have been studied in Work Package (WP) 3 task 3.4. The envisioned objective of the task is to provide a highly flexible and efficient utilization of radio resources for multicast and broadcast communication in a common platform with unicast communication. To this end, the radio protocol design should support a delivery of multicast and broadcast[2] data via both Point-To-Point (PTP) and Point-To-Multipoint (PTM) radio transmissions which envisage seamless transition or switching between PTP and PTM transmissions. Moreover, the radio access design includes intelligent logic to flexibly apply forward error correction schemes, depending on whether PTP or PTM is used for the delivery of multicast and broadcast data. Note that the flexible transition between PTM and PTP is not applicable for receive only devices since they do not support uplink that is required to exploit the PTP benefits, such as error correction from link adaptation and Hybrid Automatic Repeat Request (HARQ).

Conventional broadcast and multicast technologies may suffer from heavy packet loss during poor radio channel conditions, due to the fact that feedback systems for link adaptation and packet re-transmissions such as HARQ and Automatic Repeat Request (ARQ) are not applied. The use of feedbacks for link adaptation and re-transmissions makes broadcast and multicast systems complex, since the radio network is expected to react with the same network setting for all users that in practice have various radio channel conditions. In this deliverable, the prospect of feedback systems in broadcast and multicast transmission is investigated in the context of 5G-Xcast radio protocol design. In particular, the use of 2nd layer of Forward Error Correction (FEC), which is also referred to as layer 2 FEC[3], scheme with consideration of the feedback for

[2] The definitions of the terms multicast and broadcast are subjects of big debate in wireless communication community. In this document, the proposed RAT protocol and RRM solutions are comprehensive for broadcast (transmission to all UEs) and multicast (transmission to set of UEs that are known by the network that those UEs will be receiving the multicast data). Solutions that use uplink channel are applicable for UEs that are not "receive only". On the other hand, solutions that don't require the uplink channel are applicable for "receive only" devices.

[3] The proposed scheme is considered as 2nd layer of FEC since there is a 1st layer of FEC in 5G NR which is already standardized in 3GPP at the PHY layer. The proposed scheme is also termed as layer 2 FEC since it is proposed to be implemented at the RLC layer or above in the RAN protocol architecture which is equivalent to layer 2 of the OSI model.

retransmission of FEC packet data units is proposed to considerably reduce the packet loss rate with better feedback efficiency. For receive only devices, the 2nd layer of FEC can be applied by using redundancy to generate repair packets. Moreover, Quality of Service (QoS)-aware HARQ is proposed to optimize feedback requests for re-transmission based on the QoS requirement of the multicast and broadcast service. Furthermore, a cross-layer link adaptation in co-ordination with higher layer error correction schemes is proposed to further improve the radio efficiency of the network. For the case where the network has limited congestion and adequate resources to support multiple Modulation and Coding Scheme (MCS) settings for the same service, use of the optimized MCS settings based on grouping of UEs is proposed to improve UE's energy efficiency by allowing them to tune to the broadcast / multicast channel that serves the UE's desired MCS setting.

One of the techniques used to achieve high data rates to support data-rate-demanding applications such as multimedia, is to use radio spectrum in different domains. Traditionally, cellular technologies such as Long Term Evolution - Advanced (LTE-A) use licensed bands whereas non-cellular technologies such as WiFi use unlicensed band. In this document, the prospect of spectrum sharing in 5G-Xcast is summarized taking into account various use cases and several spectrum allocation bands.

In addition to the conceptual design of 5G-Xcast radio protocols and Radio Resource Management (RRM), coverage and system-level simulations are used to perform key performance evaluations:

The coverage simulations show a detailed analysis on proportions of areas that are not covered by conventional Digital Terrestrial Television (DTT) broadcast networks, by using a real life scenario. Herein, a distribution of cellular network based on LTE-A that can serve the areas not covered by DTT has been analysed. Furthermore, the number of users per cell is analysed at various TV transmission periods considering realistic data across various TV channels. By using cell-neighbour relation information and the number of users per cell at various TV transmission periods, the prospect of dynamic utilization of PTM or PTP is analysed.

System level simulations are used to evaluate the prospect of using 2nd layer of FEC coupled with an efficient feedback to request re-transmission of FEC packet data units. Herein, it is shown that the scheme considerably reduces packet loss rates with a negligible overhead on spectral efficiency. Besides, further system level simulations are used to evaluate link adaptation techniques with broadcast and multicast, and to analyse 5G Single Cell Point to Multi-Point (SC-PTM) schemes in comparison to LTE-A based 4th Generation (4G) SC-PTM.

5G-Xcast, along with other 5GPPP projects, has participated in the evaluation of 3GPP's NR that is set to meet IMT-2020 requirements. The calibration and final system level simulation results for IMT-2020 evaluation of NR focusing on enhanced mobile broadband use cases are included in this document.

List of Figures

List of Tables

List of Acronyms and Abbreviations

3GPP	3rd Generation Partnership Project
4G	4th Generation
5G	5th Generation
5G-PPP	5G Infrastructure Public Private Partnership
ACK	Acknowledgement
ADU	Application Data Unit
AF	Application Function
AFC	Automated Frequency Coordination
ALC	Asynchronous Layer Coding
AL-FEC	Forward Error Correction at Application Layer
AM	Acknowledged Mode
AMF	Access and Mobility Function
Auto	Automotive
BLER	BLock Error Rate
BM-SC	Broadcast Multicast Service Center
BS	Base Station
CBRS	Citizens Broadband Radio Services
CMAS	Commercial Mobile Alerts
CQI	Channel Quality Indicator (CQI)
CRC	Cyclic Redundancy Check
CU	Centralized Unit
D2D	Device To Device
DASH	Dynamic Adaptive Streaming over HTTP
DCCH	Dedicated Control Channel
DFC	Dynamic Frequency Coordination
DL-SCH	Downlink Shared Channel
DL	Down Link
DTCH	Dedicated Traffic Channel
DTT	Digital Terrestrial Television
DU	Distributed Unit
DVB-T	Digital Video Broadcasting – Terrestrial
DVB-T2	Digital Video Broadcasting – Second Generation Terrestrial
ENG	Electronics News Gathering
EU	European Union
EC	Error Correction
FLUTE	File Delivery over Unidirectional Transport
eMBB	enhanced Mobile Broadband
eMBMS	enhanced Multimedia Broadcast Multicast Service
eNB	eNodeB
ETWS	Earthquake and Tsunami Warning System
ETSI	European Telecommunications Standards Institute
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
FEC	Forward Error Correction
gNB	NR nodeB or gNodeB
GoB	Grid of Beam
GSM	Global System for Mobile
H2020	Horizon 2020 and beyond
HARQ	Hybrid Automatic Repeat Request
IoT	Internet of Things
IEG	Independent Evaluation Groups
IETF	Internet Engineering Task Force
IGMP	Internet Group Management ProtocolIMT International Mobile Telecommunication
IP	Internet Protocol
ITU	International Telecommunications Union
ITU-R	ITU Radiocommunication sector
LAA	License Assisted Access

LDPC	Low-Density Parity Check
LSA	License Shared Access
LTE	Long Term Evolution
LTE-A	Long Term Evolution - Advanced
mMTC	massive Machine-Type Communications
MAC	Medium Access Control
MBMS	Multimedia Broadcast Multicast Service
MBSFN	MBMS over Single Frequency Networks
MCE	Multi-cell/multicast Coordination Entity
MC-MM	Multicell Mixed-Mode
MCCH	Multicast Control Channel
MCS	Modulation and Coding Scheme
MIMO	Multiple-Input and Multiple-Output
MLR	Multicast Listener Report
MM-MC-MCCH	Mixed Mode-Multicell-Multicast Control Channel
MM-MC-MTCH	Mixed-Mode Multicell Multicast Traffic Channel
MNO	Mobile Network Operator
MODCOD	Modulation and Coding
MooD	MBMS operation on-Demand
MTC	Machine Type Communication
MTCH	Multicast Traffic Channel
M&E	Media & Entertainment
NACK	Negative Acknowledgment
NCGI	R Cell Global Identity
NG-RAN	Next Generation RAN
NR	New Radio
OB	Outdoor Broadcasting
PDCP	Packet Data Convergence Protocol
PDSCH	Physical Downlink Shared Channel
PDU	Protocol Data Unit
PHY	PHYsical radio layer
PHY-SI	Physical layer service integration
PMCH	Physical Multicast Channel
PMSE	Programme Making and Special Events
PoC	Proof of Concept
PTM	Point-to-Multipoint
PTP	Point-to-Point
PW	Public Warning
PWS	Public Warning System
QCI	Quality Class Indicator
QoS	Quality of Service
QFI	QoS Flow Identification
QPSK	Quadrature Phase Shift Keying
Rel	Release
RAN	Radio Access Network
RAT	Radio Access Technology
RB	Radio Bearer
RIT	Radio Interface Technology
RLC	Radio Link Control
RLNC	Random Linear Network Coding
ROM	Receive Only Mode
ROHC	Robust Header Compression
RRC	Radio Resource Control
RRM	Radio Resource Management
RRS	Reconfigurable Radio System
SC	Single Cell
SC-MCCH	Single Cell – Multicast Control Channel
SC-MTCH	Single Cell– Multicast Traffic Channel
SC-PTM	Single Cell – Point-to-Multipoint

SDAP	Service Data Adaptation Protocol
SDU	Service Data Unit
SE	Spectral Efficiency
SFN	Single Frequency Networks
SI	System Information
SIB1	System Information Block 1
SIB12	System Information Block 12
SINR	Signal to Interference plus Noise Ratio
SISO	Single Input Single Output
SMF	Session Management Function
SMRRM	Secrecy Multicast Rate Region Maximization
SRIT	Set of Radio Interface Technologies
SU-MIMO	Single User Multiple Input Multiple Output
TAI	Tracking Area Identity
TB	Transport Block
TBS	Transport Block Size
TCP	Transmission Control Protocol
TDMA	Time Division Multiplexing Access
TM	Transparent Mode
TR	Technical Report
TXRU	Transceiver Unit
TS	Technical Specification
TV	Television
TVWS	TV White Space
UE	User Equipment
UDP	User Datagram Protocol
UL	Uplink
UMTS	Universal Mobile Telecommunications System
UM	Unacknowledged Mode
URLLC	Ultra-Reliable and Low-Latency Communication
V2X	Vehicle-to-vehicle or Vehicle-to-Infrastructure
WP	Work Package

1 Introduction

1.1 Background

A rapid evolution of broadcast / multicast vertical market sectors such as multimedia and entertainment, automotive, public warning systems and internet of things, calls for the development of the broadcast / multicast communication technology to satisfy the requirements of these vertical market sectors.

Currently, 3GPP's specification work on 5G NR focuses mainly on unicast features. The specification for NR started with release 15 being purely based on unicast for 5G phase I. Even though release 16, 5G phase II, has started, the discussions on broadcast and multicast features are still at an early stage.

5G-XCast is an H2020 Phase II project focused on broadcast and multicast communication enablers for the 5G wireless systems. It has been working towards providing a comprehensive solution to support the requirements of the aforementioned vertical market sectors. Among other things, the goals of the project includes the design of a highly flexible and efficient RRM for embedding broadcast / multicast into 3GPP's 5G NR.

1.2 Objective

The main objective of WP3 task 3.4 on "RAT protocols and RRM" is to design a highly efficient and flexible 5G-Xcast Radio Access Network (RAN) protocol and RRM that fulfils the requirements of various use cases within the scope of the project, and to perform proof-of-concept performance evaluations of 5G-Xcast RRM solutions. Moreover, evaluations of 3GPP's NR that has been submitted to the International Telecommunications Union (ITU) as a candidate technology for IMT-2020 have been performed under the roof of the 5G-PPP for system configurations related to the interest of the project.

1.3 Structure of the document

The document is structured as follows. Section 2 discusses the motivation and scope of PTM protocol and RRM. Section 3 presents state of the art RRM in LTE-A PTM and NR which is so far primarily designed for unicast. Section 4 describes the designed RRM in 5G-Xcast. The performances of the RRM principles are evaluated in Section 5 by using coverage and system-level simulation. Section 6 presents guidelines for the deployment of 5G-Xcast radio access network. Section 7 presents calibration and final system level simulation results for IMT-2020 evaluation of NR focusing on enhanced mobile broadband use cases. Section 8 presents summaries and concluding remarks on the activities performed in this task.

2 Motivation and scope of PTM protocol and RRM design

2.1 Requirements on the radio access network

The study item 3GPP TR 38.913 [2] highlights the Radio Access Technology (RAT) technical requirements for future broadcast and multicast systems. Complimentarily, the radio access requirements that have impact on the 5G-Xcast use cases have been studied by WP3 working group in D3.1 [3]. From protocol and RRM point of view, the new RAT is expected to be flexible and efficient enough to support the requirements of existing services (e.g., download, streaming, group communication, TV, etc.) and new services e.g., Vehicle-to-vehicle or Vehicle-to-Everything (V2X) and services for massive Machine-Type Communications (mMTC) devices. Moreover, the new RAT is expected to support efficient multiplexing of unicast and broadcast / multicast across, at least, time and frequency domains. Furthermore, the RAT is expected to support dynamic adjustment of broadcast / multicast areas based on user distribution or service requirements.

2.2 Limitations of LTE-A Pro broadcast radio access network

The major RAN limitations in the latest LTE-A broadcast release, which is also termed as Further enhanced Mobile Boadcast and Multicast Service (FeMBMS), have been identified in D3.1 [3]. Among other things, the limitations on the RRM, latency and service scheduling have been elaborated.

In regard to RRM, the specification has limited support for feedback systems to assist the network to optimize the radio resources leading to challenges in terms of providing the required spectral efficiency and packet loss rates, which create constraints on requirements such as M&E1_R7, M&E1_R23, M&E1_R29, M&E1_R36 and Auto1_R2.

Moreover, lack of flexible switching between PTP and PTM transmission schemes as well as handover procedures between Multicast-Broadcast Single Frequency Networks (MBMSFN) areas create challenges on service continuity, which in turn could constraint requirements such as M&E1_R24, Auto1_R1 and PW1_R12.

Furthermore, there is limited flexibility on the trigger for Multimedia Broadcast and Multicast System (MBMS) service access where a trigger must come from the network side to wake up MBMS reception for saving User Equipment (UE)'s power which is relevant for such requirements as PW1_R5 in Public Warning System (PWS) applications. In many cases, it is the user who activates reception of multicast and broadcast content. But in the case of PWS, users are not aware when such a warning message is going to be broadcasted. Therefore, the trigger to start receiving warning message content needs to come from the network.

The flexibility of the solution to allow operation under different spectrum usage frameworks is important as indicated in requirement M&E1_R38. The flexibility would allow deployments in various scenarios of spectrum allocation for the networks.

2.3 Scope of the tasks on 5G-Xcast RAT protocol and RRM

To address some of RAN protocol and RRM limitations of the current broadcast / multicast systems, WP3 tasks on "RAT protocol and RRM" are allocated among participating partners based on their area of expertise and the scope of the project proposal.

5G-Xcast RAT protocol and RRM study includes investigation of feedback systems for broadcast and multicast system via link adaptation as well as efficient HARQ with consideration of the trade-off among spectral efficiency, packet loss rates and signalling overhead for the feedbacks. Moreover, the use of second layer of forward error correction scheme has been investigated in order to provide improved spectral efficiency and packet loss rates (M&E1_R7, M&E1_R23, M&E1_R29, Auto1_R2). A feedback system with lower signalling overhead can be tailored with second layer of FEC for further improvements in spectral efficiency and reduced packet loss rates.

5G-Xcast RAT protocol design also targets provision of flexible and efficient radio resource allocation methods considering QoS requirements for all services. The protocol functions have taken into account seamless transition between PTP and PTM transmission modes to guarantee service continuity requirements (M&E1_R24, Auto1_R1 and PW1_R12). Moreover, flexible and intelligent algorithm has been designed to provide optimized content delivery by exploiting adaptation of PTM transmission schemes with possibility of a dynamical defined RAN-level multicast area. Furthermore, various aspects of the RRM have been investigated by using practical and heuristic approaches.

One aspect of efficient RRM is the use of triggers from the network to initiate MBMS reception in order to provide PWS applications. Herein, a trigger from the network eliminates the need for the UE to continuously monitor the MBMS channels which in turn is expected to lower UE power consumption (PW1_R5). An example of such a trigger mechanism in E-UTRAN is the '*cmas-indication*' in the paging message, which triggers reception of cell broadcast messages. The '*cmas-indication*' indicates to the UE that System Information Block 1 (SIB1) now contains the scheduling information for System Information Block 12 (SIB12), which contains a cell broadcast warning message.

Furthermore, concrete evaluations are performed via elaborated coverage and system level simulations. The performance of the 5G-Xcast RAT protocol solution has been compared with state of the art PTP and PTM solutions. Besides, evaluations of 3GPP's "New Radio", which has been submitted to ITU as a candidate technology for IMT-2020, have been performed under the roof of the 5G-PPP for system configurations related to the interest of the project.

Spectrum sharing in 5G-Xcast is investigated to address requirements M&E1_38. For the basis of the study, selected scenarios are evaluated to study their suitability and further developed to Proof of Concepts (PoCs) to prove their suitability with real equipment.

3 State of the art RRM

3.1 RRM protocols in LTE-A PTM

The most relevant RRM-related protocol layers in LTE-A PTM [7], particularly for user plane data, are Radio Link Control (RLC), Medium Access Control (MAC) and Physical (PHY) layer, as depicted in Figure 3.1-1. Herein, the Packet Data Convergence Protocol (PDCP) is not used, i.e., transmission for multicast / broadcast operates using PDCP transparent mode. The major roles of the RLC layer are segmentation and/or concatenation of RLC Service Data Units (RLC SDUs) to fit into the available MAC transport blocks provided by the lower layers. On the other hand, the major functions of the MAC protocol in LTE-A PTM are radio resource scheduling and multiplexing of data to lower layer transport blocks.

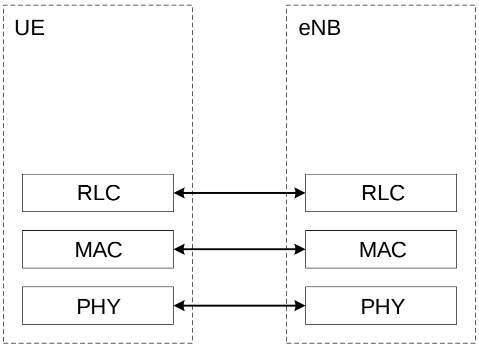

Figure 3.1-1: LTE-A RAN protocols for PTM data transmission [7].

Figure 3.1-2 shows an example of user plane data flow in LTE-A PTM. First of all, the MBMS packets from higher layers are input to RLC layer as RLC SDUs. Based on the available MAC transport block, the RLC layer concatenates or segments RLC SDUs. Next, the RLC layer appends header information to the RLC SDUs to generate RLC Protocol Data Units (PDUs). The RLC header contains information that supports the corresponding receiver RLC to assemble RLC SDUs from received RLC PDUs. After RLC PDUs are generated, the MAC layer multiplexes RLC PDUs which may come from different sources, e.g. different MBMS services, into the available MAC transport block.

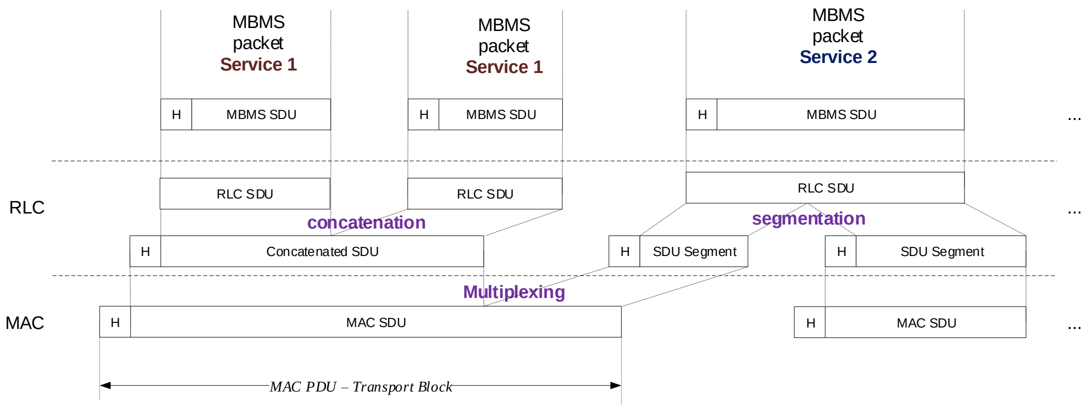

Figure 3.1-2: Packet flow for MBMS data [7].

3.2 RRM protocols in 5G NR

This section briefly outlines the radio protocols specified for the 5G NR with consideration of PTP communication. Figure 3.2-1 describes the architecture of the radio protocol function pertinent to communication between an NR gNodeB (gNB) and a UE. The specified radio protocols are Service Data Adaptation Protocol (SDAP), PDCP, RLC, MAC and PHY layers [8]. A major change is that the concatenation of packets no longer takes place at the RLC layer, but has been moved to the MAC layer. A completely new element is the SDAP layer which is used for packet marking with QoS flow ID (QFI) and mapping of QFI to radio bearers.

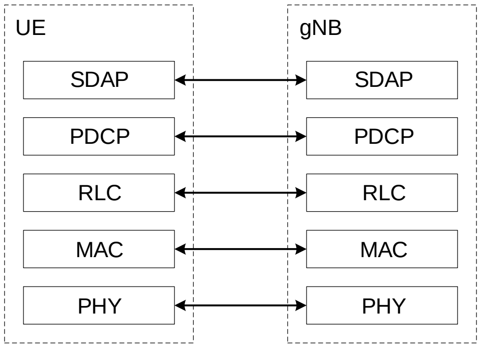

Figure 3.2-1: NR radio protocols for unicast data transmission [8].

The main functions of SDAP, PDCP, RLC and MAC in accordance with 3GPP's general description of NR [5] are listed below. Further details on the specification for layer 2 protocols (SDAP, PDCP, RLC and MAC) can be found in [9], [10], [11] and [12], respectively.

SDAP Layer

- Mapping between a QoS flow and a data radio bearer
- Marking QoS flow ID (QFI) in both Downlink (DL) and Uplink (UL) packets

PDCP Layer

- Header compression and decompression: Robust Header Compression (ROHC) only
- Reordering and duplicate detection
- PDCP PDU routing (in case of split bearers)
- Retransmission of PDCP SDUs
- Ciphering, deciphering and integrity protection
- PDCP SDU discard
- PDCP re-establishment and data recovery for RLC AM
- Duplication of PDCP PDUs

RLC Layer

- Transparent Mode (TM) or Unacknowledged Mode (UM) or Acknowledged Mode (AM)
- Segmentation (AM and UM) and re-segmentation (AM only) of RLC SDUs
- Reassembly of SDU (AM and UM)
- RLC SDU discard (AM and UM)
- Error Correction through ARQ (AM only)
- Duplicate Detection (AM only)
- Protocol error detection (AM only)

MAC Layer

- Mapping between logical channels and transport channels

- Multiplexing / demultiplexing of MAC SDUs belonging to one or different logical channels into / from transport blocks (TB) delivered to / from the physical layer on transport channels
- Scheduling information reporting
- Error correction through HARQ
- Priority handling between logical channels of one UE
- Priority handling between UEs
- Packet re-ordering with retransmissions with HARQ

Figure 3.2-2 elaborates on an example of downlink user plane data flow across the 5G NR radio protocols. First of all, higher layer Internet Protocol (IP) packets are marked with QFI and mapped to radio bearers. Then, the PDCP layer performs header compression and security (ciphering and integrity protection) and forwards PDCP PDUs to the RLC layer. After this, the RLC layer wraps RLC SDUs or segments thereof into RLC PDUs based on the available MAC layer transport block size. Unlike current PTM systems which support only UM mode communication, the 5G-NR PTP can operate in UM or AM mode where re-transmissions of lost packets can be performed via Automatic Repeat request (ARQ) procedures. Following the RLC functions, the MAC layer multiplexes RLC PDUs which may come from the same or different sources, e.g. different radio bearers, into the available MAC transport block.

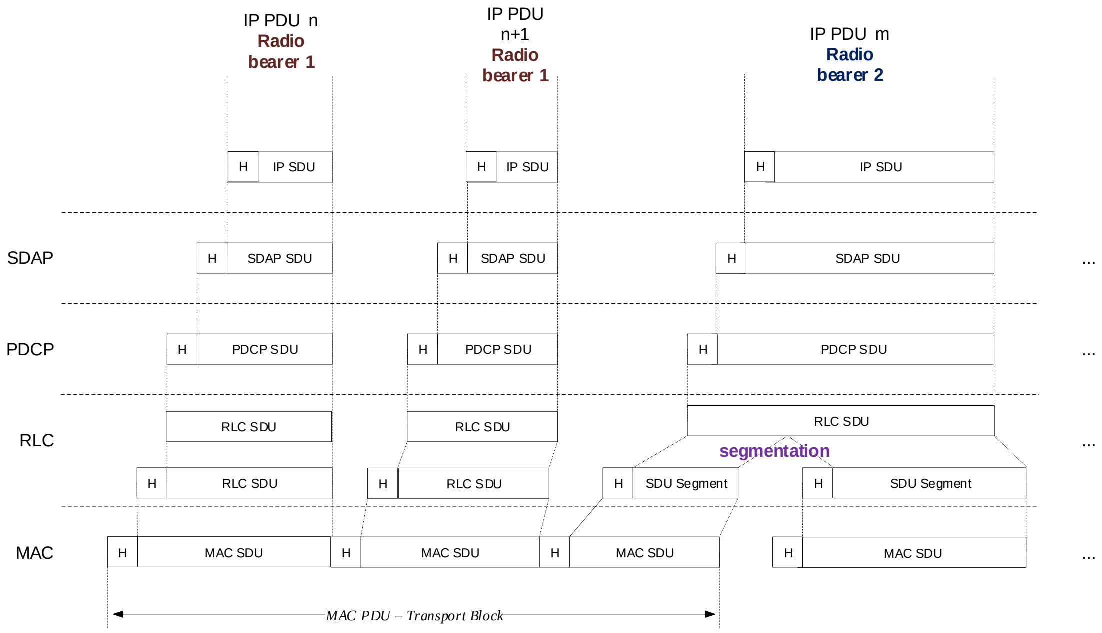

Figure 3.2-2: User plane data flow across 5G-NR radio protocols [8].

4 RRM for 5G-Xcast

4.1 Contributions

The 5G-Xcast RAT protocol and RRM functionalities are designed to support flexibility and efficiency of new radio that is required for existing and future services, and it uses 3GPP's NR as baseline for enhancement. To this end, the major contributions of 5G-Xcast RAT protocol and RRM includes

- Support for flexible delivery of multicast or broadcast data via a radio-access-level seamless transition between PTP and PTM transmission modes. In a certain geographical area, if there are a limited number of UEs consuming a service in broadcast or multicast mode, better spectral efficiency can be achieved by mapping PTM radio bearers to PTP radio bearers. Moreover, if a UE that uses PTM transmission is experiencing poor radio channel conditions, transition of UE's transmission mode to PTP transmission by mapping PTM radio bearers to PTP radio bearer may improve spectral efficiency by exploiting PTP benefits such as link adaptation and HARQ (taking the latency constraint of the service into account). Further details can be found in Section 4.2.2 and D3.3 [45] Section 6.1.3.
- Support for use of QoS-aware feedback to optimize HARQ feedback overheads in PTM bearers. In case of very high number of UEs, ACK / NACK feedbacks can be source of extremely high signalling overhead that considerably deteriorate the network efficiency in general. To alleviate the signalling overhead to some extent, the HARQ feedbacks are optimized based on QoS requirement of the service. Further detail can be found in Section 4.3.1. Moreover, a method that leverages HARQ via unicast cannel as a retransmission feature for broadcast / multicast is investigated in Section 4.4.3.
- Support for selective FEC upon transition from PTM to PTP transmission modes. In the case of using AL-FEC, a selective FEC procedure is used to make the radio access network intelligently select only source packets for the PTP radio bearer and both source and repair packets for the PTM radio bearer. Further details can be found in Section 4.2.3 and D3.3 [45] Section 5.3
- Support for feedbacks for efficient link adaptation and error corrections in PTM transmission modes. To reduce the heavy packet losses and maintain technical requirements reliably, in poor channel conditions of PTM transmission, the prospect of feedback and error correction schemes have been investigated. To this end, link adaptation in multiple antenna configuration for PTM and a layer 2 Error Correction (EC), also known as 2nd layer of FEC, in the radio access network are investigated. Further details can be found in Section 4.3.2 and Section 4.3.3.
- Support for efficient multiplexing of unicast and broadcast / multicast. Further details can be found in D3.3 [45] Section 7.2 and Section 4.2.1.
- Support for mechanism of link adaptation in co-ordination with higher layer error correction schemes such as layer 2 EC in order to achieve efficient and reliable broadcast / multicast wireless link that fulfils minimum expected Quality of Experience (QoE) requirements of a service at minimum cost in the spectral efficiency of the network. Further details can be found in Section 4.3.4.
- Support for RAN-level security. RAN-level security procedures are one of the requirements in 5G-Xcast networks. As described in D3.3 [45] Section 2, the same security procedures applied to PTP transmissions should be applied in PTM ones. For the Terrestrial Broadcast mode, if confidentiality is enforced, the relevant encryption keys should be made available to the users via other channels e.g. Unicast for non-ROM devices. Optionally, for terrestrial broadcast

with no unicast support for non-ROM, a theoretical approach that optimizes physical layer RRM with consideration of security is described in Section 4.7.
- Support for use of multiple MCS configurations targeting improvement of energy efficiency of UEs by allowing them to tune to multicast / broadcast channel that is suitable to the channel conditions of the UEs. Note that such a scheme that uses multiple MCS configuration is applicable only when there is no congestion and reasonably adequate resources are available to serve multiple broadcast / multicast channels. Further detail can be found in Section 4.4.4.
- Support for flexible 5G-Xcast spectrum sharing that proposes a reasonable allocation of various spectrum bands taking into account various use cases considered in the 5G-Xcast project. Further details can be found in Section 4.6.

Contribution on coverage and system-level simulations for PTM includes:

- The coverage simulations show a detailed analysis on proportions of areas that are not covered by conventional DTT broadcast networks by using a real life scenario. Moreover, a distribution of 3GPP-based network that can serve the areas with potential users that are not covered by DTT has been analysed. Furthermore, the number of users per cell are analysed at various TV transmission periods considering realistic data of various TV channels. By using cell-neighbour relations and the number of users per cell at various TV transmission periods, prospect of dynamic utilization of PTM or PTP is analysed. Further details can be found in Section 4.4.2, 5.1.1, and 5.2.1.
- System level simulations are used to evaluate the prospect of using 2nd layer of FEC coupled with efficient feedback to request re-transmission of FEC packet data units. Besides, system level simulations are used to evaluate link adaptation techniques with broadcast and multicast, and to analyse 5G Single Cell Point to Multi-Point (SC-PTM) schemes in comparison to LTE-A based 4th Generation (4G) SC-PTM. Further details can be found in Section 5.1.2 and 5.2.2.

5G-Xcast, along with other 5GPP projects, has contributed in the evaluation of 3GPP's NR that is set to meet IMT-2020 requirements. The steps in the IMT2020 evaluation in 5G-Xcast include:

- Calibration of the system-level simulator used in this project against that provided by 3GPP members. Further details can be found in Section 7 and Annex A.
- Evaluation of NR focusing on enhanced mobile broadband. Further details can be found in Section 7 and Annex B.

4.2 Flexible resource allocation methods for 5G-Xcast

4.2.1 Multiplexing among unicast and PTM transmission schemes

This section is covered in D3.3 Section 7.2 [45].

4.2.2 RAN-level seamless transition between xcast modes

To optimize utilization of radio resources in scenarios that have considerable diversity of traffic volume, the radio system should be able to flexibly transition or switch between PTP and PTM transmission modes. To this end, it is proposed that 5G-Xcast RAT supports RAN-level flexible switching between PTP and PTM transmission modes.

In 3GPP, a flexible delivery of content as a unicast or broadcast service is specified as MBMS Operation on Demand (MooD) feature in 3GPP TS26.346 [43] which is also referred to as “MBMS offloading”. Herein, the decision for transition between xcasting delivery modes (PTP via unicast or PTM via broadcast or multicast) is made at the core-level targeting optimisation and balancing of traffic volumes in the core network. The unicast to/from broadcast switching anchor is the Broadcast Multicast Service Center (BM-SC) which utilizes user service consumption reports from UEs to make switching decisions. Further study of the MooD feature in 5G-Xcast can be found in D4.1 [44]. Supplementary to this, RAN-level seamless transition between PTP and PTM transmission modes is proposed in 5G-Xcast to optimise utilization of radio resources.

For delivery of IP multicast / broadcast content, RAN-level transmission modes include PTP or PTM radio transmission in association with various RRC states as described in D3.3 [45]. The PTP transmission mode utilizes UE-specific dedicated Radio Bearer (RB) for control and data signals whereas the PTM transmission mode is not dedicated to a specific UE. Accordingly, delivery of IP multicast / broadcast data to UEs in RRC connected state can be realized by mapping multicast Radio Bearer (RB) to unicast RB which uses PTP transmission whereas delivery of IP multicast / broadcast data to UEs in RRC Inactive or RRC Idle is done via the multicast RB which uses PTM transmission mode. Herein, UE’s RRC inactive state has considerable benefit over UE’s RRC idle state since it maintains UE’s connection of the RAN to the core network and transition of RRC inactive state to RRC connected state can be performed with extremely low latency. However, care should be taken not to waste the dedicated link established with the core if the UE is in RRC_Inactive state for extremely longer time.

The criteria for switching between unicast RB and multicast RB can be the number of UEs demanding multicast service and/or UEs’ QoS requirement. For example, the criteria on the number of UEs can be implemented by re-using LTE-A’s counting procedure 3GPP TS 36.300 [7]. However, unlike LTE-A which uses the counting function to disable (suspend) or enable (resume) multicast RB transmission, in 5G-Xcast the counting function can be used to make the decision of RB switching between unicast and multicast RBs.

Figure 4.2-1 describes high-level RAN procedure to switch from unicast RB to multicast RB to deliver IP multicast data. First, information on the number of multicast UEs is collected by using counting functions. Then, decision for switching bearer is made based on a threshold configured by the network operator or network planner. Following bearer switching decision, the new RB re-configuration is sent to the UEs that consume multicast service. Moreover, the buffered data in the unicast RB are copied to the multicast RB. Accordingly, the buffered data and newly arriving IP multicast data are transmitted over the multicast RB. Note that with unicast RB transmission the RLC SDU buffer can be different for various UEs since UEs have independent dedicated radio link. Hence, copying the buffer for the UE with highest buffer size is more crucial to avoid packet loss.

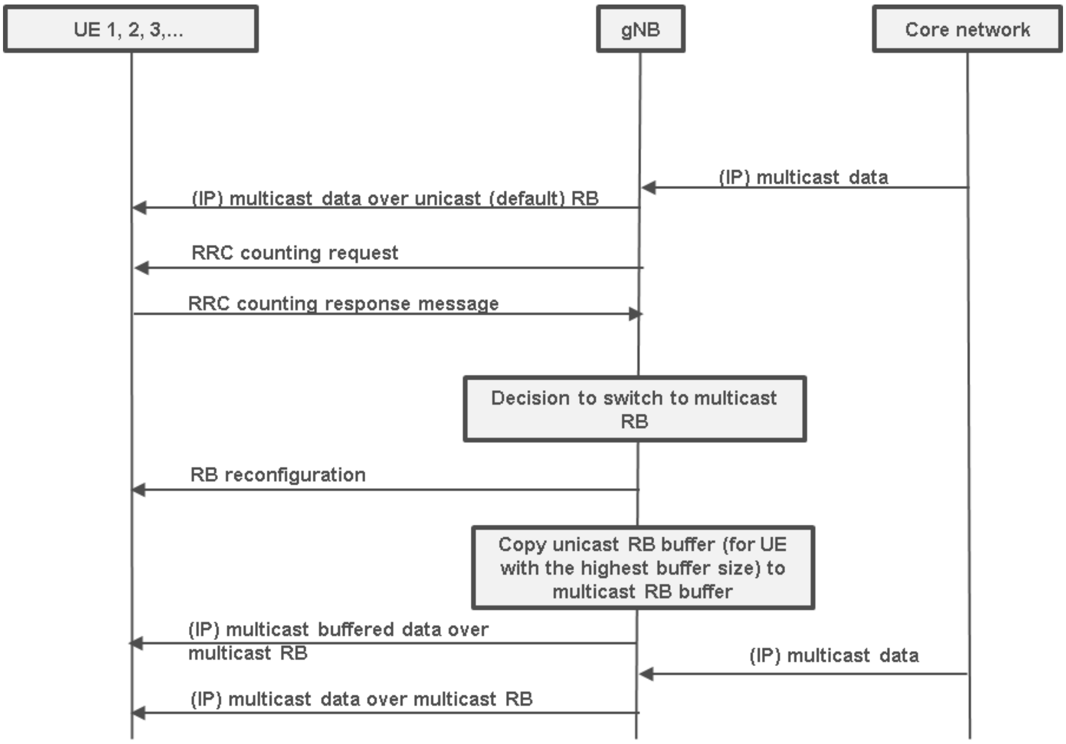

Figure 4.2-1: High-level RAN procedures to switch from unicast RB to multicast RB.

Figure 4.2-2 demonstrates high level RAN procedure to switch from multicast RB to unicast RB. Similar to previous case, counting procedure is used to collect counting results. If the number of multicast UEs are lower than a threshold set by network operator or network planner, the decision to switch from multicast RB to unicast RB can be made in order to exploit improved spectral efficiency from unicast RBs that use feedbacks or HARQ. However, the latency requirement of the considered services should be taken into account before introducing features such as HARQ while using the unicast RB. After the switching decision, RB reconfiguration can be sent to the UE. Besides, multicast RB buffer is copied to unicast RB buffer for each UEs that is being served by the gNB. Accordingly, buffered data and newly arriving multicast data are transmitted via unicast RB.

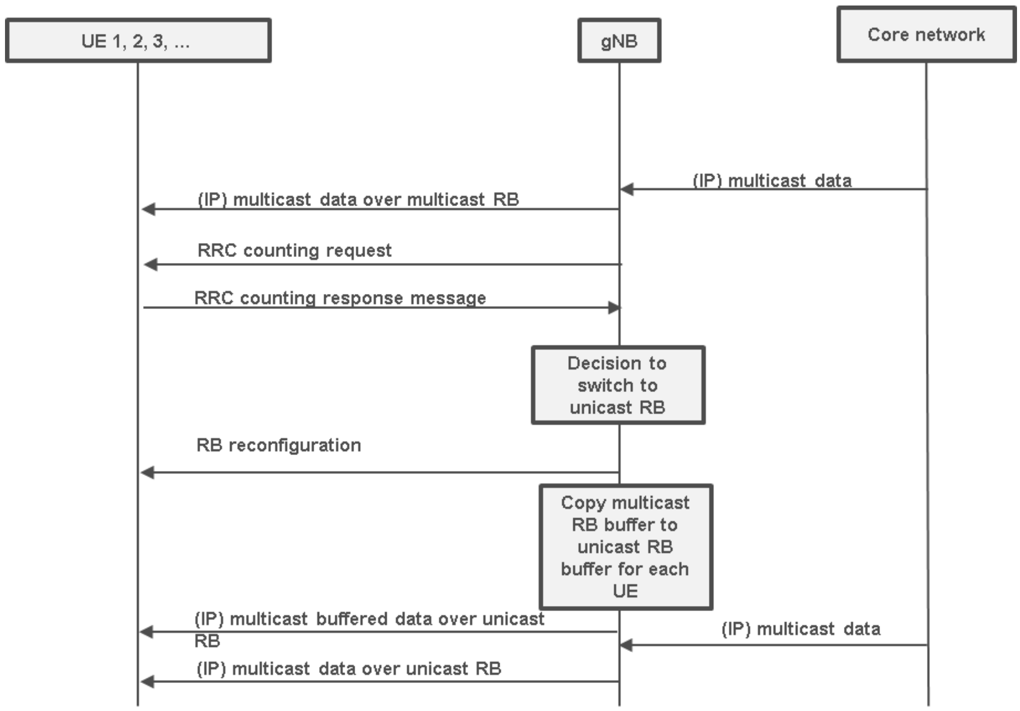

Figure 4.2-2: High-level RAN procedures to switch from multicast RB to unicast RB.

In addition to the above procedures that perform RAN-level seamless switching between unicast or multicast RB across all UEs being served by the gNB, UE-specific switching between unicast and multicast RB can be done depending on the channel condition of the user. For example, if the UE is experiencing severe degradation of received signal quality while being served by a multicast RB, it can request the gNB to switch transmission from multicast RB to unicast RB. To assist the gNB for UE-specific switching decisions, the UE has to provide signal measurement to the gNB. In particular, such switching functionalities are crucial for cloud RAN deployment with central unit and distributed unit. Herein, implementing the switching in distributed unit considerably reduces the signalling over F1-interfaces that would have been needed for UE-specific switching decisions at the central unit. Further details on implementation of such a switching function can be found in the deliverable D3.3 [45].

Figure 4.2-3 demonstrates the radio protocol enhancement to support seamless switching between unicast RB and multicast RB. The unicast control and data bearers use logical channels Dedicated Control Channel (DCCH) and Dedicated Traffic Channel (DTCH), respectively. On the other hand, the multicast control bearer uses Single Cell Multicast Control Channels (SC-MCCH) or Mixed-Mode Multi-Cell Multicast Control Channels (MM-MC-MCCH) for SC-PTM and Multi-Cell Mixed Mode (MC-MM) transmissions, respectively. Moreover, the multicast data bearer uses Single Cell Multicast Traffic Channels (SC-MTCH) or Mixed-Mode Multi-Cell Multicast Traffic Channel (MM-MC-MTCH) for SC-PTM and MC-MM transmissions, respectively. The logical channels for mixed mode multi-cell transmission, MM-MC-MCCH and MM-MC-MTCH, use physical layer numerology that is designed based on 5G unicast numerology enhancement as descried in D3.2 [46]. Seamless transition of unicast and multicast radio bearers is facilitated by mapping the relevant logical channels, DCCH, DTCH, SC-MCCH, SC-MTCH, MM-MC-MCCH, MM-MC-MTCH, onto shared transport channel, i.e., Downlink Shared Channel (DL-SCH).

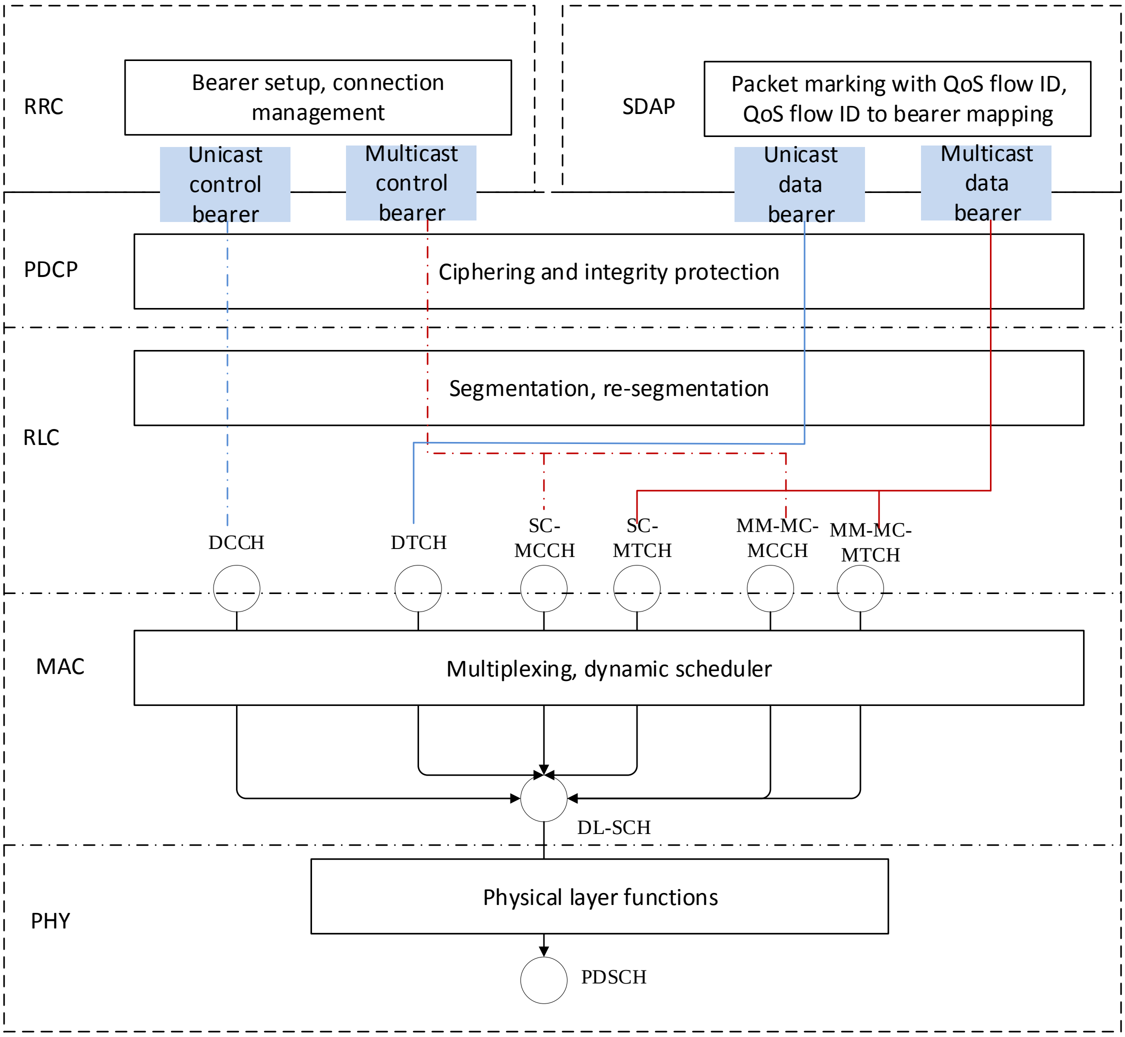

Figure 4.2-3: Radio protocol enhancement to support seamless switching between unicast RB and multicast RB.

For the case where a multicast RB is used, the decision to use SC-PTM or local MC-MM transmission can rely on geographical distribution of the multicast UEs. The geographical data may be derived from location services such as GPS. Location services may have privacy constraints where users don't consent to provide location. In such cases, UE measurement should be used to measure interference levels in order to assist switching decisions. If a considerable number of UEs are receiving multicast data via SC-PTM at the border of two or more cells, coordinating the cell for multi-cell transmission avoids interference between cells, which in turn improves spectral efficiency. For environments where location data is not available (due to privacy/

current deployment of UE level GPS information to RAN) or where it is not accurate enough, e.g. cities that have strong variation of user interference levels due to shadowing, UE measurement data may be used to assist in determining switching between SC-PTM and local MC-MM.

4.2.3 Selective FEC

Background

FEC coding is used at various layers of protocol stack. Unidirectional communication often uses FEC at higher layers to minimize impacts of packet loss between the communication end-points. One example of system using FEC at higher layers is enhanced Multimedia Broadcast and Multicast System (eMBMS). eMBMS utilizes File Delivery over Unidirectional Transport (FLUTE) for various MBMS user services such as 3GP-DASH, file download 3GPP TS 26.346 [5]. This section describes the higher layer FEC protocol supporting multicast transport. The use of unicast or multicast / broadcast transmission is autonomously determined to efficiently deliver IP multicast data. In 5G-Xcast deliverable D3.3 [45] section 5.3 - Selective FEC in 5G-Xcast RAN Architecture, the architectural aspects are explained. In the deliverable D3.3 [45] section 6.1 - L2 architecture and bearer selection in Cloud-RAN it is proposed that unicast and multicast / broadcast transmission can use different RLC entities driven by a switching function that can select between unicast and multicast transport channels.

The protocol architecture is as shown in Figure 4.2-4, considering both distributed and centralized deployments with a mix of multicast and unicast flows. In this example the DU-1 of gNB-1 will schedule the source and FEC flows using Xcast radio bearers with multicast transmissions over-the-air. In another example the DU-2 of gNB-2 will be able to drop the FEC flow from transmission and schedule only the source flow using unicast transmissions over-the-air. These flows can be multiplexed with other existing unicast flows within the gNB / DU as well.

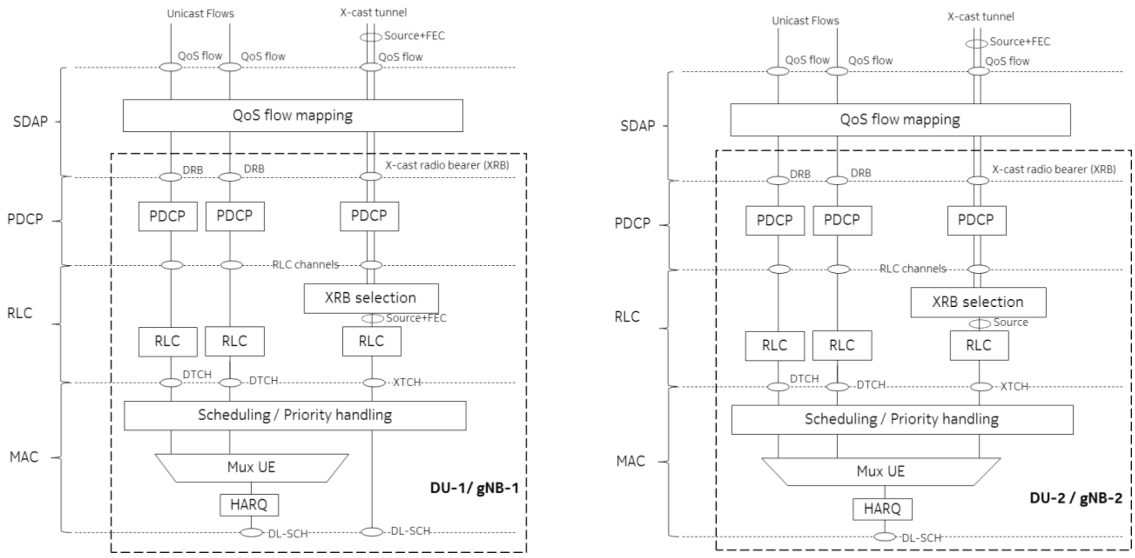

Figure 4.2-4: Selective FEC protocol architecture.

Selective FEC protocol description

This section describes the selective FEC for an IP PDU session type where it can be assumed that the PDU session modification was completed for creating a multicast context in the network. An N3 tunnel is established for the transport of multicast data (e.g. for IP multicast group). The RAN can store associations between the multicast context and all PDU sessions for which Internet Group Management Protocol (IGMP) /

Multicast Listener Report (MLR) triggered the PDU session modification procedure to create the association with the multicast context.

The Application Function (AF) invokes Npcf_PolicyAuthorization_Create request to create an application session context at the PCF in 3GPP 29.514 [49]. As part of the request, the AF specifies MediaComponet and MediaSubComponents. The MediaSubComponent can be modified to include an optional attribute indicating whether the MediaSubComponent is for a source flow or an FEC flow. The MediaSubComponent includes the fNum attribute, which is an ordinary number of IP flow. The fNum attribute can be used to refer from a MediaSubComponent carrying the FEC flow to the MediaSubComponent carrying the source flow. The MediaSumComponent would include an attribute (e.g. fNumSource) and the value of this attribute would be fNum of MediaSubComponent carrying the source flow.

If the AF invokes Npcf_PolicyAuthorization_Crate request at the time when the SMF allocated resources for multicast session and if the PCF decides that a modification is needed, then the PCF invokes Npcf_SMPolicyControl_UpdateNotify request in 3GPP TS 29.512 [50]. The PCF provides SMF with a PCC rule (PccRule) for one or more source flows and a PCC rule for one or more FEC flows for the source flows. For example, the PCC rule for FEC flows includes a reference to the PCC rule for the source flows.

The Session Management Function (SMF) decides on QoS flow mapping. The PCC rules for source flows may be mapped to one QoS flow, i.e. all source flows are aggregated, if the flows have same QoS characteristics. Similarly, the PCC rules for FEC flows can be mapped to one QoS flow. The SMF also has the option of mapping one PCC rule to one QoS flow. It should be noted that the maximum number of flows per PDU session is currently 64.

SMF message options are

- N1N2MessageTransfer
- NonUeN2MessageTransfer
- McastContextMessageTransfer

The N1N2MessageTransfer option requires that the SMF initiate the procedure for all impacted PDU sessions. The SMF initiates Namf_N1N2MessageTransfer as per the PDU Session Modification procedure in 3GPP TS 23.502 [51]. The Access and Mobility Function (AMF) initiates the PDU Session Resource Modify procedure over NG interface by sending a PDU SESSION RESOURCE MODIFY REQUEST containing the PDU Session Resource Setup Request Transfer IE in which the list of QoS flows for this PDU session is provided. The list of QoS flows for multicast may be included in the message. Each entry in the list representing the QoS flows for FEC flows can include a reference (QoS Flow Id) to the corresponding QoS flow for source flows.

The NonUeN2MessageTransfer uses Tracking Area Identities (TAIs), NR Cell Global Identities (NCGIs) and global RAN node IDs for routing N2 messages to the RAN nodes in 3GPP TS 29.518 [52] . The SMF would need to know the TAIs, NCGIs or global RAN node IDs of serving UEs that should receive the multicast. The N2 message should include a multicast context ID and QoS flows configuration as discussed above for the N1N2MessageTransfer case.

One option is to introduce a new resource in the AMF API definition for multicast contexts (e.g. ../multicast-contexts). The SMF initiates a multicast context modification procedure (e.g. McastContextMessageTransfer request by HTTP POST to ../multicast-contexts/{multicastContextId}). The multicast context ID could be for example IP multicast group defined by IP multicast address for any source multicast or by IP multicast address and source multicast addresses for source specific multicast. The multicast context ID could be a temporary ID allocated for multicast group. The AMF forwards the N2 message, which would include the multicast context ID and the list of QoS flows, to RAN nodes serving the multicast context.

Upon the reception of the information about QoS flows carrying application source flows and corresponding FEC flows, the (R)AN can decide to not transmit the QoS flow carrying application FEC flows when the (R)AN decides to transmit data to the UE using unicast bearers that use HARQ or RLC retransmissions.

In the step 14 in Figure 4.2-5, the IP multicast data is delivered to the Next Generation RAN (NG-RAN) over a data tunnel. The decision to not transmit the QoS flow carrying application FEC flows is made above the RLC, e.g. in the protocol entity where the dynamic selection is made between unicast and multicast RLC entities and transport channels for the transmission according to Figure 4.2-5. Consequently, in case of the CU/DU split architecture with F1 fronthaul interface, it is possible to place the functionality in the DU and carry the decision to not to transmit FEC flow for unicast. At least the QoS flow payload type, the one or more sets of QoS parameters and the relation between QoS flows would need to be provided over F1 interface to the DU.

A network node (e.g. DU) can decide to deliver the multicast data using only unicast transmission to all UEs e.g. due to small number of UEs and/or geographically separated UEs. In this case the network node may notify an upstream network node (e.g. CU) about the decision and the notification may include identities of QoS flows carrying application source flows. Upon receiving such notification, the upstream network node may decide not to transmit QoS flows carrying FEC flows. When the network node later decides to transmit multicast data using a multicast / broadcast transmission and the upstream network node suspended transmission of FEC flows, the network node must requests the upstream node to transmit QoS flows carrying FEC flows.

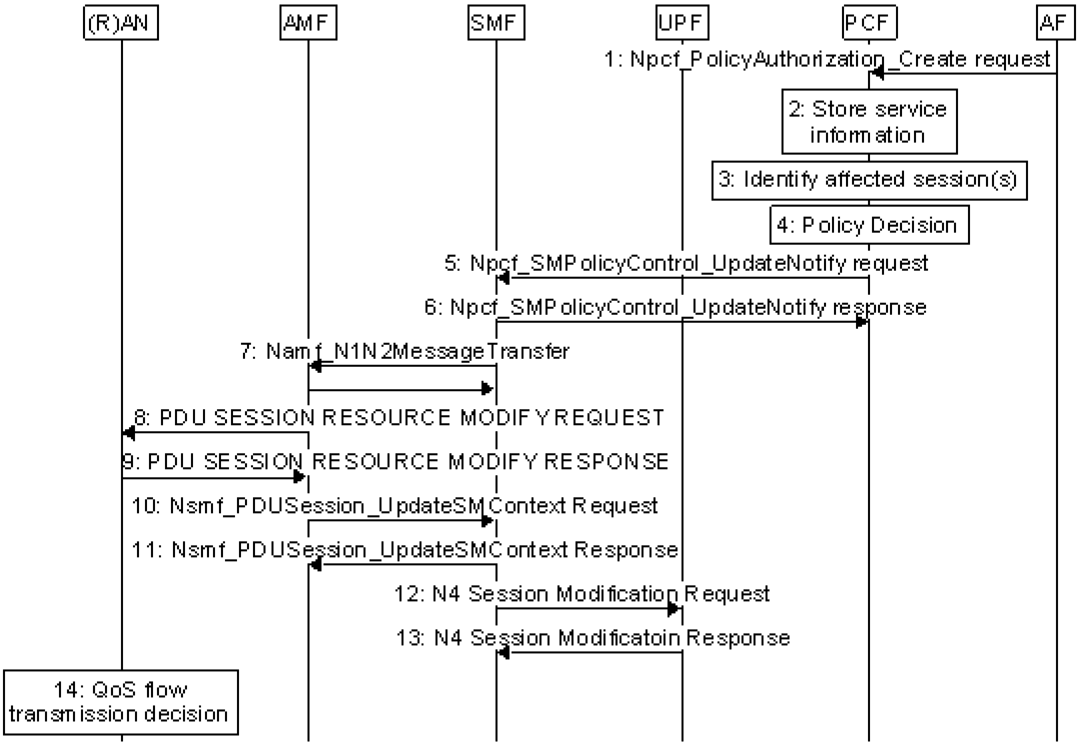

Figure 4.2-5: Signalling flow provisioning the RAN for information about the source and FEC flows.

4.3 Prospect of feedback schemes and FEC for PTM

4.3.1 Feedback schemes with QoS

In 5G, the need for feedback from the IP multicast transmission is identified as one of the key enablers for improved system performance. The system efficiency depends on the amount of data transmitted via the system in a period of time and for unicast the HARQ can significantly increase the system efficiency when an appropriate coding scheme is selected. The robustness of the selected coding scheme and the retransmission rate define the capacity of the communication channel.

In LTE-A, the UE is not aware of QoS information associated with an MBMS bearer. The QoS of the MBMS bearer is terminated at the Multi-cell/multicast Coordination Entity (MCE) which uses the received Quality Class Indicator (QCI), i.e. QoS characteristics, together with allocation and retention priority, maximum bit rate, and guaranteed bit rate in the admission process. In order to improve feedback in terms of received QoS, the availability of QoS parameters at the UE receiver is required. The QoS parameters can be provisioned to the UE either using RRC signalling or for example extending the User Service Description information. The RRC signalling is performed by the gNB based on QoS flow information received from the core network.

In a conventional unicast HARQ scheme, the receiver sends back an Acknowledgment (ACK) or Negative Acknowledgement (NACK) depending on successful packet reception. When the conventional HARQ method is applied to multicast, multiple ACK channels need to be assigned in the uplink for multicast users resulting in uplink feedback overhead which increases with the increasing number of multicast users. If the transmitter selects the same channel coding schemes as in the case of unicast transmission and the receivers are in similar radio conditions, i.e. the probability of erroneous reception is assumed to be the same, then the probability of the transmitter receiving NACK increases in the proportion of the number of receivers. For example, if the receiver correctly receives the transmission with probability of 70% (Pc = 0.7), then the probability of correct reception by two receivers is only 49% assuming the receiving processes are independent processes. In case of 3 receivers, the probability of correct reception by all receivers is 34.3%. Therefore, the number of retransmissions increases which impacts the system efficiency in proportion to number of receivers. Using more robust channel coding reduces the probability of NACKs but also decreases the spectral efficiency.

The QoS parameters (e.g. packet loss rate or delay budget) and the observed QoS at the receiver can be used to decide whether ACK, NACK or no feedback shall be sent for an erroneously received transport block (IP multicast packet). Literature in this field provides examples on how to optimise HARQ for multicast, for example by automatically retransmitting a packet a predetermined number of times without ACK / NACK feedback, and then performing the HARQ operation in the conventional HARQ method. The optimal number of autonomous retransmissions can be based on limited feedback from the UEs [48]. Furthermore, assuming that the users receiving IP multicast traffic are mainly interested in the QoS, the receiver NACK can be sent only if the received data is not going to meet the target QoS. If the receiver failed to decode the block but re-transmission is not necessary to meet the QoS, then the receiver may send ACK if the HARQ process at the transmitter requires explicit acknowledgment to proceed with transmission next data. It is also possible to send no feedback if the HARQ process at the transmitter operates with implicit ACKs where ACK is assumed if NACK is not received. In this case the block decoding can fail but the re-transmission is not necessary since the QoS requirement is fulfilled. In 5G, the SDAP layer maps the QoS flows to RBs and it is possible to continuously monitor QoS and adjust the MAC configuration accordingly.

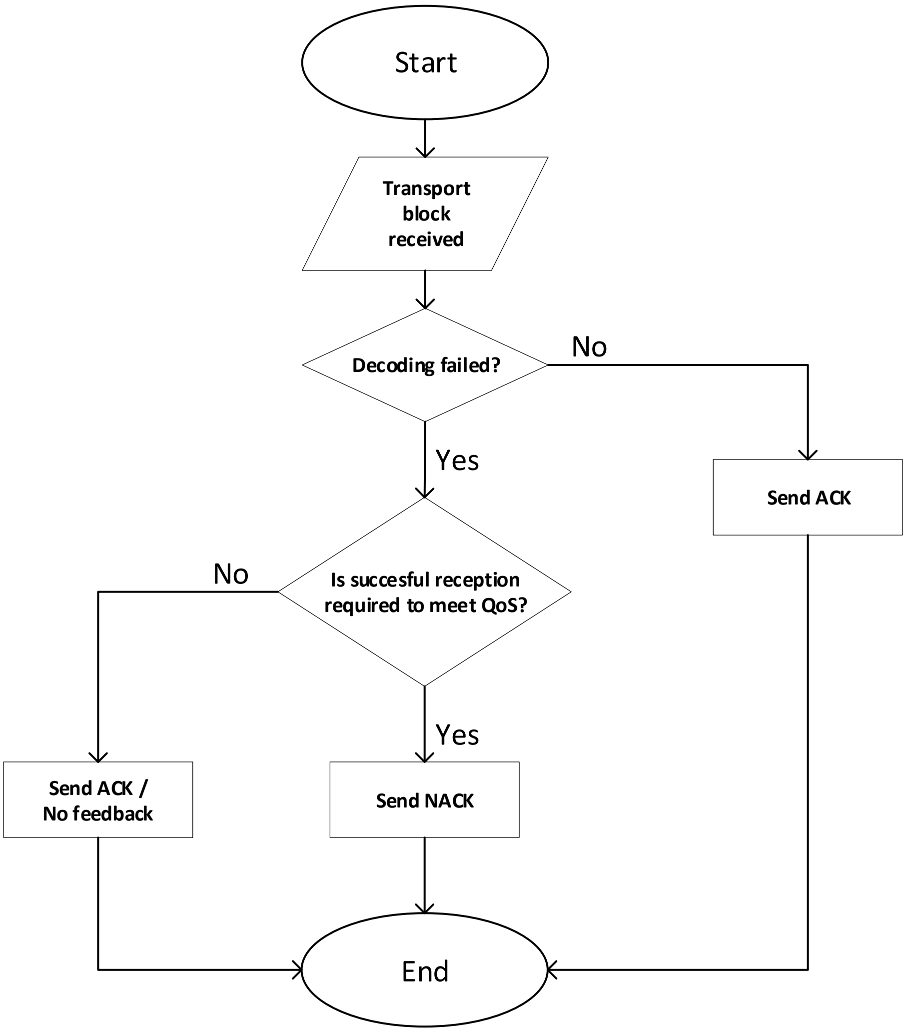

Figure 4.3-1: MAC protocol enhancement to HARQ feedback scheme according to QoS parameters.

According to the proposed feedback scheme, the capacity of the communication channel can be improved when the QoS parameters indicate user satisfaction. The radio performance and user perceived QoS can be aligned by the type of delivered service, since different performance situations (to retransmit or not) will have different impacts on QoS. This approach is suitable for different RLC modes and is especially suitable for streaming multimedia traffic in PTM communications where users are in different conditions. The transmission mode and retransmissions are not all the time driven by the HARQ process of the user in worst condition if the QoS is otherwise acceptable. The proposed scheme allows for service and media specific extensions which could consist of methods for user grouping, multiplexing the feedback channel, user selection for feedback etc.

4.3.2 Link adaptation for PTM

The performance of wireless systems depends on the conditions of the radio links. In order to cope with the changing conditions of the radio links and provide minimum QoS for the services consumed by a user terminal, proper MCS have to be chosen. The mechanism that executes the process of dynamic adjustment of these schemes is known as link adaptation.

Instant modification of MCS following change of radio channel (due to fast fading, etc.) is usually termed as inner loop link adaptation. Modified MCS based on the current radio channel conditions are typically applied in the following transmission time intervals. Hence, the modified MCS does not necessarily lead to improved performance. In such cases, it is crucial to use packet re-transmission (e.g., via HARQ) and perform modification of MCS via outer loop link adaptation techniques.

In the existing 3GPP specification for LTE-A and NR, link adaptations are typically applied for unicast communication and their suitability with broadcast and multicast communication is quite complex and it has been open area of research for quite some time. Moreover, 5G systems are expected to work with significantly large antenna arrays (as compared to LTE-A); this adds more dimensions to the complexity of multicast and broadcast systems that conventionally use simpler antenna array configurations. Section 5.2.2 presents detailed system level simulation-based study of link adaptation in 5G-Xcast system with consideration of higher antenna array configurations.

4.3.3 2nd Layer of FEC in RAN for PTM

Motivation

In the current LTE-A PTM specification HARQ feedback is not used. Proprietary implementation of dynamic link adaptation based on Channel Quality Indicator (CQI) feedback is possible for SC-PTM e.g. based on the worst UE in the cell. Based on these two restrictions a rather large margin has to be applied in selection of the MCS leading to inefficient use of the radio resources. In fact, 3GPP performed a detailed study in 3GPP TR 36.890 [6] on PTM with group-based uplink feedback for link adaptation and HARQ. Moreover, the HARQ feedback messages are reported from each UE to the network whenever a packet is received. The number of CQI and HARQ ACK / NACK messages scale with the number of UEs, leading to a high feedback load in scenarios with high number of users where PTM is typically a suitable option. Even more importantly, the HARQ-based scheme of 3GPP TR 36.890 [6] becomes very inefficient as the number of UEs grows as packet loss events at different UEs are largely statistically independent such that different UEs will typically ask for retransmissions of different packets.

The work in [13] proposed exclusion of the HARQ ACK / NACK feedback and use of only CQI feedback to achieve an improved performance via enhanced outer loop link adaptation techniques, but by construction lacks the capability to deliver data with high spectral efficiency with very high reliability, as there are no means to reliably fix packet losses e.g. due to channel variations not predicted by the CQI reports. Hence, an alternative error correction scheme with minimal overhead of feedback messages that at the same time provides high reliability is desirable.

Accordingly, an alternative technique that can provide the required performance via FEC schemes is proposed. It is based on Random Linear Network Coding (RLNC) which is selected due to its suitability for radio channels that induce packet losses, which is described by T. Ho, et al [14], and the flexibility of decoding with or without packet re-ordering as long as the required number of network coding PDUs is available at the receiver. Unlike block codes such as Raptor codes [5], RLNC offers the capability to perform successive en- / decoding and recoding [15]. The entailed feature of recoding makes RLNC interesting option to scale it to co-operative / Device to Device (D2D)-assisted broadcasting, which is however beyond the scope of this deliverable.

Protocol function implementation for 2nd layer of FEC

This section discusses the feasible options for implementing 2nd layer FEC in the 5G-Xcast protocol stack.

Feasible location for FEC Sublayer

The viable candidate locations to install FEC sublayer function are demonstrated in Figure 4.3-2. With RLNC, one of the necessary requirements for decoding is that received RLNC PDUs have fixed size (which is a design parameter). In other words, reception of variable length FEC PDUs from the same generation sequence is not suitable for decoding. As such, FEC sublayer location #1 is flexible enough to perform RLNC en- / decoding under the constraint of fixed FEC PDU size. On the other hand, FEC sublayer location #2 and #3 are not feasible candidates as both options don't guarantee forwarding of fixed FEC PDU sizes to their respective lower layers, as described in the following subsections.

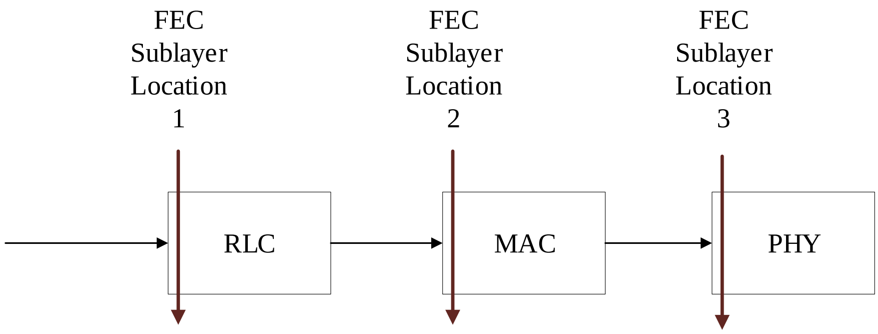

Figure 4.3-2: The viable candidate locations to install FEC sublayer function.

Figure 4.3-3 demonstrates the limitation of installing FEC sublayer functions at the entry of MAC sublayer in the radio access. Herein, RLC PDUs will be inputs to the FEC sublayer function. The generated FEC PDUs will have fixed size equal to the maximum of RLC PDUs plus FEC header information, as shown by the blue dashed boxes in the figure. However, the FEC PDUs will in general not be able to fit into the transport block provided by the lower layers.

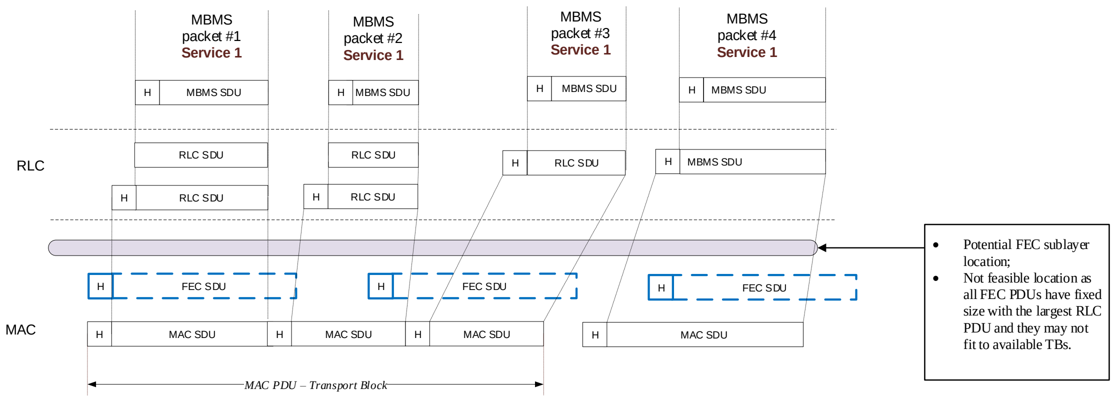

Figure 4.3-3: Potential limitations if FEC sublayer function is installed in MAC layer.

Figure 4.3-4 demonstrates the potential limitations if FEC sublayer functionality are installed as one of the initial physical layer procedures. As described in [4], 5G-NR physical layer has a set of procedures that perform 1st layer of FEC to provide bit-level robustness of transmitted data against lossy channel conditions. Herein, the major procedures of the specified error correction scheme are segmentation of a transport block into equally sized code blocks of a given maximum size, Cyclic Redundancy

Check (CRC) for decoding failure detection in each code block, and use of Low Density Parity Check (LDPC) codes for error correction on data channels. The potential location for a 2nd layer of FEC in this case is after segmentation of the transport block into code blocks. Generally, the 2nd layer of FEC packets should be distributed across different transport blocks to provide more robustness against fading processes. However, based on the characteristic of equally sized FEC PDUs across an entire generation the generated FEC PDUs would in general not fit into the allocated physical transmission resources, unless the amount of resources allocated to every transmission is selected such that it can, without significant padding, carry an integer number of complete FEC PDUs[4].

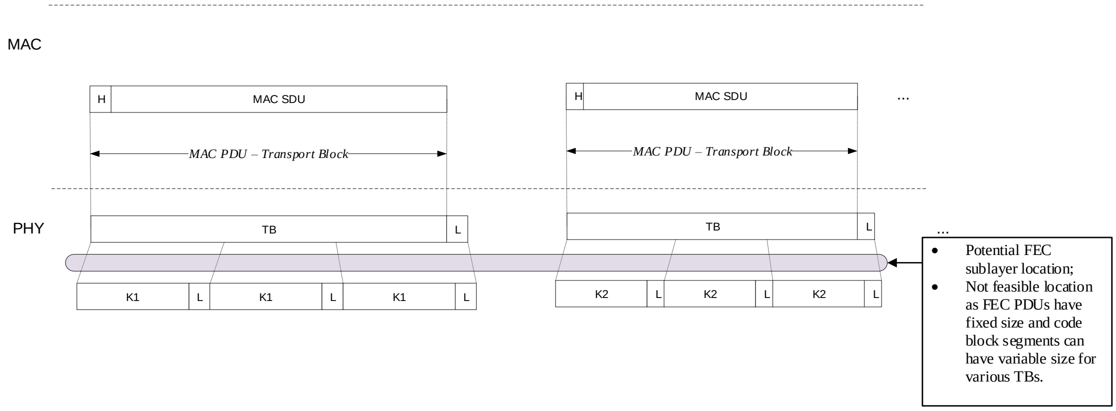

Figure 4.3-4: Potential limitations if FEC sublayer function is installed in physical layer.

Options of FEC Sublayer

Assuming the feasible FEC sublayer location #1, two major options of FEC sublayer functions are investigated as follows.

Figure 4.3-5 shows the first feasible option to perform RLNC functions inside the radio protocol at FEC sublayer location #1. Herein, MBMS packets, which in general have variable sizes, are received at the FEC sublayer as FEC SDUs. Then, the RLNC encoder generates at least as many fixed size FEC PDUs as the number of input FEC SDUs. In this case, the size of an FEC PDU is the maximum of sizes of the encoded FEC SDUs, which directly constitutes a disadvantage of this approach. The major advantage of this option is that it allows instantaneous encoding based on the available FEC SDUs.

[4] Note that the PTM design is based on 5G NR protocol even though NR design so far focused for unicast. In 5G-NR, concatenation is not any more done at the RLC but is coupled with multiplexing function at the MAC layer. Hence, based on the available resource at the lower layers, the MAC layer multiplexes RLC PDU (full PDU or segment) to constructs the transport block. This transport block size can vary based on the physical resource allocated by the lower layers. However, if a fixed physical resource allocation is assumed as in the case of LTE-A's MBSFN where dedicated time-frequency resources are used with fixed MCS, the 2[nd] Layer FEC can be possible to implement at top of the physical layer since fixed time-frequency resource allocation and fixed MCS leads to the same transport block size.

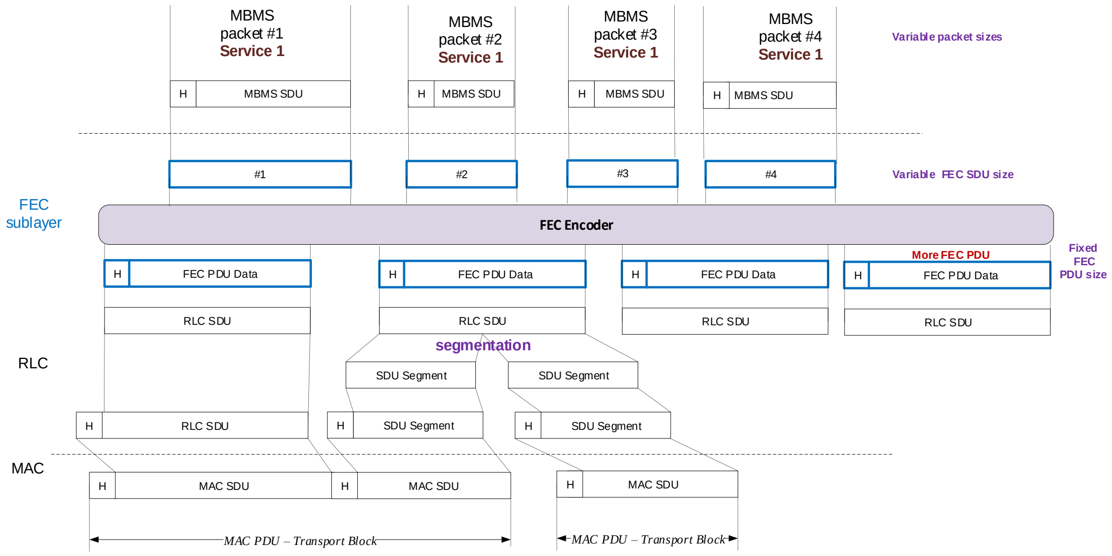

Figure 4.3-5: 1st feasible option for RLNC-based FEC functions placement above RLC layer (location #1).

Figure 4.3-6 describes the second feasible option to perform RLNC functions inside the radio protocol at FEC sublayer location #1. In this case, a fixed SDU size is configured with the same size as the fixed size FEC PDU payload. As a result, MBMS packets from the higher layers are segmented and / or concatenated to fit into the fixed-size FEC SDU, e.g. see packet #2, #3, and #4 in the FEC SDUs. Then, the FEC SDUs are encoded by the RLNC encoder to generate FEC PDUs. The main drawback of this option is the fact that FEC SDU sizes are fixed and FEC SDUs need to be filled with complete or segments of incoming packets. If an FEC SDU is only partially filled, it waits for more incoming packets before it is passed on to the encoder. In essence, it may incur delays in the process. However, a minimum affordable delay can be configured to stop waiting for incoming packet and use padding instead.

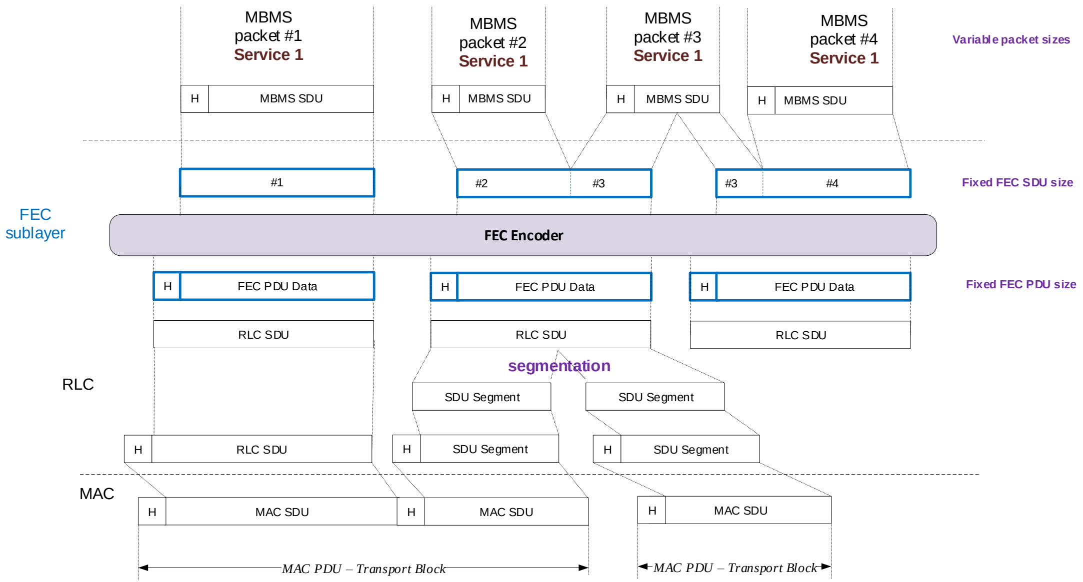

Figure 4.3-6: 2nd feasible option for RLNC-based FEC functions placement above RLC layer (location #1).

Due to its efficient radio resource utilization capability, option #2 (cf. Figure 4.3-6) is selected as a way forward in the implementation of RLNC-based 2nd layer of FEC in RAN for PTM communication.

Proposed FEC implementation

The main requirement for a UE to decode RLNC encoded data is to receive at least as many FEC PDUs as the number of encoded FEC SDUs. However, some FEC PDUs can be lost due to lossy wireless transmission channel. Hence, a certain number of extra FEC PDUs will have to be sent to the UE to compensate for the loss of packets. Existing approaches like the AL-FEC standardized for LTE-A do this only in a pre-emptive manner, which may transmit more than needed in some situations and still not be sufficient in others. Hence the proposal of this work is to use feedback from the UEs to signal how many more PDUs would be required. While the work in [28] proposes the use of sliding window to optimise the en- / decoding complexity in the application layer, it is proposed herein to stick to the use of a sequence of generations of fixed size successive en- / decoding in order to maintain en- / decoder history without incurring delays related to block-wise encoding.

Figure 4.3-7 depicts a simplified functional diagram for 2nd layer of FEC in RAN. Herein, higher layer data units are grouped into generation sequences upon which successive encoding is performed. Then, the encoded FEC PDUs are forwarded to the lower layers to be transmitted to UEs via the gNB multicast / broadcast channel. After UEs successfully receive the FEC PDUs from lower layer, the generation sequence of the FEC PDU is read from the PDU header and the corresponding decoder is used to perform decoding in order to extract service data units which are forwarded to higher layers.

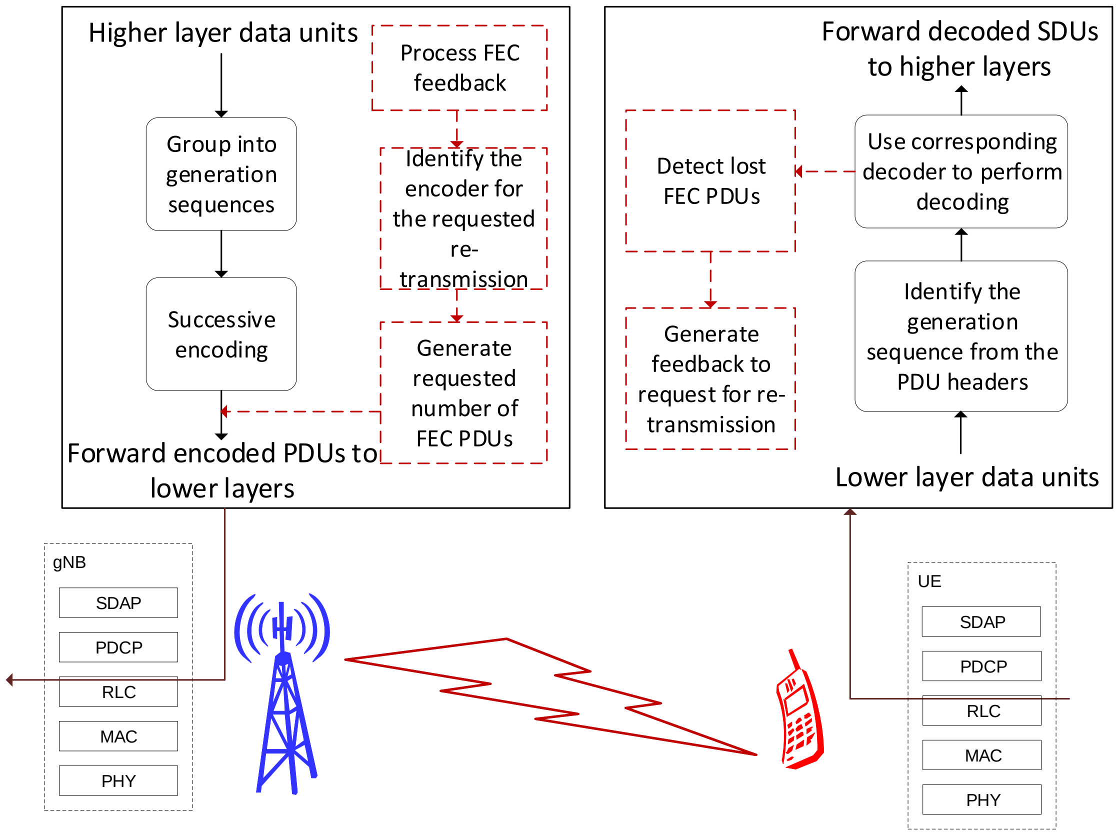

Figure 4.3-7: Simplified functional diagram for 2nd layer of FEC in RAN

If the UE is unable to decode all FEC SDUs of a certain generation after reception of a given number of FEC PDUs, it can use uplink feedback to signal to the network the number of extra FEC PDUs required for that generation. Then, the network transmits additional FEC PDUs from the notified generation, doing so again over the multicast / broadcast channel, which is a clear improvement over the conventional packet specific HARQ considered in [6]. Unlike HARQ ACK / NACK feedback messages that are triggered with every reception of a packet, the FEC uplink feedback is triggered only if the UE is unable to decode after reception of FEC PDUs that are outputs of a successive encoding of SDUs from a certain generation sequence. For efficiency this checking and reporting can be restricted to be performed only with a certain periodicity depending on the latency requirements of the service, e.g. 50ms.

In this process, the network can take into account already transmitted additional PDUs if multiple FEC feedback messages have been received from different UEs for the same generation, as additional PDUs requested by one UE are again multicasted / broadcasted and may hence also be received by other UEs. One aspect of such an implementation is that the network tracks the count of transmitted PDUs for the last N generations. Ideally, the network would track all previous generation; however, it is expensive in terms of memory requirements to maintain the entire history, and may also not be required depending on latency requirements for the service. Hence, maintaining the count of additionally sent PDUs for the last N generations, which is a design parameter, is indispensable. Great care must be taken with the computation of how many additional PDUs are required in order to provide the truly required number without on the other hand overloading the transmit buffer with an exorbitant number of additional PDUs.

4.3.4 Cross-layer link adaptation in coordination with higher layer error correction schemes

Motivation

Conventionally, broadcast / multicast in LTE networks is operated with a quite static configuration of the layer 1 MCS to provide sufficient robustness against fading channel variations since ACK / NACK based error correction methods do not work well with multicast / broadcast services. For further protection against fading variation in LTE, 3GPP has specified application layer FEC (AL-FEC) based on Raptor codes in [5] as a higher layer EC scheme. Herein, AL-FEC is assumed to be conventionally implemented at the application servers above the UDP / IP layer and it allows more flexible MCS setting alleviating conservative MCS. However, the two FEC layers (PHY / MAC and AL) are operating independently and, therefore, often cause a disproportionate radio resource utilization which in turn causes low spectral efficiency, as well as more adverse interference situations and worse system performance. As long as UEs don't provide the network with feedback, MCS selection is still based on the potentially worst condition that a served UE may be; hence, a conservative MCS which only provides low spectral efficiency is used.

In 5G-Xcast project, a 2nd layer of EC in RAN described in Section 4.3.3[5] is investigated as higher layer EC, which can be located e.g. above or in RLC layer [53]; herein, the mechanisms of MCS modification in co-ordination with 2nd layer of EC in RAN, needs to be investigated in order to achieve high overall spectral efficiency.

[5] Note that the terminology '2nd layer EC in RAN' refers to the same scheme described as '2nd layer FEC in RAN' in Section 4.3.3. The term EC is introduced due its general sense that error correction can be forwarding looking (e.g., FEC) or it can be with feedback (not forward looking).

Hence, a mechanism that co-ordinates MCS modifications via cross-layer Link Adaptation (LA) at a scale comparable to higher layer EC (such as AL-FEC or 2nd layer of EC in RAN) operation is indispensable in order to maximize efficiency of the network while providing robustness that fulfils Quality of Experience (QoE) requirements.

Implementation

Higher-layer EC such as AL-FEC or layer 2 EC perform successive or block-wise encoding on a block of EC Service Data Units (SDUs) which are typically grouped into blocks / generations. The scheme proposes a practical mechanism for co-ordinating MCS modification with higher layer EC schemes.

To practically realize the proposed scheme in a real wireless communication system, it is proposed that UEs perform measurement on EC Protocol Data Unit (PDU) loss rate to monitor the EC PDU Loss Rate (PLR) within higher layer FEC block. Event-based or periodic reporting can be used to deliver the PLR measurements from UEs to the network. Then, the network processes PLR measurement reports from multiple UEs to adjust MCS settings that are to be used for PTM bearers that are applied for transmission to all UEs being served by the network.

Since UEs have various channel conditions, care should be taken in the MCS adjustment since PLR measurements are received from multiple UEs. First of all, the PLR measurement report from each UE should maintain a measurement report Sequence Number (SN) to provide information about the EC block sequence number that is measured. Before the network performs MCS adjustments, PLR measurement reports of the same SN should be received from all UEs being served by the PTM bearer. To this end, the network maintains a timer, referred in this document as ‘multiple user report timer’, which is started when the PLR measurement of a new SN is received from a UE.

Figure 4.3-8 demonstrates the functional description at the RAN to process PLR measurements from multiple UEs being served by PTM bearers. Upon reception of PLR measurement reports, the RAN extracts the report SN, ‘rx_report_SN’, and the PLR measurement value, ‘rx_PLR_measurement’. Then, it updates its current report SN, ‘current_SN’, with the received report SN, ‘rx_report_SN’. Next, the current report SN ‘current_SN’ is compared with previous report SN ‘previous_SN’. If ‘current_SN’ is greater than ‘previous_SN’, it signifies reception of the first new measurement report from one of the UEs. Consequently, the network starts ‘multiple user report timer’ to monitor measurement reports with the same SN from multiple UEs. Besides, the network initializes the current maximum PLR value, ‘current_max_PLR_value’, with received PLR measurement ‘rx_PLR_measurement’; and it updates previous report SN ‘previous_SN’ with the current report SN ‘current_SN’. On the other hand, if the current report SN is not new, i.e., ‘current_SN’ is not greater than ‘previous_SN’, the network compares the received PLR value, ‘rx_PLR_measurement’ with the current maximum PLR value, ‘current_max_PLR_value’. If the received PLR value, ‘rx_PLR_measurement’ is higher than the current maximum PLR value, ‘current_max_PLR_value’, the network updates the current maximum PLR value, ‘current_max_PLR_value’ with the received PLR value, ‘rx_PLR_measurement’. If not, the network maintains the current maximum PLR value, ‘current_max_PLR_value’.

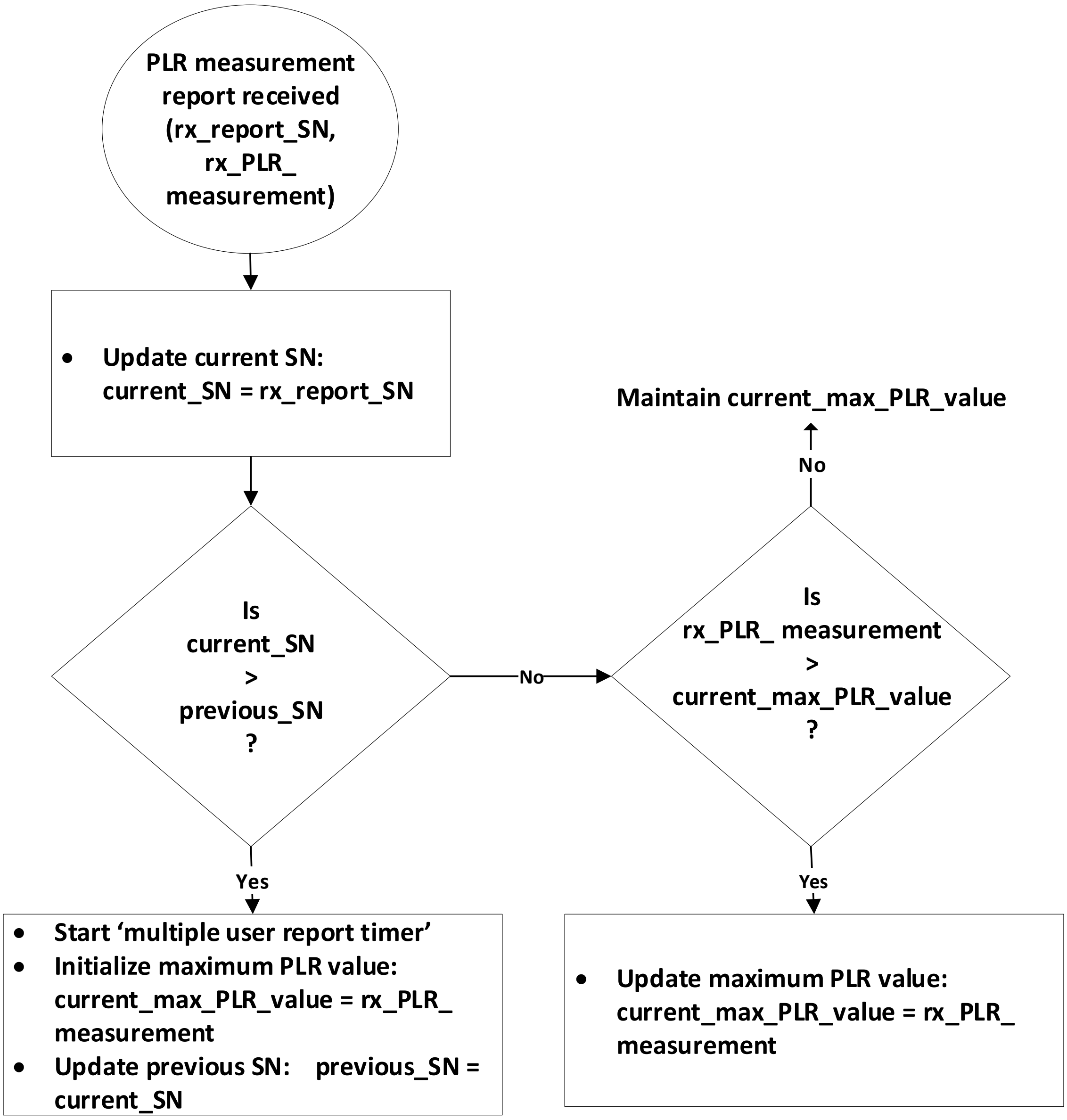

Figure 4.3-8: Functional description at the RAN to process PLR measurements from multiple UEs being served by PTM bearer.

Upon expiry of timer 'multiple user report timer', the network performs a decision to update MCS settings. Accordingly, the network planner or operator should be able to define PLR value thresholds at the radio access to compare with the current maximum PLR value. Herein, higher threshold, 'threshold_higher', and lower threshold, 'threshold_lower', are defined to assist the network on the decision of decreasing and increasing MCS settings, respectively. To avoid fluctuation effects on the MCS settings, the lower threshold, 'threshold_lower' should be configured with a value lower or equal to the higher threshold, 'threshold_higher'.

Figure 4.3-9 describes the MCS modification procedures upon expiry of 'multiple user report timer'. Herein, if the current maximum PLR value 'current_max_PLR_value' is greater than higher threshold 'threshold_higher', the MCS setting is decreased by MCS decrement offset 'mcs_delta_offset_decrement'. If not, the current maximum PLR value 'current_max_PLR_value' is compared with the lower threshold 'threshold_lower'. If the current maximum PLR value 'current_max_PLR_value' is smaller than the lower threshold 'threshold_lower', the MCS setting is incremented by MCS increment offset 'mcs_delta_offset_increment'.

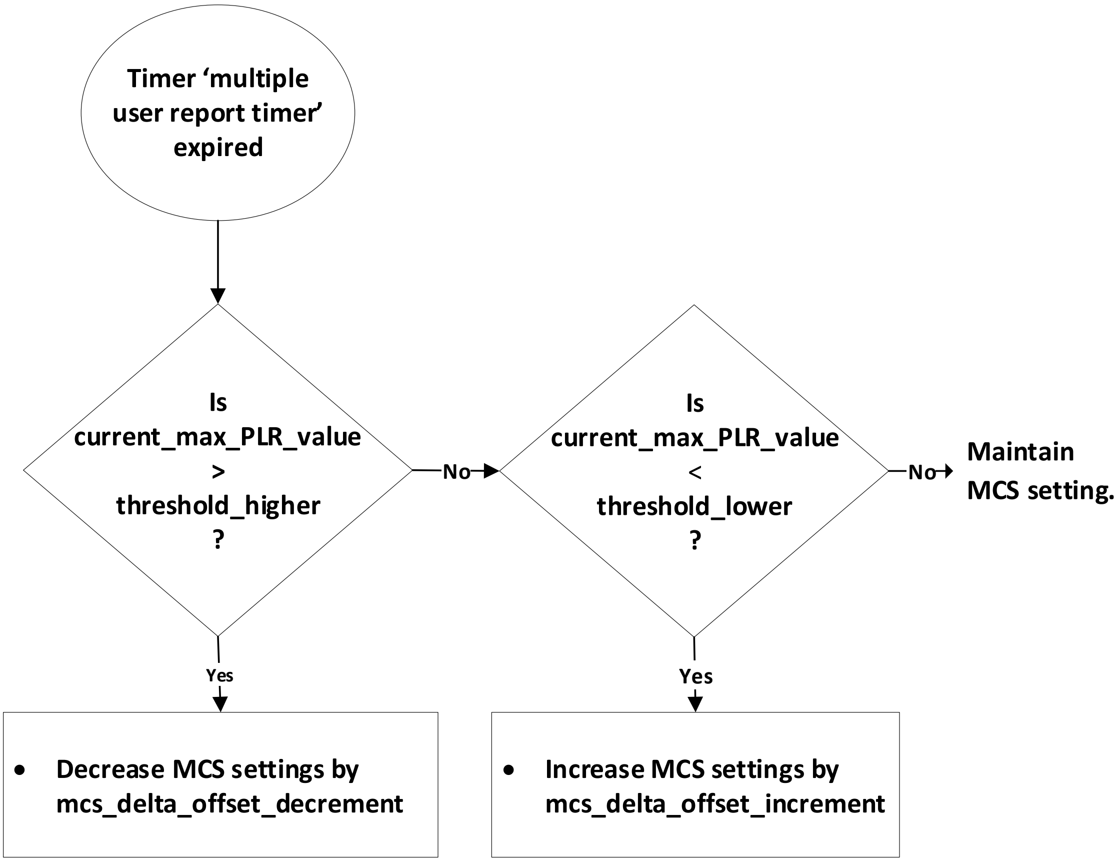

Figure 4.3-9: MCS modification procedures upon expiry of 'multiple user report timer'.

4.4 Efficient use of radio transmission methods

4.4.1 Protocol level analysis of dynamically defined multicast area

This section is described in the 5G-Xcast deliverable D3.3 [45].

4.4.2 Coverage impact on resource efficiency aspects

The realistic coverage analysis and results presented in sections 5.1.1 and 5.2.1 are the basis of the following considerations on resource efficiency. As shown in section 5.2.1, the number of potential users (TV sets) of a 5G mobile video streaming service in partial substitution of Digital Video Broadcasting – Second Generation Terrestrial (DVB-T2) transmissions varies dramatically from few TV sets per cell to hundreds, since the substitution can be applicable only when DVB-T2 signal is below threshold, therefore involving some cells only marginally. Moreover, an investigation to determine the neighbouring relations between cells to better characterize the impact of the 5G-Xcast PTM features in relation to Single Frequency Network (SFN) allocation shows that cells with a medium to high number of potential TV sets (candidates to activate PTM features) tend to have several cell-neighbours. In this scenario the usage of a MC-MM or terrestrial broadcast solutions with an SFN allocation strengthen the broadcast signal and reduces interferences between cells, obtaining a higher resource efficiency of the network. Therefore, this kind of allocation seems to be preferable in the scenario simulated. Generally speaking, the use of SFN MC-MM or terrestrial broadcast (key innovations proposed by the project) can be then suggested for a certain number of cells from the ones where the service will be always on (with a high number of potential TV sets) to those where the PTM is dynamically activated when needed (with a medium number of potential TV sets). To better understand the initial configuration or dynamic switching between PTP and MC-MM transmission, a methodology has been devised to determine the exact number of cells where PTM solutions are preferable. As described in sections 5.1.1 and 5.1.2 the methodology utilizes audience data considering two 2018 periods when the maximum and minimum number of TV sets are on, as well as for the yearly average. For each considered scenario, the TV channels with minimum and maximum audience share and the average audience share is taken into account.

The step-by-step analysis of the considered scenario in the coverage simulation presented in sections 5.1.1 and 5.1.2 can be generalized and used by an operator when launching a streaming TV service to decide in which cells activate PTM features (MC-MM or terrestrial broadcast).

1. For each cell, knowing with a planning tool the area size and type where the service can/will be activated, the number of potential reachable receivers can be determined:
    - from area size & area type population density, the number of residents reached can be derived;
    - from number of TV-sets per population, the number of potential TV sets reached can be derived.
2. For each cell, from TV audience data the number of viewers can be established:
    - from the percentage of TV sets on in the observed period, the number of TVs on for each cell can be derived;
    - from audience share (percentage) of the considered channel or set of channels to be streamed in the observed period, the number of viewers per cell can be derived.

The criteria of activation can lead to widespread activations (if based on audience peaks, maximum TVs on period), limited activations (if based on audience lows, minimum TVs on period) or can be based on average data. Finer dynamic period activations can be considered knowing daily audience behaviour. These first assumptions could be subsequently validated by actual requests of streaming in a live network. It can be noted, though, that in the terrestrial broadcast solution, where no feedback channel is present, receivers may not have uplink capability. In such case, the requests of streaming and the actual number of viewers per cell exploiting the service would not be available.

While it’s up to operators to find a suitable criterion to decide in which cell activate a PTM feature, some general consideration on resources allocation can be drawn. The different PTM solutions (MC-MM) in SFN will have different overhead and be more resource saving with respect to PTP when multiple users require the streaming (with MC-MM being the leanest). Even though PTM solutions keep resources usage low, they are uselessly burdensome in cells with very low (less than 1) number of viewers (e.g. a number of viewers of 0.1 means that a transmission is needed only 10% of the time). In this scenario broadcast resources are underutilized and therefore somehow unexploited. The proposed methodology can clarify where (in which cell) and when (under what circumstances) saving with PTM solutions is significant, and where/when it is not and PTP approaches are preferable. The PTM-PTP policies can be implemented by operators considering a variety of aspects; the methodology presented shows the boundaries within which operators should move.

4.4.3 Partial HARQ retransmission for broadcast

In 5G, the terminals receiving a broadcast / multicast transmission also have broadband (one to one) connections with the base station. A possible improvement would be to consider an uplink transmission of NACKs in case the receiver does not correctly decode the packet as described in Section 4.3.1. This can be done by adding HARQ process, in which the first transmission is not in a point-to-point mode but in a broadcast mode. Once this broadcast detection fails, a NACK is transmitted and the HARQ process takes over the retransmission and decoding of this packet by using, for instance, incremental redundancy. The advantage of this type of scheme is that the base station does not need to build on the weakest receiver in the multicast group. It may be more advantageous to use a higher modulation and coding scheme, even if some of the transmissions need to be repeated for only few users.

The scheduler selects one transmission rate to satisfy the maximum number of users. However, users near the cell edge experience bad channel quality due to power attenuation. In this case, the scheduler can choose between two extreme cases: either to use a data rate that fits the good channel quality cluster or to use a low data rate that can be decoded by all users, including the ones at the cell edge. The former case excludes users at the cell edge; the latter case is inefficient because a low data rate is imposed on all users. The goal is then to find the best compromise to maximize the delivery time.

The total time needed to deliver a correct packet to all users is given by a sum of the time needed for the broadcast phase and the time needed for the unicast (retransmission via HARQ) phase. The retransmission time depends on the number of frequency resources allocated to the HARQ process. For simplicity, it is assumed that the same number of resources is allocated to both the broadcast and the HARQ components.

In an ideal scenario, where all users have the same channel quality, the base station can choose a retransmission rate that guarantees reception for all users. As the supported data rates across users may vary, it is difficult for the base station to find a rate that fits all users, unless it selects the lowest rate that all users can decode. However, this penalizes users with good channel conditions and increases the total delivery time. In 5G-Xcast, an optimisation algorithm is proposed that selects the best data rate based on the channel qualities of the users [56].

A single cell scenario using multicarrier transmissions with Orthogonal Frequency-Division Multiplexing (OFDM) is considered. It is assumed that mobile receivers are located randomly within the cell. Receivers suffer both from slow fading due to attenuation and shadowing as well as fast fading. Receivers close to the base station will have a better signal-to-noise ratio (SNR) than users on the cell edge. We assume that the base station has knowledge of all user channels for each subcarrier, for example CQI.

A mixed broadcast / multicast protocol divided into two steps is considered, as shown in Figure 4.4-1. In a first broadcast step, the base station transmits a packet to all users using the same carrier (group of adjacent subcarriers) frequency. Some users will be able to decode the packet and transmit an ACK whereas some other users will not be able to decode the packet and shall transmit a NACK. In a second step, the base station addresses the users that transmitted a NACK using unicast. Each user can be addressed with a suitable (robust) MODCOD (Modulation and Coding) that will enable it to decode the packet after one retransmission.

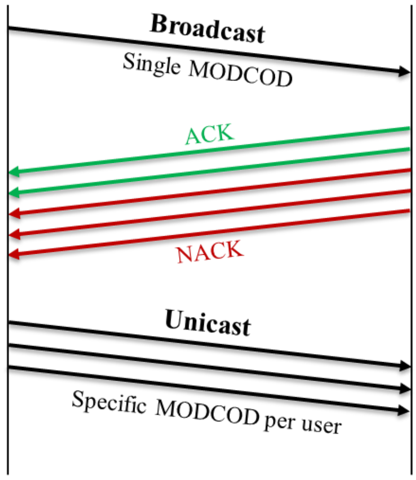

Figure 4.4-1: Mixed broadcast / multicast protocol with retransmissions

4.4.4 Improving the UE side broadcast and multicast receiving efficiency

In 5G NR, a UE is able to report its experienced channel quality through uplink Channel State Information (CSI) [32]. The channel quality is indicated by a numeric value of CQI, indexed from 0 to 15, in a quality increasing order, where CQI 15 stands for the best channel quality and CQI 0 means the UE is out of the range of the cell served by the Base Station (BS). The BS may configure on which channels the UE should report back the CQI indexes. Ideally the UE will report the highest one which stands for the best quality channel available. The feedback mechanism provides a means to perceive the quality of multicast channels and help the decision on MCS and Transport Blosck Size (TBS) selection.

A different approach to that in the previous subsection is taken which is computational, by taking into account the channel quality distribution among the broadcast and multicast service receivers, and proposes an optimal channel allocation and MCS/TBS selection scheme on the available one or more multicast channels [57].

In NG-RANs, during a scheduling period, each multicast channel uses a designated MCS/TBS together with other associated transmission parameters. The waveform is such configured to achieve a most efficient channel usage. On the other hand, the UEs may report different CQIs to the BS. Suppose there are s CQIs reported by a group of UEs. The BS needs to select the most efficient MCS/TBS for the multicast channel(s).

The first case is that there is only one multicast channel available. To ensure a fully satisfying reception by all UEs in a multicast group, as a straightforward solution, the BS may select MCS/TBS against the lowest CQI reported by group UEs. E.g. we have 100 UEs reporting different CQIs that may range from 1 to 15. Then we use CQI 1 as the perceived channel quality and use, for example, QPSK and a code rate 78/1024, as specified by 3GPP TS 38.214 [32]. All UEs will use same MCS including those even reporting a CQI as high as 15, who should tune themselves downwards to match the low rate. Alternatively, according a Service Level Agreement (SLA) signed with the users, an MCS against higher CQI could be selected, while scarifying the poor channel quality UEs and achieving an overall satisfying reception.

The next case is there is more than one available channel. This is the case where the BS has redundant resources that can be employed to transmit the multicast service. The BS has more choices in using different MCS/TBS for different channels. For a simple example, if some UEs report CQI 1, while the other UEs report CQI 15, then the BS can use QPSK and code rate 78/1024 on one channel, and use 64QAM and code rate 948/1024 on another. Then the UEs reporting CQI 1 tune themselves to the QPSK channel and UEs reporting CQI 15 tune themselves to the 64QAM channel to achieve the best receiving efficiency.

The question coming up is how to allocate the UEs into the available channels. The UEs need to be subgrouped and each subgroup is allocated into a dedicated separate channel. The term “subgroup” is used because all the UEs in question are already in a same multicast group.

UEs with the same CQI can be put into a subgroup. If the number of reported CQIs is greater than the number of available channels, then some UEs with different CQIs need to be put into a same subgroup.

Some UEs with smallest CQI may be excluded to some extent, in order to achieve an overall maximum throughput without violating coverage requirements. The excluded UE can still try to use a higher rank MCS but will expect higher BLock Error Rate (BLER). Even though with excluded UEs, the agreed service coverage should still be satisfied (e.g. at least 95% of UEs are well served). The throughput is traded off to guarantee the service coverage.

Here is an example illustrating the problem. Suppose there are 119 UEs reporting CQIs ranging from 1 to 15. At the BS side there are 3 available multicast channels. Also there are additional objectives to guarantee at least 95% of the UEs’ (equivalent 114 UEs’) reception, and maximise the whole multicast group’s throughput. The question is how to allocate the 119 UEs into the 3 channels, i.e. against which CQIs to select the MCS/TBS. The question is formalized as below:

$$\mathrm{T} = \sum_{cqi=1}^{15} \mathrm{f}(cqi) \cdot \mathrm{N}_{cqi}$$

where N_{cqi} is the number of UEs reporting *cqi*, and *f(cqi)* is the weight function, meaning that for a UE using a channel correspondent to *cqi*, its gain is *f(cqi)*. Here the UE’s real channel quality must be equal to or greater than *cqi* so as to achieve the full capacity of the channel. In this case *f(cqi)* is the throughput gained by the UE. T is the valuation function and the question is to maximize T. Alternatively there can be other valuation functions, such as BLER in average.

The algorithm is described in Figure 4.4-2 and Figure 4.4-3.

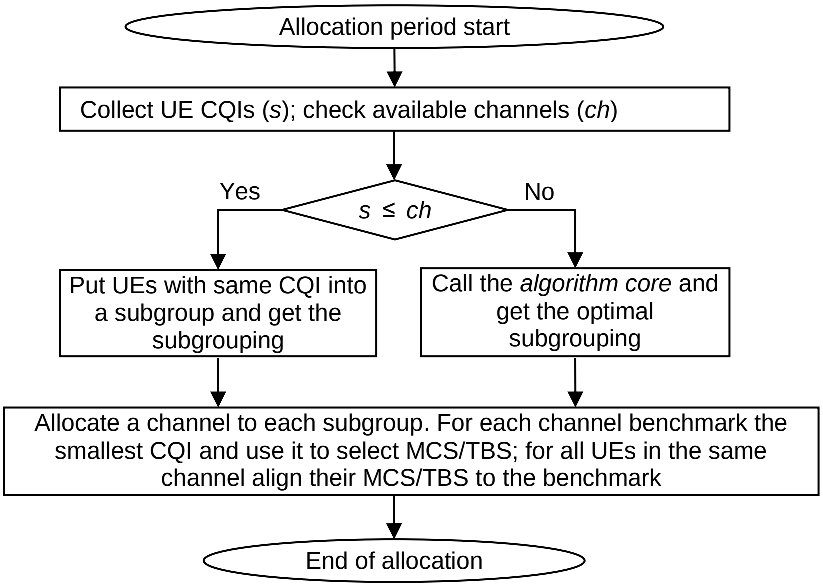

Figure 4.4-2: General channel allocation algorithm flow chart

1) Use *gr* as iterator with initialized value1, 1 ⩽ *gr* ⩽ *ch*.
2) Consider there are only *gr* CQIs to subgroup, there can only be *gr* subgroups.
3) Suppose we have got the optimal subgrouping for *n*-1 CQIs, solve the case *n*:
4) Memorize optimal subgrouping with each CQI as highest one, thus we have memorized *n*-*gr* subgroupings, starting from *gr*, *gr*+1, *gr*+2, …, *n*-1.
5) If CQI *n* will be selected, then we need to use the optimal *n*-1 subgrouping with *gr*-1 subgroups.
6) If CQI *n* will not be selected, then we need to get the best from the *n*-1 subgrouping with *gr* subgroups.
7) Compare the results of 5) and 6), the better is the optimal *n*-subgrouping.
8) Incrementally execute from step 2) and get the optimal subgrouping.

Figure 4.4-3: Core algorithm of UE subgrouping

To further validate the effectiveness of the proposed method, calculations were conducted with comparisons to an ordinary single multicast channel allocation without any optimisation. The results further verify the observation obtained in previous work [33]. The acceptable effects come from the middle CQIs, as plotted in Figure 4.4-4, for the sample used in the calculations, between 4 and 12. However, as CQI against which the selected MCS increases, QoS for UEs with low CQIs are deteriorating. The covered UEs drop from 95.8%to 46.2% for the CQIs from 4 to 12. Although the proposed tailoring method is for multiple channel multicasting, it also works well for the single channel. A horizontal line is plotted which stands for the best T value that can be obtained with one channel (labelled “1 channel optimal”), which is for CQI 4 and the T value is 456, where 95.8% UEs’ service is guaranteed.

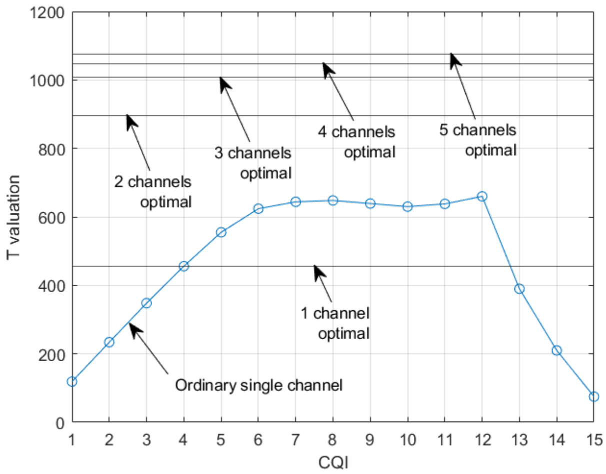

Figure 4.4-4: The T valuation for optimal multiple channel transmission vs. ordinary single channel transmission

As the proposed tailoring multicast's aiming at using multiple channels, the optimal T values for 2, 3, 4 and 5 channels are also plotted in Figure 4.4-4. As shown, 2 channels outperform all possible ordinary single channel configurations, with a T value 896 and the coverage of 95.8% UEs. Interestingly, the gains do not increase linearly with the number of employed channels, although a liner weight function is defined as *f(cqi) = cqi*. The extra gains brought in by a 4th or 5th channel become very marginal, where T value is 1046 for 4 channels and 1076 for 5 channels. In view of the minor profit of more extra channels, it is recommend that using the minimum amount of channels that just outperforms the ordinary method. This also depends on the CQI distribution among the UEs.

Obviously more channels are employed for a same multicast service at BS side and at the transmission side more resources are occupied. But that is a worthy cost to pay for the very desirable gain at the UE side, which is proved by the results represented by the T valuation. Furthermore, tuning into a most suitable channel for a UE saves energy and active time when receiving multicast service, especially when the UE is in a dual connectivity or multiple connectivity mode when it needs to allocate more receiving resources for other services, and even with other BS(s).

4.5 Trigger methods for MBMS reception in PWS applications

Since the UE is not aware of an (upcoming) broadcast of a PWS message, the RAN needs to trigger the UE to start MBMS reception.

Section 5.3.2 of 3GPP TS 36.331 [47] specifies a notification method in the paging message to trigger reception of Earthquake and Tsunami Warning System (ETWS) and Commercial Mobile Alert Service (CMAS) messages. This notification requires the UE to obtain SIB1 with scheduling information to acquire the SIBs that contain the ETWS or CMAS message.

An MBMS indication in the paging message is proposed to notify the PWS application in the UE to start MBMS reception.

The notification should either explicitly or implicitly contain the Temporary Mobile Group Identity (TMGI). The PWS application will request the file(s) for the PWS service for the TMGI from the MBMS middleware. A TMGI that is contained in the MBMS notification is more flexible since it allows broadcasting of multiple concurrent MBMS content.

An alternative method to trigger MBMS reception in the UE is to use Cell Broadcast messages that trigger the PWS application to initiate reception of the MBMS content. This method is outside the scope of the present document.

4.6 Spectrum sharing in 5G-Xcast

While state of the art on spectrum sharing can be referred to Annex D, this section focuses on prospect of spectrum sharing in 5G-Xcast.

The 5GXCast use cases differ greatly in terms of required coverage, bit rate and quality of service. The spectrum allocation options have been studied and analysed for each use case in different spectrum bands and with different spectrum allocation methods, ranging from exclusive licensing to spectrum sharing and unlicensed spectrum. The type of operator who would have most benefit in the selected combination of use case and spectrum assignment has also been studied.

The spectrum allocation and usage options are described under the following categories and allocation options: use cases (M&E1, M&E2, M&E3, PW1, Auto1, IoT1), spectrum bands (470-694 MHz, 700 MHz, 2.3 GHz, 3.5 GHz, 3.8-4.2 GHz, 6 GHz, 26 GHz and above), allocation/usage options (Nation-wide long-term licenses, Local and temporary licenses, CBRS, Licensed Shared Access, Concurrent Shared Access, Unlicensed spectrum.), and operator (MNO, broadcaster, other). The different spectrum bands, spectrum allocation methods and types of operators for the considered use cases have been studied. The spectrum bands are divided into three groups: coverage bands below 1 GHz, mid-capacity bands between 1 to 6 GHz and high capacity bands above 6 GHz. Sections 4.6.1 to 4.6.3 discuss these bands, section 4.6.4 the options for spectrum allocation and section 4.6.5 the operator types.

4.6.1 Sub-1 GHz bands: 470-694 MHz & 700 MHz

The frequency bands of mobile networks are traditionally divided into frequency bands by characteristics such as typical coverage and capacity able to be provided. Wide area coverage bands are generally accepted as best at frequencies below 1 GHz. At these frequencies, propagation over long distances is good and these bands are economical for a mobile operator to build out a good nation-wide coverage. The bandwidth in the coverage bands is narrow; hence, it is difficult to provide broadband connectivity or support large numbers of data-hungry applications in the same cell simultaneously. Capacity bands have been utilized for several tens of years now. In order to provide coverage bands for 5G, these lower frequency bands need to be cleared from existing use. The coverage frequency bands are difficult to share with other types of spectrum users and the primary spectrum assignment method for coverage frequency bands is exclusive licensing. In practice, many of the coverage frequency bands have been used by terrestrial television. The pioneer 5G coverage band globally is 700 MHz band, and it may be extended to cover lower digital TV UHF bands 470-694 MHz in the future.

4.6.2 1 to 6 GHz range: 2.3 GHz, 3.5 GHz, 3.8-4.2 GHz, and 6 GHz

The capacity bands begin from 1 GHz and extend to higher frequencies. In some cases, the frequency bands between 1 and 2 GHz may be used as coverage bands by the mobile operators. The capacity bands offer wider bandwidths than coverage bands making them possible for mobile broadband services. The cell sizes of the capacity bands are smaller than those of coverage bands making it easy to build high capacity network areas, but uneconomic to build nation-wide coverage. As the capacity bands are not expected to be deployed with full coverage, spectrum sharing with other spectrum users becomes feasible.

The mid-band of the capacity bands is limited to 6 GHz in the high end. The first pioneer capacity mid-band is 3.5 GHz. It will be extended to cover 3.4-4.2 GHz. Also the LSA band 2.3 GHz will be used for 5G and 6 GHz is being harmonized for unlicensed use. The countries which are able to clear these bands before assigning them to 5G can assign nation-wide licenses or in some cases a part of the spectrum is dedicated to private LTE/5G networks. Most countries will not be able to clear all mid-capacity bands and different spectrum sharing methods will be used depending on the characteristics of the incumbent spectrum user. For static incumbents, static sharing using license terms is the prevailing method and for the dynamic incumbents, dynamic spectrum sharing is required.

4.6.3 Above 6 GHz: 26 GHz and above

The high-frequency capacity bands are above 6 GHz. Although, the band naming begins on 6 GHz, the pioneer band is 26 GHz, and it will be followed by even higher frequencies. They are often called millimetre waves. The bandwidths are very wide compared to any other communication system allowing gigabit/second -level wireless bitrates. The connectivity between the base station and user equipment requires a line of sight, the cell sizes are very small and the beams can be very directive. The millimetre wave bands are very suitable for spectrum sharing. Italy is the first European country, which included club use -type of spectrum sharing as a part of the 26 GHz auction rules.

4.6.4 Spectrum allocation options

The considered allocation options are exclusively licensed spectrum, nation-wide long-term licenses, local and temporary licenses, and shared spectrum. Following the 5G Spectrum Position Paper [GSMA 5G Spectrum Public Policy Position. November 2018. https://www.gsma.com/spectrum/wp-content/uploads/2018/12/5G-Spectrum-Positions-1.pdf] of GSMA, the primary spectrum management approach for 5G remains exclusively licensed spectrum. Practically, all mobile network bands are currently on exclusively licensed bands. Spectrum sharing and unlicensed bands complement that. An option for assigning spectrum for industrial users are local licenses for private LTE/5G networks.

Linear TV services have been offered on UHF terrestrial TV band at 470 - 862 MHz for several decades and it unsurprising that the same frequency band is recommended also in this study. M&E1 shares the same basic characteristics as linear TV and due to that the mapping of linear TV and M&E1 are generally the same. Virtual and augmented reality requires very high bitrate and due to that they fit best to the highest capacity bands. Remote live production benefits from high uplink capacity. On the other hand, live production in a remote location needs coverage. The coverage is best achieved on the coverage bands and utilization of the current primary PMSE camera link band, 2.3 GHz for shorter communication distances, could be a practical combination. Public warning should reach as many people as possible, so coverage bands are preferred. The media services require more capacity than the coverage bands can offer, so the mid capacity bands could be used for providing them.

4.6.5 Mapping use cases, spectrum bands, allocation options and operators

Table 4.6-1 combines the results of the mappings in the previous tables. Linear TV and hybrid broadcasting fit best to the similar spectrum use as the TV services have been using for decades. The coverage bands below 1 GHz, nation-wide exclusive licenses having either broadcaster or MNO as the operator would work best considering also that societies have been using them for TV broadcasting. The virtual and augmented reality services require very high bitrates, which can only be provided on the highest capacity bands beginning from around 3 GHz. All spectrum allocation options are feasible. The operator for the services is most likely MNO, but other local operators can provide them in private LTE/5G networks, as well. Remote video production has two sides: one is remoteness and the other is bandwidth requirements of video. Remote can easily be translated to coverage band, i.e. 700 MHz and video production to 2.3 GHz which is used for that purpose by broadcasters and production companies. Any allocation method providing even a little bit higher availability than unlicensed should be considered. The spectrum license holder can be MNO, broadcaster or a private LTE/5G license holder. Public warning system requires highest coverage and availability limiting the choices to nation-wide exclusive licenses on 700 MHz and provided by MNO or broadcaster. Media services to vehicles could be provided in the 3.5 GHz or 6 GHz bands using any other allocation method but concurrent, which is expected here to be available only on 26 GHz. The media services to cars could be provided by broadcaster, MNO and other companies dedicated to roadside communications.

Table 4.6-1. Use case – spectrum band – allocation option – operator

	Linear TV	M&E1	M&E2	M&E3	PW1	Auto1
Band	< 1 GHz	< 1 GHz	> 3 GHz	700, 2300 MHz	700 MHz	3.5, 6 GHz
Allocation	Nationwide	Nationwide	All	Nationwide, local, LSA	Nationwide	All, but concurrent
Operator	Broadcaster, MNO	MNO	MNO, other	MNO, Broadcaster, other	Broadcaster, MNO	Broadcaster, MNO, Other
Notes						

PTM transmissions (broadcast / multicast) could present a more efficient delivery mechanism in many scenarios when compared to PTP transmission schemes (unicast). 5G-Xcast project develops architecture for PTM in 5G and has identified different use cases, or use case families, which cover the scenarios where the highest benefits of 5G PTM could potentially be achieved. The use cases belong to the following 5G vertical market sectors: Media & Entertainment, Public Warning, Automotive and Internet of Things. Different 5G use cases and applications differ greatly in terms of coverage, bit rate and quality of service they require. Thus, the combination of spectrum bands and spectrum quality they need is different in each use case.

This section has analysed spectrum allocation options in different frequency bands for the six different PTM use cases. The use cases have been analysed against the spectrum bands they could use, then the spectrum bands have been analysed against the different allocation options (ranging from exclusive licensing to spectrum sharing and unlicensed spectrum), and the use cases were analysed against the allocation options. Finally, all of these were brought together in use case - spectrum band - allocation option - operator mapping.

4.7 RRM with consideration of security

In this section, RRM algorithms are investigated from the point of view of RAN physical layer signal processing in the PTM scenarios, with consideration of physical layer security.

4.7.1 Current and potential PTM RRM with consideration of security

State-of-the-art investigations on the topic of RRM with service security either focus on the PTM or PTP transmissions. However, motivated by a growing consumers' desire for high-quality multimedia UEs (such as 4k hand-held devices and 3D augmented reality for the M&E vertical market sector), serving these UEs shall take into account: a multicast service, which is subscribed to by all users, and a confidential unicast service, which is subscribed to by a dedicated user to prevent unauthorized access from the unsubscribed users and the dedicated eavesdroppers. A heuristic approach is to combine these two services into one integral service over one transmission time block, which is defined as physical layer service integration (PHY-SI). In a PHY-SI system, these two coexisting services can share the same radio resources by exploiting the physical characteristics of wireless channels to significantly enhance the spectral efficiency. However, in general, the confidential and multicast (or public) services must be available to different user groups to satisfy their own demands. Thus, it is critical to guarantee reliable transmission for the confidential unicast service without sacrificing the quality of the multicast service.

Take a multi-antenna multicasting system as an example, the transmit beamforming/precoding is typically designed to ensure an efficient transmission of the common messages that all users can decode to maximize the sum-rate while maintaining the desired QoS level for all users. Due to the vulnerable nature of the wireless broadcast channel to eavesdropping, physical layer security techniques are becoming increasingly important. They achieve high secrecy performance without secret key distribution and management that may lead to security vulnerability in wireless channels. The key feature of physical layer security is that the channel for the legitimate user must be better than the eavesdropper's channel to guarantee a positive secrecy rate that is defined as the mutual information difference between the legitimate user's channel and the eavesdropper's channel to the transmitter. Recently, various secure transmission strategies against eavesdropping have been developed based on information-theoretical studies, where multi-antenna wiretap channels have been investigated to take advantage of the additional degrees of freedom and diversity gains. The existing techniques in multi-antenna secrecy channels aim to design the optimal transmit beamforming vectors, and to introduce more interference to degrade the eavesdroppers' link (i.e., artificial noise and cooperative jammer), thus improving the achievable secrecy rate in multi-antenna secrecy channels.

Unlike many works focused on the PHY-SI from the viewpoint of information theory, the work in 5G-Xcast [42], focusing on signal processing techniques, proposes a design of transmit covariance matrices to achieve the capacity region service information. More specifically, an artificial noise-aided RRM algorithm is designed to maximize the achievable secrecy rate to find secrecy-multicast trade-off performance. The proposed algorithm enables the specification of variant target quality of multicasting service and the maximization of the corresponding achievable secrecy rate, as well as provides the optimality of transmit beamforming via showing that the confidential optimal covariance matrix is of rank-one. More details on the performance analysis of the state-of-the-art and proposed PTM RRM with consideration of security will be given in Section 4.7.2.

4.7.2 Optimisation of RRM algorithms in the PTM scenarios with consideration of security

As mentioned in Section 4.7.1, most of the current RRM strategies only tackle the PHY-SI problem from information theoretic aspects, where the main goal is to derive capacity results or to analyse coding schemes that achieve certain rate regions. However, to pave the road for practical implementation, it is also important to investigate PHY-SI from signal processing aspects and identify the optimal transmit strategy for the transmitted integrated services to maximize the achievable secrecy rate regions.

Specifically, the fundamental limit on the achievable rate region in a PHY-SI system is investigated subject to the secrecy constraint. The optimal integration of both open-multicast and confidential-unicast services is investigated in a discrete memoryless broadcast channel, to bidirectional relay networks. The achievable secrecy rate region under channel uncertainty in a compound broadcast channel represents a robust PHY-SI transmit strategy. Compared to the current strategies such as the TDMA-based scheme and the power splitting scheme, a Secrecy-Multicast Rate Region Maximization (SMRRM) problem is formulated subject to the transmit power and the energy harvesting constraints by incorporating perfect or imperfect channel estimation. The goal is to jointly design the optimal input covariance matrices for the energy beamforming, the open-multicast service, the confidential-unicast service, and the artificial noise. The detailed system model and evaluation configurations can be found in [42], while the representative evaluation results are shown here.

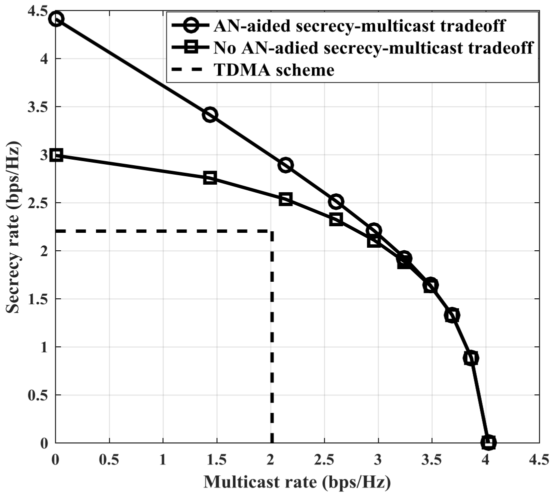

Figure 4.7-1: Secrecy-multicast performance tradeoff with transmit power 10dB

From Figure 4.7-1, it can be observed that the proposed artificial noise (AN)-aided scheme outperforms no-AN scheme. The striking gap indicates that AN indeed enhances the security performance without compromising the multicast rate. Nonetheless, with the increasing demand for multicast rate, the two curves tend to be coincident, which implies that AN is prohibitive at high multicast rate regions. The prohibition of AN reveals an inherent difference between PHY-SI and PHY-security: the use of AN must be more prudent due to the demand for multicast rate. The proposed scheme yields a significantly larger region than the TDMA-based one, which implies the inherent advantage of PHY-SI over traditional service integration. Also it can be seen that the performance gap between the power splitting suboptimal scheme and the real secrecy rate region is negligible. This observation demonstrates that the power splitting scheme can achieve a near-optimal performance with higher implementation efficiency.

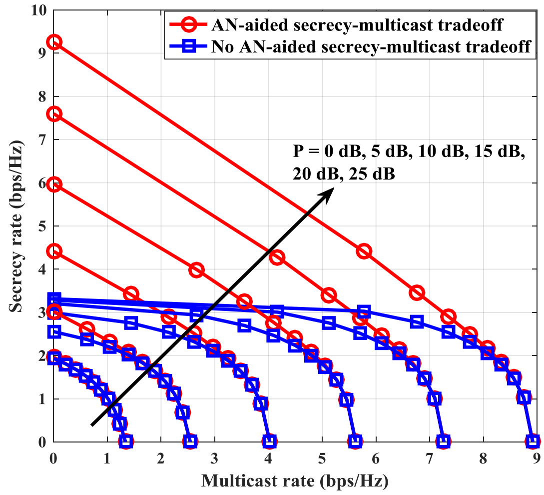

Figure 4.7-2: Secrecy-multicast performance tradeoff with different transmit powers

From Figure 4.7-2 it can be observed that the proposed AN-aided scheme achieves a secrecy rate region larger than the no-AN one, even under low transmit power. However, the gap between these two strategies is dramatically reduced with the transmitting power. This is due to AN's dual role in PHY-SI, i.e., in order to guarantee the multicast rate, AN must decrease to reduce the interference at all receivers. The second observation is that the secrecy rate regions with AN expand more strikingly when the transmitting power increases. On the contrary, the secrecy rate regions without AN practically expand in the horizontal direction, so that the increasing transmitting power mainly contributes to the multicast message transmission, rather than the confidential message transmission.

4.8 RRM for terrestrial broadcast

This section is described in D3.3 [45] Sections 5.2.5 and 6.1.5 as well as D3.2 [46] Section 3.4.

5 PTM performance evaluation

5.1 Evaluation methodology

5.1.1 Coverage simulation methodology

This evaluation focuses on a scenario where mobile operators transmit video streams to users in areas where one or more DTT broadcast services are not available or with signal strength below a determined QoS threshold. To study the effects of such potential demand on the cellular network and how 5G-XCast Mixed Mode solution can help in an efficient delivery of TV contents a real-life scenario is considered.

The observed area is ≈ 37 km x 37 km and includes an Italian city of 130,000 inhabitants surrounded by countryside (see Figure 5.1-1). Two DTT broadcast sites (shown in Figure 5.1-1) are present in the area, one in the city and one in the countryside near the urban area. The two sites transmit 5 different DTT services (local[6] and national).

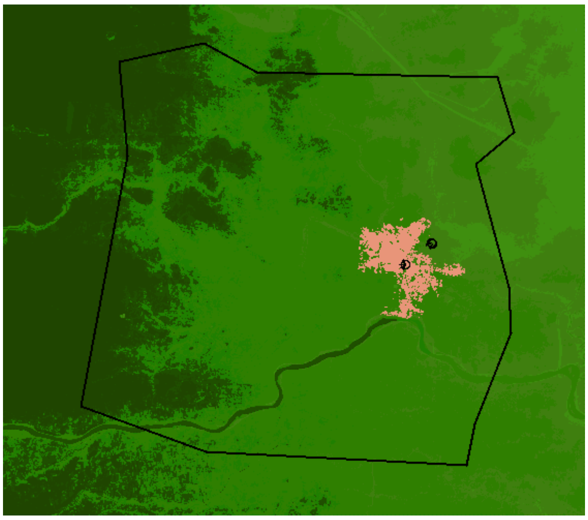

Figure 5.1-1 Area focus of the coverage simulation

The area is also well covered by 4G LTE-A mobile service. As an example of a real mobile network the 800 MHz LTE-A layer is considered as its cell distribution is supposed to be similar to the upcoming 5G 700 MHz layer. To take into account the border effect, mobile cell sites located in a ring around the considered area are included in the LTE-A layer.

A coverage simulation of the area is performed for the 5 DTT services and the 800MHz LTE-A layer with a proprietary planning tool. For the study, the considered area is analysed with a set of more than 550,000 representative points placed on a grid of 50 x 50 metres. The DTT signal simulated in each representing point is checked against the minimum field strength at fixed reception location of 45.0 dBµV/m [55]. In particular a DVB-T2 broadcast in Band IV/V is considered with 256-QAM modulation (with an indicative bit rate around 35/40 Mbps) and a minimum C/N of 19.7 dB. The points with at least one DTT service below the threshold are then considered for a 5G video stream substituting the DVB-T2 broadcast. In each pixel of the LTE-A layer the cell which is the best server is determined. The study analyses how many cells can be potentially involved in this transmission (with the area involved within each cell) as well

[6] Note that local services tend to limit their audience to the city area and its surroundings.

as the maximum number of potential users (TV sets) that could be reached by the 5G streaming. The latter data is obtained by considering the population density in the different area types (from ISTAT the Italian National Statistical Institute) and the number of TV-sets per population (from the web-site of DeAgostini Geografia, an Italian Geography Publisher). The cells with a medium to high number of potential users are then analysed for neighbour relations, under the assumption that a 700 MHz 5G NR layer would have a similar number and arrangements of cells to the considered 800 MHz LTE-A layer. Cell-neighbour relation is defined as follow: Cell A and Cell B are neighbour if they belong to the same site or if and only if Cell B can guarantee the service within at least 20% of Cell A best server area and vice versa.

PTM solutions activation can be suggested for a certain number of cells. To determine the exact number of those cells a more in-depth methodology is devised and an analysis is performed. To assume the number of active channel streams in cells with a medium to high number of potential users (TV sets) television audiences' data can be used. Italian TV audiences' data are available from Auditel (a totally independent and impartial company that measures television audiences in Italy on a national and regional level through the various broadcasting modes). Based on rigorous statistical methods, Auditel has set up a panel of families selected to represent the Italian population as a whole. As in many advanced countries, to measure television audiences, data is collected automatically by means of an electronic meter (people-meter) linked to each TV set in the sample home. The 2018 Italian TV audiences' data, publicly available (http://www.auditel.it/dati/), is used in the analysis.

5.1.2 System level simulation methodology

Nomor

Nomor's system-level simulator "RealNeS", that is used in WP3 to develop and analyse RAN methods and protocols and to perform the system-level simulations required in the context of the evaluation 3GPP's proposal for IMT-2020, can also be operated as a live demonstrator with a graphical user interface on top of it. It is primarily focused on the user plane covering the various protocol layers from a large set of data traffic generators over UDP or Transmission Control Protocol (TCP) and down to a detailed emulation of the physical layer. Accurate spatial channel models are used, where the latest model from [31] is added in the course of this project. It allows for simulation of both generic environments such as "dense urban", "rural" or "indoor" as defined in [29] and [31] as well as real-world scenarios, where actual geographical, building and mobility data can be imported for more illustrative demonstrations. RealNeS is actually not a single simulator for a particular RAT, but it covers various technologies, namely LTE-A, 802.11 and NR, the latter obviously being under heavy development as 3GPP is in the process of standardizing NR. It also features a multi-RAN framework that facilitates the simultaneous simulation of several networks of various RATs each operating in a unique and mutually exclusive frequency band. This facilitates various studies on e.g. traffic steering across networks, mobility and network convergence.

This simulator is used for evaluation of link adaptation with PTM, 2^{nd} layer of FEC in RAN and for IMT-2020 system level evaluation of NR.

Samsung

Samsung R&D Institute UK has developed an in-house abstract system level simulator called 5G-PySim, written in Python. The simulator simplifies higher layers implementation and focuses on L1 and L2. It uses link level simulation results as part of configuration that gives basic characteristics of 4G LTE-A or 5G NR. The system level simulation results can be represented in IPython with any available tool, such as Jupyter. It allows carrying out the simulations intended to investigate the performance of 4G LTE-A and 5G NR with comparable settings.

In the aspect of simulating RAN deployment, as 5G-PySim can support, a typical hexagonal deployment of base stations (BSs) is conceived involving an adequate number of sites and UEs where service traffics are broadcasted from BSs to UEs. The layout can be a few rings surrounding a designated BS, which would involve 19 sites in the case of 2 concentric rings (1 site in the centre + 6 immediate neighbouring sites as the first ring + 12 sites as a further second ring). MCS (Modulation and Coding Scheme), Transport Block Size (TBS) and other parameters can be flexibly set at the BS. In the aspect of channel models, the latest 3GPP technical specifications and technical reports are used. CQI (Channel Quality Indicator) values can be manually set for all UEs, which allow it to be observed how the RAN would perform in all the different channel conditions. The actual simulation settings are given in sub-section 5.2.2.

This simulator is used for throughput and block error rate evaluation as a function of CQI, comparing 5G-MM against 4G SC-PTM.

5.2 Evaluation results

5.2.1 Coverage simulation results

The first result of the realistic coverage simulation is that the points in the grid with at least one DTT service field strength below the considered threshold of 45.0 dBμV/m, are 35.9% of the total (corresponding to an area of 502 km^2). After superimposing to those points the cellular layer, it can be observed that the cells containing points below threshold (where a cellular video streaming could be activated) amount to 69.5% of the total number of cells in the considered area (123 cells involved). Figure 5.2-1 depicts the distribution of the cells in relation to the percentage of their respective best server areas involved in possible transmissions of video stream.

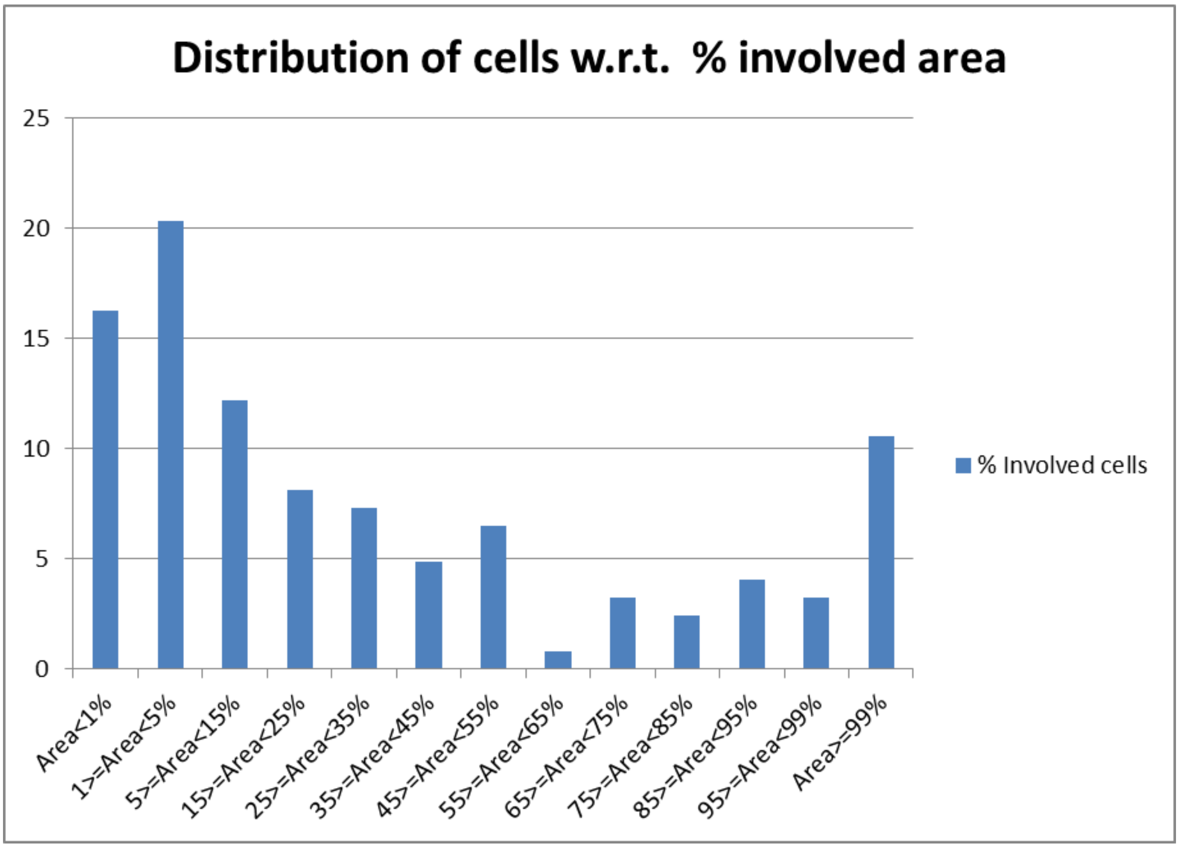

Figure 5.2-1 Distribution of cells in relation to the % of their area involved in DTT substitution

It can be noted that more than 36% of the involved cells are only marginally engaged since only 5% or less of their best server areas (two most-left bars) contains points below DTT threshold. On the other hand, more than 14% of the involved cells are almost totally engaged in DTT substitution as the percentage of their involved best server area is above 95% (two most-right bars). The remaining 50% of the engaged cells have different percentages of their involved areas with a majority of them in the range between 5% and 35% of their areas.

The above merely geographical considerations can be extended with an analysis on the number of TV sets potentially reachable by the mobile service within each cell. The maximum number of TV sets potentially interested in a mobile service substituting DTT can be derived from the population density in the different area types (to obtain the population in each cell) and the statistical number of TV-sets per population. The distribution of cells in relation to the number of reachable TV sets for video streams is depicted in Figure 5.2-2

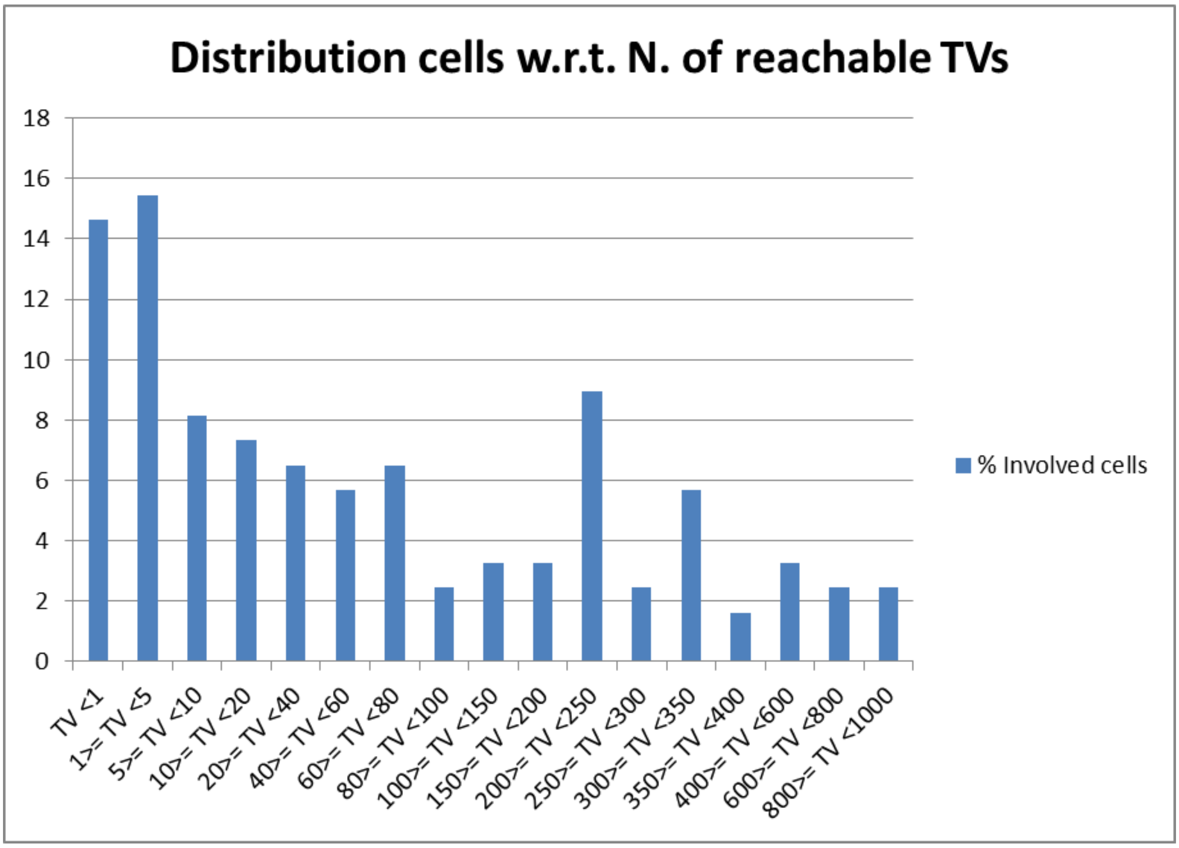

Figure 5.2-2 Distribution of cells in relation to the number of reachable TV sets in their area concerned with possible receptions of video streams in DTT substitution

As expected from the geographical analysis, 38% of the cells presents a very low (<5) or low (<10) number of reachable TV sets. For those cells, a sporadic PTP transmission would probably satisfy the very low demand. Moving towards a greater number of TVs per cell, 27% of the cells could reach 200 and more TV sets each, leveraging on PTM transmissions. Furthermore, 8% of the cells present a number of reachable TV sets above 400.

To understand if 5G-XCast PTM solutions (MC-MM and terrestrial broadcast) could benefit from the use of SFN resource allocation, a neighbour analysis has been also performed. The number of cell-neighbour relations found for the 123 cells according to the criteria defined in Section 5.1.1 is 545. Cells are mostly neighbouring each other as the following graphs considering cells with different number of potential TV sets show (see Figure 5.2-3). Cells name are anonymized but AAn cells belong to the same site AA.

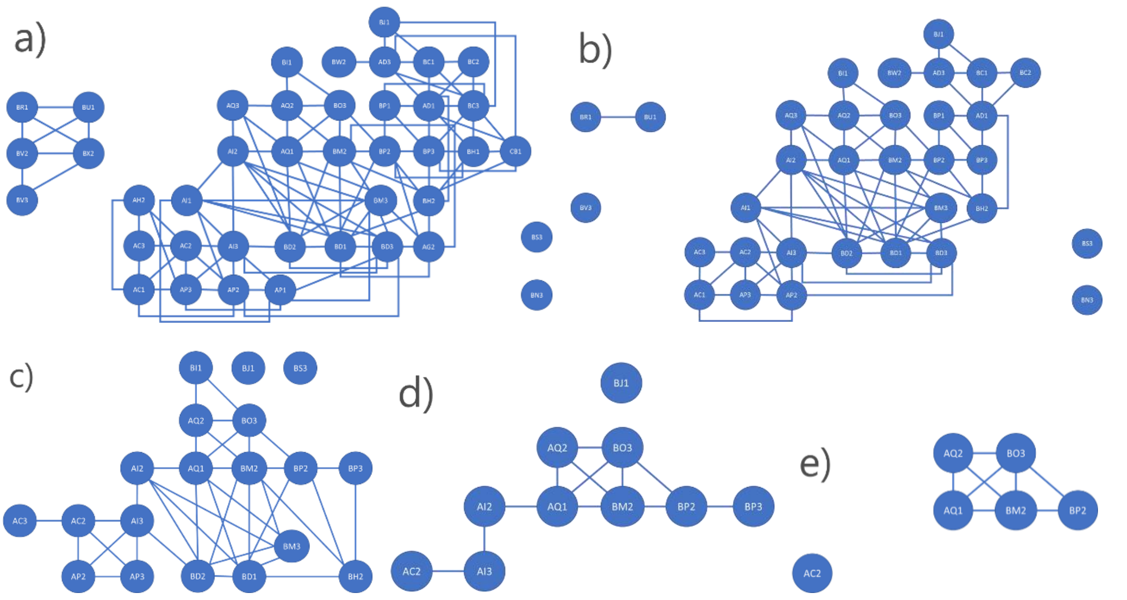

Figure 5.2-3 Neighbour relation graphs within cells with number of potential TV sets greater than a) 100 (41 cells), b) 200 (33 cells), c) 300 (19 cells), d) 400 (10 cells) and e) 500 (6 cells)

From the graphs it can be seen that the cells with a medium to high number of potential TV sets are typically interconnected (i.e. they tend to have several cell-neighbours). In this scenario the usage of a MC-MM or T-Broad solution with a SFN allocation strengthens the broadcast signal and reduces interferences between cells obtaining a higher resource efficiency of the network. The use of SFN MC-MM or terrestrial broadcast can be adopted for a certain number of cells: from the ones where the service will be always on (with a high number of potential users) to those where the PTM is dynamically activated when needed (with a medium number of potential users). The methodology devised to determine the exact number of cells where PTM solutions are preferable has been applied to the case study and Auditel data of the different day periods in the observation interval of 19 hours per day (between 7 and 2 AM) for 12 months of 2018 has been examined. Accordingly, the time period between 7 and 9 AM in the month of August 2018 is the one with the minimum number of TV sets on, only 7.15 % of TVs, while the time period between 8:30 and 10:30 PM in the month of February 2018 is the one with the maximum number of TV sets on, 46.26% of TVs. The yearly average of 2018 amounts to 20.82% of TVs on. For each considered minimum and maximum TVs on period and the 2018 average, the relative audience figures of the seven major national Italian TV channels has been collected and, within each set, the TV channels with minimum and maximum audience and the average audience of the channels in the set has been singled out (see Figure 5.2-4).

7.15% TVs On

August 2018	7:00-9:00
Rai1	**19.9%**
Rai2	4.2%
Rai3	5.52%
Rete4	**1.02%**
Can5	15.69%
Ita1	1.36%
La7	4.07%
Average	*7.39%*

46.26% TVs On

February 2018	20:30-22:30
Rai1	**25.08%**
Rai2	4.76%
Rai3	5.18%
Rete4	**3.64%**
Can5	15.53%
Ita1	5.2%
La7	4.39%
Average	*9.11%*

20.82% TVs On

Average 2018	All periods
Rai1	**16.81%**
Rai2	5.64%
Rai3	6.87%
Rete4	**3.57%**
Can5	15.35%
Ita1	4.52%
La7	3.7%
Average	**8.07%**

Figure 5.2-4 Shares of the seven major national Italian channels in the TVs on periods analysed with indications of the minimum, maximum and average shares.

The number of viewers of a selected channel in each cell can be determined by the following formula:

N. of potential TVs x percentage of TVs on x the selected channel share (percentage).

The maximum, minimum and average number of viewers is determined by the channel most watched, the channel least watched and average share values of the corresponding case.

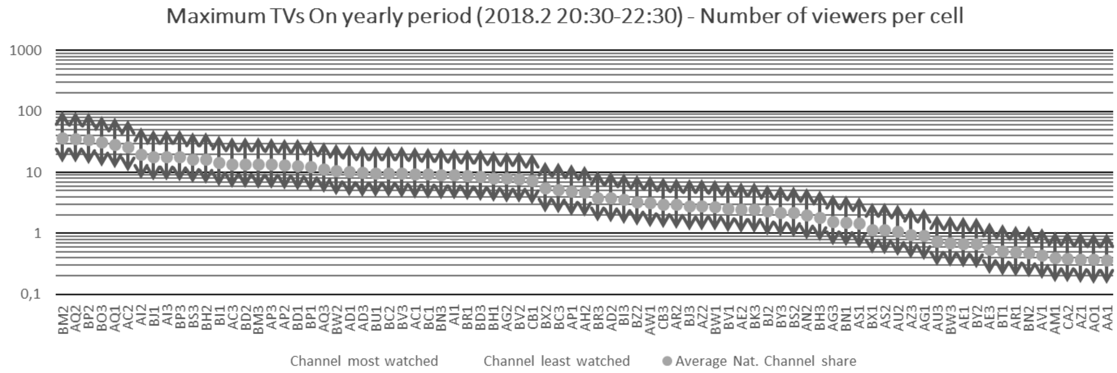

Figure 5.2-5 Sorted list of the cells with one or more viewers of the most watched channel (Rai 1) in the period of the maximum number of TV sets On

In Figure 5.2-5 a sorted list of the 81 cells with one or more viewers of the most watched channel (Rai 1) in the period of the maximum number of TV sets on is presented. For each cell listed three values are shown: the number of viewers of the most watched channel, the number of viewers of the least watched channel, and the number of viewers considering the average national channel share. It can be noted that the number of viewers for the least watched channel it’s below 1 only for cells with fewer TV sets while tops 100 for the most watched. In this period PTM activations should apply to a lot of cells (e.g. top 50, 60 cells or all 81 cells listed if considering the streaming of the most watched channel). The remaining 42 cells (34% of the total) have less than one viewer of the most watched channel per cell.

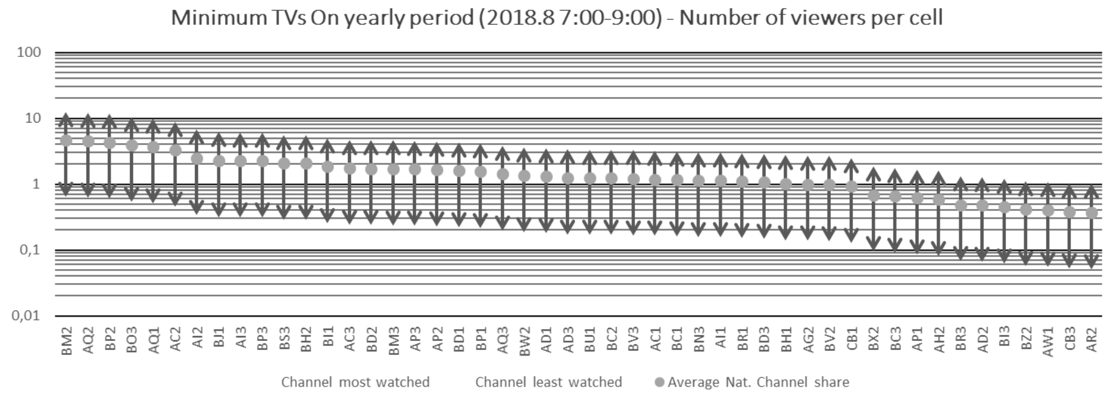

Figure 5.2-6 Sorted list of the cells with one or more viewers of the most watched channel in the period of the minimum number of TV sets On

In Figure 5.2-6 a sorted list of the 48 cells with one or more viewers of the most watched channel in the period of the minimum number of TV sets on is presented. For each cell listed three values are shown: the number of viewers of the most watched channel, the number of viewers of the least watched channel, and the number of viewers considering the average national channel share. It can be noted that the number of viewers for the least watched channel it's below 1 in every cell while tops 12 for the most watched. In this period PTM activations could be limited to few cells with higher numbers of TV sets (e.g. top 10, 20 cells or all 48 cells listed if considering the streaming of the most watched channel). The remaining 75 cells (61% of the total) have less than one viewer of the most watched channel per cell.

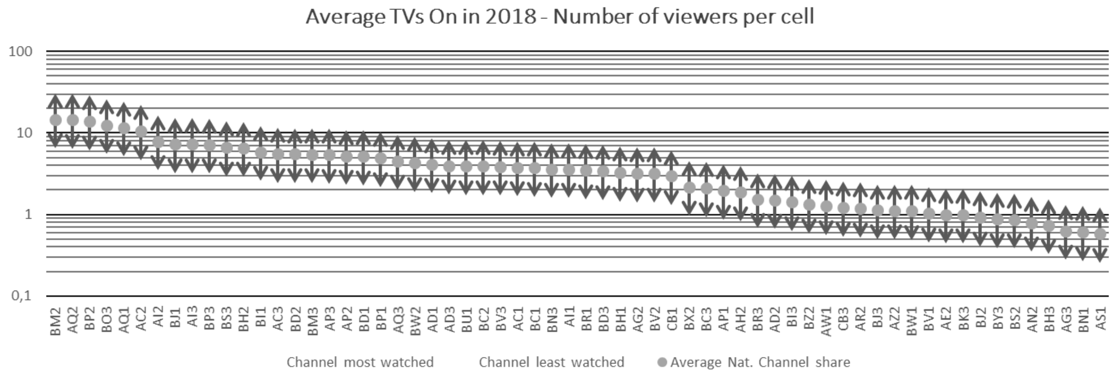

Figure 5.2-7 Sorted list of the cells with one or more viewers of the most watched channel considering the average TV sets On and audience shares (of 2018)

In Figure 5.2-7 a sorted list of the 62 cells with one or more viewers of the most watched channel considering the average TV sets on and audience shares (of 2018) is presented. For each cell listed three values are shown: the number of viewers of the most watched channel, the number of viewers of the least watched channel, and the number of viewers considering the average national channel share. It can be noted that the number of viewers for the least watched channel it's below 1 only for cells with fewer TV sets while tops 30 for the most watched. Considering the 2018 average audience data, PTM activations should apply to a significant number of cells (e.g. top 41 with average above 2 viewers or all 62 listed cells if considering the streaming of the most watched channel). The remaining 61 cells (50% of the total) have less than one viewer of the most watched channel per cell.

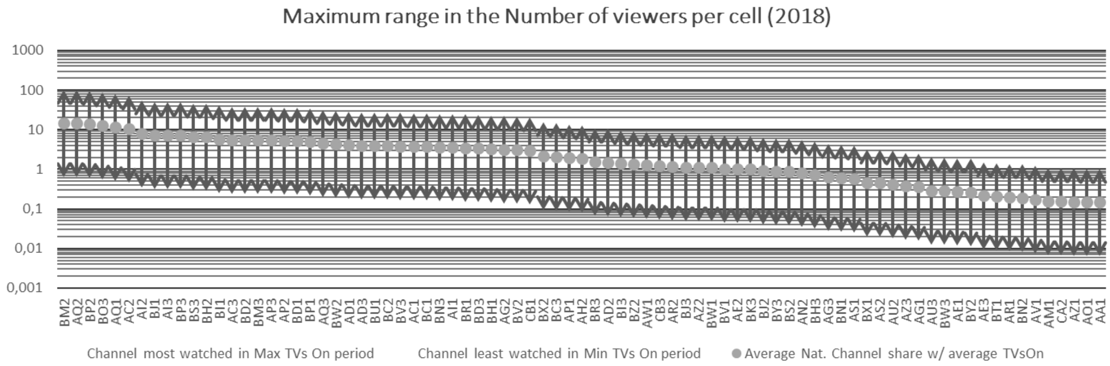

Figure 5.2-8 Sorted list of the cells with one or more viewers of the most watched channel in the period of the maximum number of TV sets On depicting the maximum range in cell viewers according to 2018 audience data (the lower limit is the least watched channel in the period of minimum number of TV sets On and the average is the nat. channel average share with average TVs On)

Finally, in Figure 5.2-8 sorted list of the 81 cells with one or more viewers of the most watched channel in the period of the maximum number of TV sets on is presented depicting the maximum range in cell viewers according to 2018 audience data. For each cell listed the upper limit is the number of viewers of the most watched channel in the period of the maximum number of TV sets on and the lower limit is the number of viewers of the least watched channel in the period of the minimum number of TV sets on while the average is the national channel average share of 2018 with the average TVs on. The remaining 42 cells (34% of the total) have less than one viewer of the most watched channel per cell. This representation can be seen as a recap of the previous analysis and any considerations regarding PTM activation policies apply as well. Alternatively, new considerations can be drawn (e.g. top 41 cells with average above 2 viewers always active plus other 10, 20 cells activated dynamically in prime-time only).

5.2.2 System level simulation

Link adaptation for PTM simulations

Simulation settings

System-level simulations of communication networks mimic geographically confined parts of a communication network consisting of multiple base stations and numerous UEs, gateways, application servers etc., i.e., including layer-2/3 and possibly higher layer protocol functionalities. This allows for evaluation of aspects such as RRM, interference between different concurrent transmissions or higher-layer consideration, such as the impact of radio network performance on TCP connections or user experience at the application level. This may take into account UE distributions or mobility according to synthetic models or in ”real-world" scenarios.

Of the various test environments defined for IMT-2020 evaluations [29], urban dense test environment is used for performance evaluation. Table 5.2-1 summarizes the main simulation parameters, which are derived from the “Dense Urban” scenario.

Parameters	***Value***
Total BS transmit power	51 dBm
System bandwidth	100 MHz
Carrier frequency	4 GHz

gNB antenna configuration	[M, N, P] = [8, 4, 2][7]
Transceiver Units (TXRUs) at gNB	8 ([Mp, Np, P] = [1, 4, 2])[8]
Inter-site distance	200 m
Number of UE antennas	8
TXRUs at UE	8 (1-to-1 mapping)
UE mobility model	3 kmph, randomly uniform distribution
BS noise figure	5 dB
UE noise figure	7 dB
BS antenna gain	14 dBi
BS antenna elevation 3dB beamwidth	10°
BS antenna azimuth 3dB beamwidth	65°
UE antenna element gain	0 dBi
PTP traffic model	Full buffer
PTM traffic model	8 Mbps, packet arrival rate 100 Hz
Channel model	3GPP TR 38.901 [31] (= IMT-2020 model B)

Table 5.2-1: System-level simulation parameter settings

Simulation results

Link adaptation evaluation for PTM transmission includes investigation of MCS modifications (with heuristic fixed offsets as well as adaptive MCS via CQI report from a UE that has the worst radio link), Single User Multiple Input Multiple Output (SU-MIMO) precoder and rank. The precoder and rank settings are realized via configuration of Precoding Matrix Indicator (PMI) and Rank Indicator (RI). The fixed PMI setting refers to use of fixed precoder to Physical Resource Block (PRB) association throughout the simulation. On the other hand, the cyclic PMI refers to adaptive use of a precoder via cyclic access to the PMI codebook, leading to diversity benefits. Such evaluation of multi-antenna PTM schemes is interesting since NR is expected to have large antenna arrays; for example, ITU-Radiocommunication sector (ITU-R) even allows up to 256 antenna elements per transceivers (TxRP) in dense urban scenarios [29].

The evaluation targets achieving optimal trade-off between coverage and spectral efficiency. Herein, coverage refers to the percentage of UEs for which the probability that Packet Loss Rate (PLR) greater than the minimum allowed loss rate θ is lower than a certain Quality of Experience (QoE) threshold ϵ, configured by the network operator or planner. In other words, the coverage refers to percentage of UEs for which

$$\Pr(PLR_{1sec} > \theta) < \epsilon .$$

The QoE threshold ϵ is assumed to be 1%. The minimum allowed targeted PLR θ is a design parameter in combination with higher layer FEC schemes such as AL-FEC and 2nd layer of FEC in RAN. For example, for the 2nd layer of FEC in RAN, which uses RLNC, the PLRs are measured on RLNC PDUs over 1 second interval, which is the higher layer FEC interval. The targeted coverage is 95 % or above.

Figure 5.2-9 demonstrates the percentage of UEs that violate the QoE threshold $\epsilon = 1\%$, i.e, $\Pr(PLR_{1sec} > \theta) > 1\%$. The MIMO rank is configured to 1. Various targeted PLR θ are analysed for fixed and cyclic PMI settings. For the same performance in spectral efficiency (1.8bps/Hz) and at lower target RLNC PLR (e.g., $\theta = 1, 10$), the cyclic PMI has higher percentage of UEs that violate the QoE threshold. The reason is that cyclic PMI adaptively changes beams and affect PLR of most of the UEs at a lower targeted PLR θ. On the other hand, for higher target RLNC PLR (e.g., $\theta = 20, 30$), the

[7] M, N and P refer to the number of vertical, horizontal and polarization arrangement of antenna elements, respectively.

[8] The 8 vertical antenna elements for each polarization are hard-wired and are fed to a TXRU.

cyclic PMI has lower violation of QoE criteria as compared to fixed PMI setting due to diversity benefits. Figure 5.2-10 shows the CDF of the overall PLR in the system. Herein, it is shown that the cyclic PMI has overall lower probability of high PLR as compared to fixed PMI. Comparing the overall QoE threshold at various PLR per second, cyclic PMI provides >95% coverage as compared to fixed PMI.

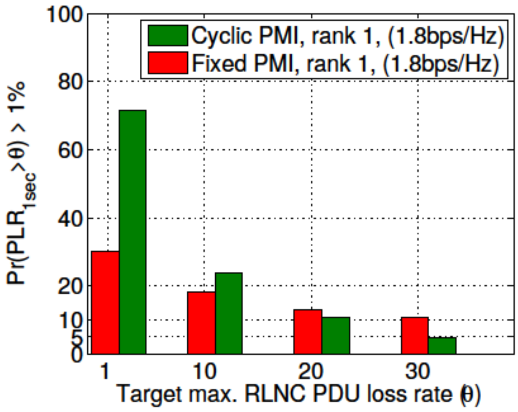

Figure 5.2-9: The percentage of UEs that violate the QoE threshold $\epsilon = 1\%$ as a function of targeted PLR θ for fixed and cyclic PMI.

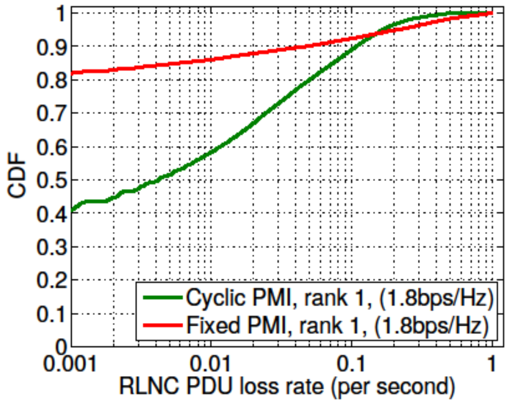

Figure 5.2-10: The overall PLR in the system for fixed and cyclic PMI.

Figure 5.2-11 adds with respect to previous comparison the consideration of MIMO rank 1 and 2. At lower targeted PLR θ, rank 2 has lower percentage of UEs that violate the QoE threshold as compared to rank 1 for the same PMI setting. The main reason is the improved diversity benefits from rank 2. At higher targeted PLR θ, the diversity benefits from rank 2 saturates since most of the diversity benefits are already exploited by using cyclic PMI. Accordingly, cyclic PMI with both rank settings 1 and 2 provide coverage >95 %. Figure 5.2-12 shows the overall PLR in the system for fixed and cyclic PMI with consideration of rank 1 and 2. Similarly, the diversity benefits of rank 2 are exhibited at lower packet loss rates for the same PMI settings; on the other hand, the diversity benefits saturate at higher packet loss rates.

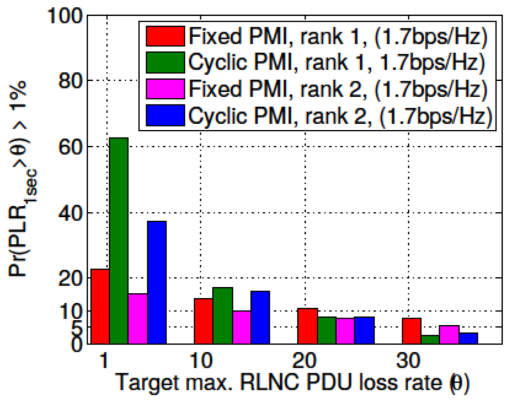

Figure 5.2-11: The percentage of UEs that violate the QoE threshold $\epsilon = 1\%$ as a function of targeted PLR θ for fixed and cyclic PMI with consideration of rank 1 and 2.

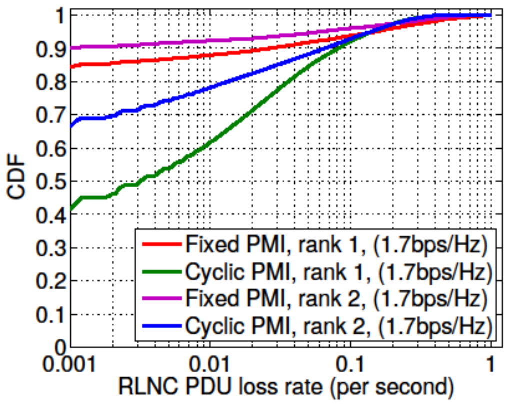

Figure 5.2-12: The overall PLR in the system for fixed and cyclic PMI with consideration of rank 1 and 2.

Figure 5.2-13 shows the percentage of UEs that violate the QoE threshold $\epsilon = 1\%$,, i.e., $\Pr(PLR_{1sec} > \theta) > 1\%$ as a function of targeted PLR θ for fixed MCS with fixed and cyclic PMI as well as adaptive MCS that takes into account cyclic and worse-UE PMI. A heuristic link adaptation offset of 13 dB is used with the adaptive MCS to have a similar spectral efficiency (~1.8 bps/Hz) across all the considered settings to have fair comparison. The adaptive MCS uses inner loop link adaptation based on worse-UE's feedback on the PTM channel as well as an outer loop link adaptation based on PTP-based ACK / NACK feedbacks from the worse-UE. The major observation is that the worst-UE PMI does not avoid very high packet loss rates since a setting based on the worse-UE makes other UEs suffer from packet loss rates, i.e., it concentrates packet loss on UEs creating new worse UE. On the other hand, the cyclic PMI randomizes the PMI settings with UE channel conditions. Moreover, the adaptive MCS modification settings do not provide the coverage requirement of >95% for all considered PMI settings. Hence, adaptive MCS which occurs at a faster time scale than higher layer FEC schemes does not give a meaningful benefit in regards to providing the expected percentage of UEs that fulfil QoE requirement of 1%. Figure 5.2-14 shows the overall PLR in the system for fixed MCS with fixed and cyclic PMI as well as adaptive MCS that takes into account cyclic and worse-UE PMI at heuristic link adaptation offset of 13 dB. For adaptive MCS, the worse-UE PMI settings outperform cyclic PMI settings in particular at both lower and higher packet loss rates.

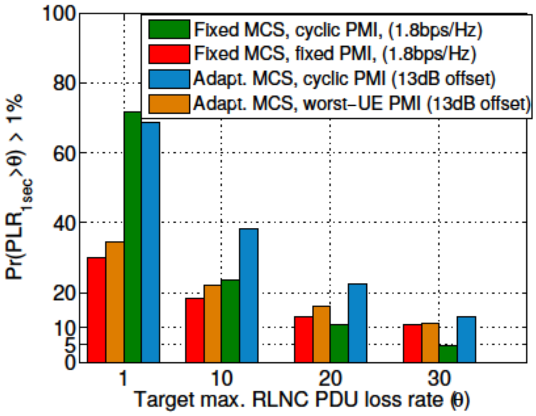

Figure 5.2-13: The percentage of UEs that violate the QoE threshold $\epsilon = 1\%$ as a function of targeted PLR θ for fixed and cyclic PMI as well as adaptive MCS that takes into account cyclic and worse-UE PMI at heuristic link adaptation offset of 13 dB.

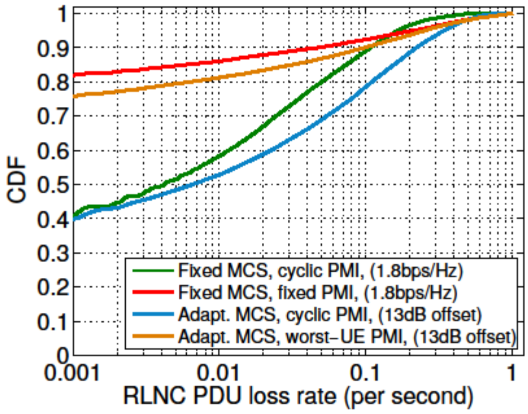

Figure 5.2-14: The overall PLR in the system for fixed and cyclic PMI as well as adaptive MCS that takes into account cyclic and worse-UE PMI at heuristic link adaptation offset of 13 dB.

2nd layer of FEC in RAN simulations

Simulation settings

The system level simulation parameters for the network deployment is the same as the setting in Table 5.2-1. 10 UEs per cell are dropped random uniformly. Multiple drops are considered to collect stable statistics for performance evaluation. The 2nd layer FEC mechanisms all operate in the Galois Fields GF(256) and with a generation size of 100 symbols, i.e., over 1 s.

Simulation results

By using elaborated system-level simulations, the newly proposed feedback-based 2nd layer FEC scheme is compared against two reference schemes:

- **No AL-FEC**: Operation without any kind of AL-FEC.
- **AL-FEC**: Operation with LTE-like AL-FEC, i.e., a systematic fountain code. Herein, a systematic RLNC code with optimal decoding is used, e.g. Gauss-Jordan elimination based. A comparison with actual Raptor codes as standardized for deployment in LTE-A is shown below.

Figure 5.2-15 compares the LTE-A raptor codes against the systematic RLNC code with optimal decoding used in this evaluation as feedback-less AL-FEC reference scheme. The comparison is done with respect to probability of failure ($p_{fail,i}$) as a function of the reception overhead (i), i.e., the number of PDUs or packets received in excess of the theoretical absolute minimum required for decoding a generation. Like in [30], a 40% packet loss rate is assumed in the simulation.

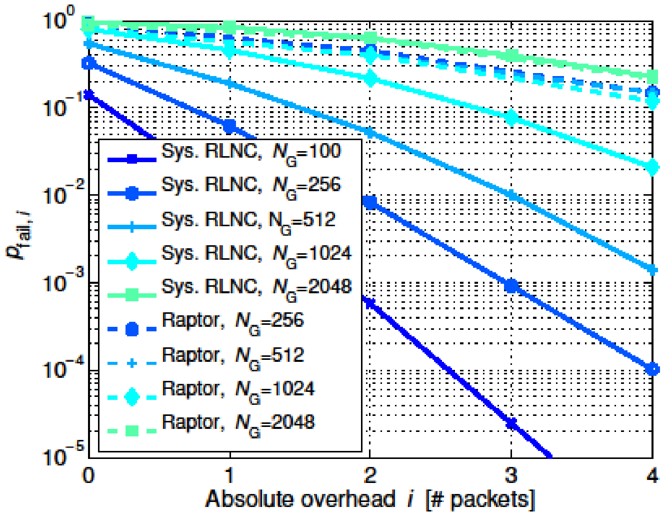

Figure 5.2-15: Comparison of systematic RLNC code and Raptor code with 40% packet loss rate.

As it can be observed, the optimally decoded systematic RLNC code clearly exhibits better performance in terms of decoding failure rate vs. reception overhead than Raptor code. However, a true Raptor code could in principle be used just as well as basis of the feedback-based 2nd layer FEC scheme. The difference in terms of average overall spectral efficiency as considered below is very small. For the generation size of 100 symbols considered in this work, the average reception overheads for the Raptor code and optimally decoded systematic RLNC code computed via (1) are approximately lower or equal than ~2% and 0.1%, respectively.

$$-1 + \sum_{i=0}^{\infty} \frac{N_G + i}{N_G} (1 - p_{fail,i}) \cdot \prod_{j=0}^{i=1} p_{fail,j} \qquad (1)$$

Figure 5.2-16, Figure 5.2-17 and Figure 5.2-18 show the CDF of application layer spectral efficiency, application layer packet loss rate and RLC SDU loss rate, respectively, for AL-FEC with 10% redundancy at application layer and different MCSs at layer-1 FEC as well as no AL-FEC for one sample MCS. As it can be seen in Figure 5.2-16, the CDFs of the spectral efficiency are simple step functions due to the fixed modulation and coding parameters. With more aggressive MCS settings, i.e., higher spectral efficiency of the layer-1 transmission scheme, both the application layer packet loss rate and the RLC SDU loss rate would increase.

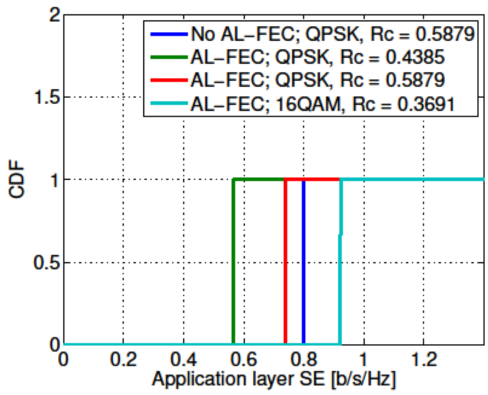

Figure 5.2-16: CDF of application layer SE [b/s/Hz] for no AL-FEC and AL-FEC with various MCS settings.

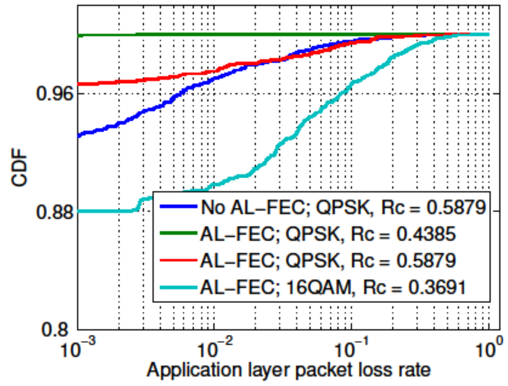

Figure 5.2-17: CDF of application layer packet loss rate for no AL-FEC and AL-FEC with various MCS settings.

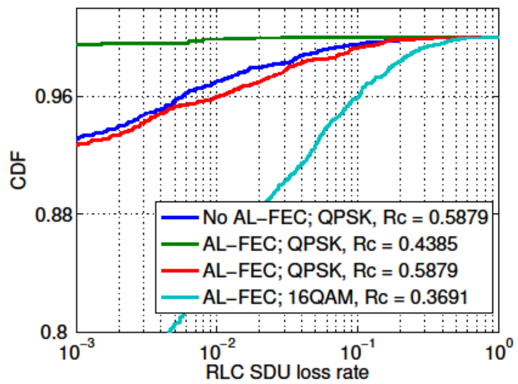

Figure 5.2-18: CDF of RLC SDU loss rate for no AL-FEC and AL-FEC with various MCS settings.

Comparing AL-FEC with no AL-FEC for the same MCS, one can observe that at the application layer (Figure 5.2-17), AL-FEC provides improved (lower) packet loss rate, as it is able to repair smaller packet loss events. On the other hand, as expected there is no impact at the RLC layer in terms of the RLC SDU loss rate (see blue versus red curves in Figure 5.2-18).

Figure 5.2-19 and Figure 5.2-20 show the CDF of application layer spectral efficiencies and packet loss rates, respectively, for no AL-FEC and AL-FEC with the same MCS setting of ‘QPSK, Rc = 0.59’ for various levels of redundancy of repair packet: 10%, 20% and 30%. Herein, AL-FEC considerably improves (reduces) the application layer packet loss rate at a cost of reduced spectral efficiency. It can also be observed how higher levels of redundancy are able to fix higher packet loss rates in the lower layers.

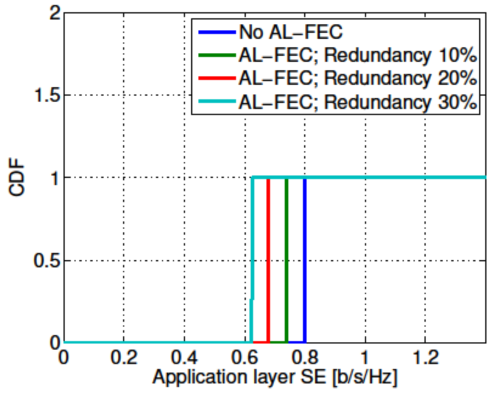

Figure 5.2-19: CDF of application layer SE [b/s/Hz] for no AL-FEC and AL-FEC with various redundancy levels and 'QPSK, Rc = 0.59'.

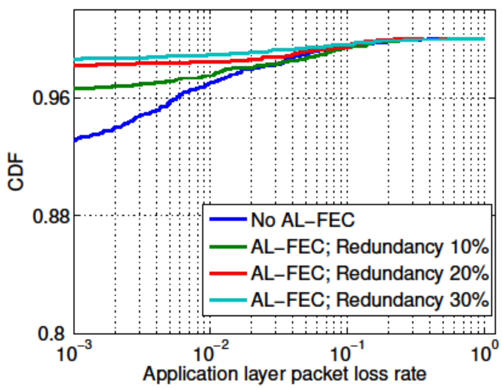

Figure 5.2-20: CDF of application layer packet loss rate for no AL-FEC and AL-FEC with various levels of redundancy 'QPSK, Rc = 0.59'.

Figure 5.2-21 and Figure 5.2-22 show CDF comparison of 2nd layer of FEC in RAN against AL-FEC and no AL-FEC, in terms of application layer spectral efficiencies and packet loss rates, respectively, for a sample MCS setting of 'QPSK, Rc = 0.59'. The 2nd layer of FEC in RAN utilizes periodic feedback (50ms) for triggering transmission of appropriate numbers of additional RLNC PDUs to compensate for lost packets. Consequently, the 2nd layer of FEC in RAN exhibits further improved packet loss rate performance as compared to conventional AL-FEC. At the same time, the spectral efficiency for 2nd layer of FEC in RAN is higher than that of AL-FEC because in the 2nd layer of FEC additional RLNC PDUs are not sent pre-emptively but are generated and sent only based on request. Accordingly, with the current configuration, in approximately 60% of all drops no additional RLNC PDUs are required for decoding, and in less than 10% of all drops, the overall spectral efficiency is lower than that of conventional AL-FEC, but with the benefit of having zero (at least < 10−3) packet loss rate.

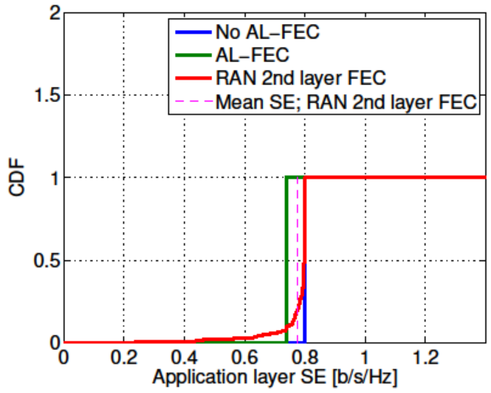

Figure 5.2-21: CDF of application layer SE [b/s/Hz] for no AL-FEC, AL-FEC and 2nd level of FEC in RAN with 'QPSK, Rc = 0.59'.

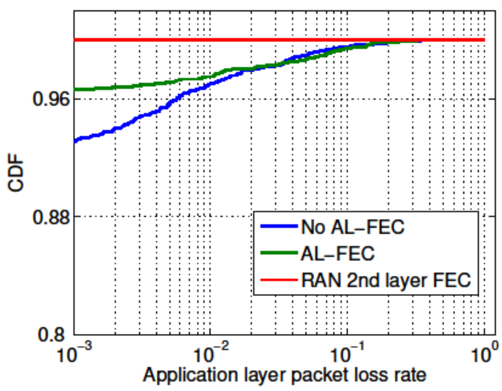

Figure 5.2-22: CDF of application layer packet loss rate for no AL-FEC, AL-FEC and 2nd level of FEC in RAN with 'QPSK, Rc = 0.59'.

Figure 5.2-23, Figure 5.2-24 and Figure 5.2-25 show the delay analysis of 2nd layer of FEC in RAN as compared to AL-FEC. Figure 5.2-23 contains the inverse-CDF (I-CDF) of the difference between the reception time of an application layer packet and its transmission time for AL-FEC with various MCS settings and for 2nd layer of FEC. For AL-FEC, the application layer delay is higher for less conservative MCS settings since the packet loss rate is higher and some application layer packet are recovered after reception of repaired packets, which delay others by the triggered reordering process. Note that for the conventional AL-FEC, no delays beyond 1.1s occur, as the reordering in the receiver is implemented to assume that 1.1s after the reception of the first PDU of a generation no more PDUs from that generation will be received[9]. Comparing 2nd layer of FEC to AL-FEC for the same MCS setting of ‘QPSK, Rc = 0.59’ (i.e. light blue curve versus green curve respectively), the application layer delay distribution is considerably more favourable for 2nd layer of FEC due to the fact that repaired packets are re-transmitted on the fly based on the periodic opportunity for feedback. Such application layer delay can be crucial in determining the quality of experience in watching a video, where a play-out buffer is installed to avoid frequent stalling of playback while in normal operation incurring some buffering delay.

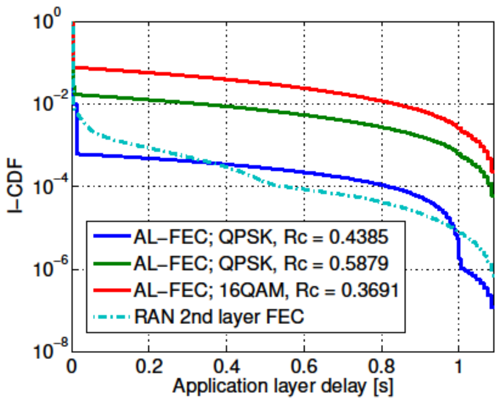

Figure 5.2-23: I-CDF of application layer packet delay for AL-FEC with various MCS settings and for 2nd layer of FEC with ‘QPSK, Rc = 0.59’.

In order to achieve a high quality of experience the target is to minimize the buffering delay while keeping the frequency of stalling events and the total relative time of stalls, i.e., the aggregated stalling time normalized by the observation window length, low. Figure 5.2-24 and Figure 5.2-25 show the CDFs of packet stalling frequency and relative packet stalling period, respectively, assuming a play-out buffer size / stalling threshold of 1.1s. This value is slightly larger than what is covered by one generation of NC SDUs to allow the repaired packets of the systematic AL-FEC code sent at the end of the generation to repair also losses on all packets of the generation. In this case, the 2nd layer of FEC provides a better performance in terms of packet stalling frequency and packet stalling period as compared to AL-FEC. Furthermore, it can be observed that it exhibits delay characteristics very similar to those of conventional AL-FEC with ‘QPSK, Rc = 0.44’, while the overall spectral efficiency is about 30% higher compared to this reference scheme.

[9] Recall that each generation contains SDUs generated over an interval of 1sec.

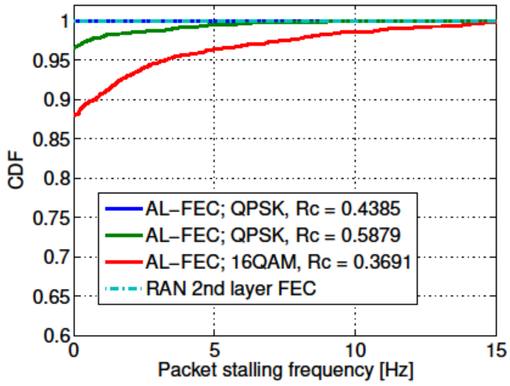

Figure 5.2-24: CDF of application layer packet stalling frequency for AL-FEC with various MCS settings and for 2nd layer of FEC with 'QPSK, Rc = 0.59'.

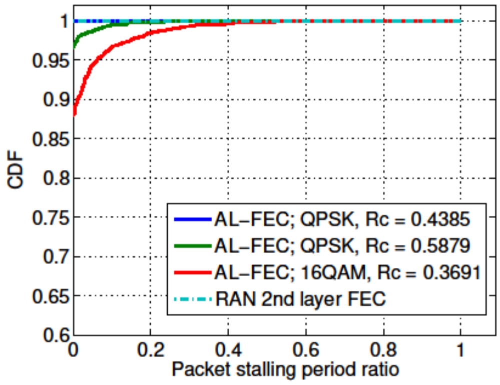

Figure 5.2-25: CDF of application layer packet stalling period ratio for AL-FEC with various MCS settings and for 2nd layer of FEC with 'QPSK, Rc = 0.59'.

The above key findings are included in [53].

Cross-layer link-adaptation in co-ordination with higher layer EC

Simulation settings

The system level simulation parameters for the network deployment is the same as the setting in Table 5.2-1. The PLRs are measured on EC PDUs over 1 second interval, which is the higher layer EC interval. The sample measurement trigger used is periodic measurement reporting defined as 'multiple user report timer', which can be configured to tens of ms. In particular, herein, it is configured to 100 ms to ensure reception of all potential measurement reports from multiple UEs.

Simulation results

Figure 5.2-26 shows the CDF of application layer SE of 'no EC' (Operation without any kind of EC scheme), 'AL-FEC' (with no Link Adaptation (LA)) and 'AL-FEC + LA' (where both AL-FEC and LA are switched on). Practically, network operator can flexibly configure 'threshold_higher', denoted by TH^H, and 'threshold_lower', denoted by TH^L. To ease demonstration of the proposed scheme, simulations are performed for various sample values of a threshold where TH^H and TH^L are assigned to same value. The MCS decrement offset 'mcs_delta_offset_decrement' and increment offset 'mcs_delta_offset_increment' are configured to 1 and 0.1, respectively. The mean SE values corresponding to 'AL-FEC + LA' are shown by dashed lines with the same colour as the corresponding CDF plots. On the other hand, the corresponding performance on the application layer packet loss rate and EC PDU loss rate are shown in Figure 5.2-27 and Figure 5.2-28, respectively. In cases of no LA, a fixed sample MCS setting with QPSK and coding rate = 0.59 is used. The AL-FEC is assumed to use 20% packet redundancy for the repair packets. The major observations are

- 'AL-FEC' provides better robustness sacrificing spectral efficiency as compared to 'no EC'. For more than 95% of the cases, the application layer packet loss rate is lower than 0.1 with AL-FEC, but 20% spectral efficiency is sacrificed for repairing packets.
- 'AL-FEC + LA' provides similar robustness with improved spectral efficiency as compared to AL-FEC. 'AL-FEC + LA' that allows MCS setting modification at $TH^H = TH^L = 3\%$' shows in more than 95% of the cases that the application layer packet loss rate is lower than 0.1%, while the spectral efficiency sacrificed for repairing packets is compensated by improved adaptation of MCS settings, hence leading to no loss in spectral efficiency.

- With ‘AL-FEC + LA’, care should be taken in the configuration of MCS modification thresholds. Configuration of higher values, e.g., $TH^H = TH^L = 9$ %, could lead to lower <95% robustness coverage at around 0.1% application packet loss rate as compared to ‘AL-FEC’, even though better spectral efficiency is achieved.

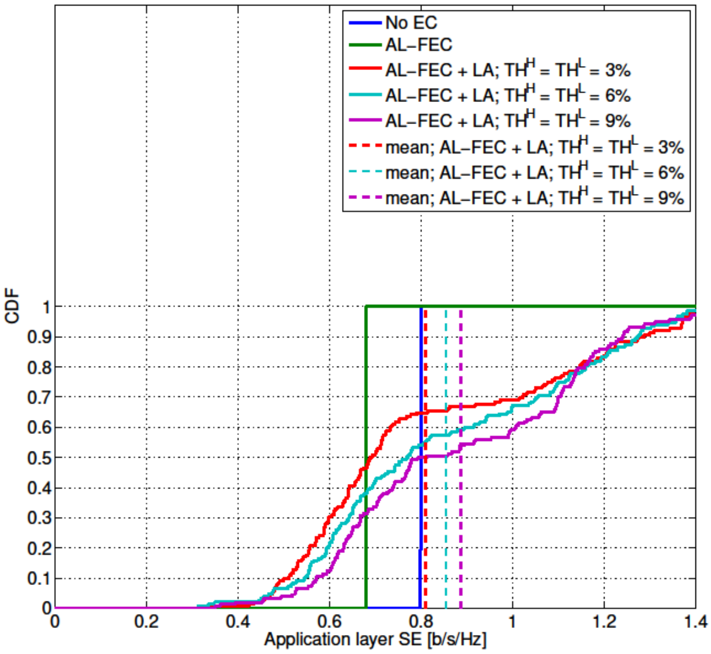

Figure 5.2-26: CDF of application layer Spectral Efficiency (SE) comparing ‘no EC’, ‘AL-FEC’ (with no LA) and ‘AL-FEC + LA’ (where both AL-FEC and LA are switched on).

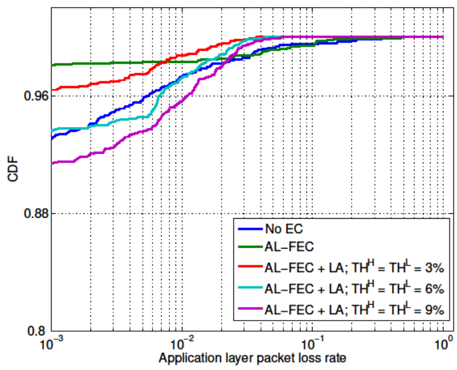

Figure 5.2-27: CDF of application layer packet loss rate for ‘no EC’, ‘AL-FEC’ (with no LA) and ‘AL-FEC + LA’ (where both AL-FEC and LA are switched on).

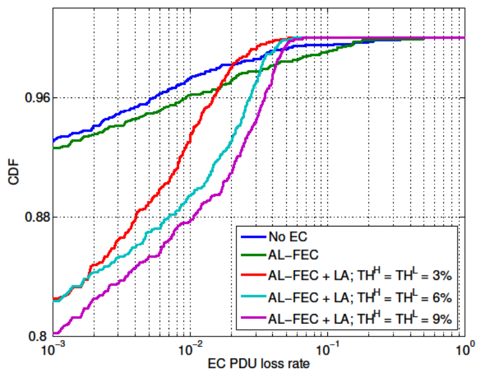

Figure 5.2-28: CDF of EC PDU loss rate for ‘no EC’, ‘AL-FEC’ (with no LA) and ‘AL-FEC + LA’ (where both AL-FEC and LA are switched on).

Figure 5.2-29 presents the CDF of application layer SE of ‘no EC’, ‘2nd Layer EC’ (with no LA) and ‘2nd Layer EC + LA’ (where both Layer 2 EC and LA are switched on). Practically, network operator can flexibly configure ‘threshold_higher’, denoted by TH^H, and ‘threshold_lower’, denoted by TH^L. To ease demonstration of the proposed scheme, simulations are performed for various sample values of a threshold where TH^H and TH^L are assigned to same value. The corresponding performance on the application layer packet loss rate and EC PDU loss rate are shown in Figure 5.2-30 and Figure 5.2-31, respectively. The PLRs are measured on EC PDUs over 1 second interval which is the higher layer EC interval. The major observations are:

- '2nd layer EC' without LA provides much higher robustness with small sacrifice on spectral efficiency as compared to 'No EC', i.e., '2nd layer EC' shows that >99.9% of the cases the application layer packet loss rate is kept below 0.1% at a cost of ~3% sacrifice on spectral efficiency. The main reason is that with layer 2 EC, repaired packets are used via on-demand re-transmission of EC PDUs (no redundant repair packets as in AL-FEC).
- '2nd layer EC + LA' can further achieve more spectral efficiency with much higher robustness as compared to 'no EC'; i.e. aggressive threshold values, e.g. $TH^H = TH^L$ = 20% or 30% can be used to harvest higher spectral efficiency, with around 31% gain as compared to 'No EC', while providing nearly 100% practical robustness via re-transmission of EC PDUs.

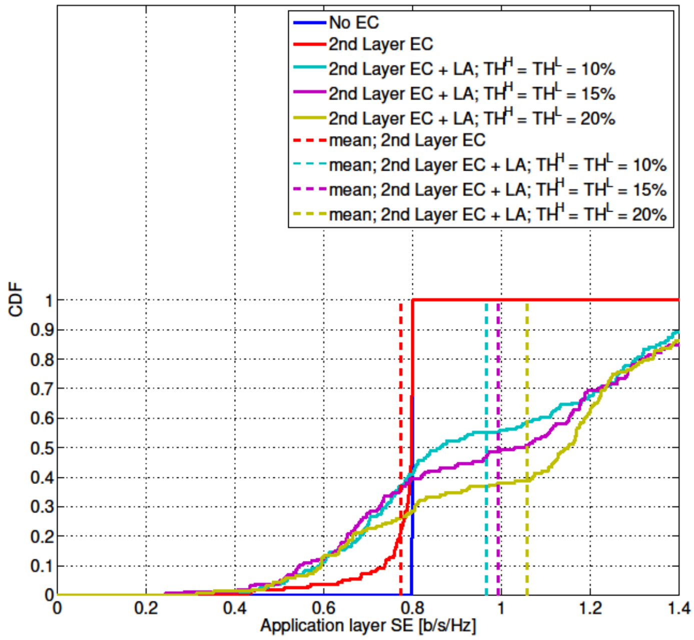

Figure 5.2-29: CDF of application layer Spectral Efficiency (SE) comparing 'no EC', '2nd Layer EC' (with no LA) and '2nd Layer EC + LA' (where both Layer 2 EC and LA are switched on).

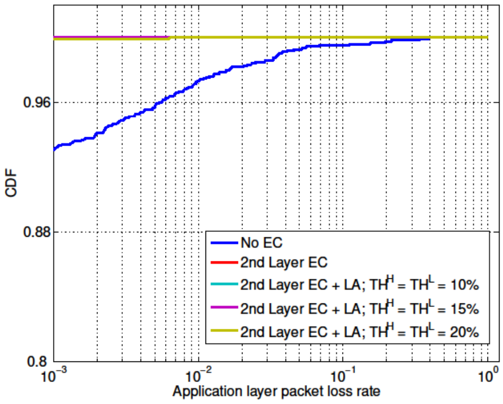

Figure 5.2-30: CDF of application layer packet loss rate for 'no EC', '2nd Layer EC' (with no LA) and '2nd Layer EC + LA' (where both Layer 2 EC and LA are switched on).

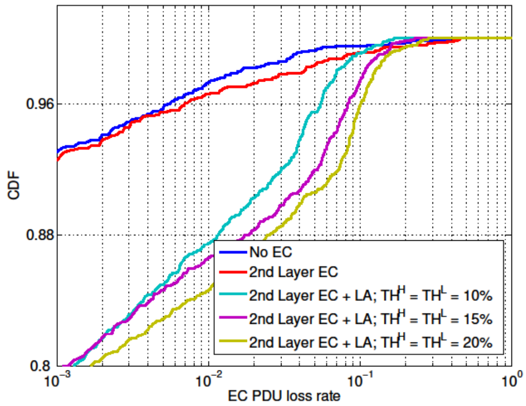

Figure 5.2-31: CDF of EC PDU loss rate for 'no EC', '2nd Layer EC' (with no LA) and '2nd Layer EC + LA' (where both Layer 2 EC and LA are switched on).

To identify the optimal MCS modification threshold, further analysis of '2nd Layer EC + LA' (where both Layer 2 EC and LA are switched on) as a function of threshold $TH^H = TH^L$ values are performed and the result is shown in Figure 5.2-32.

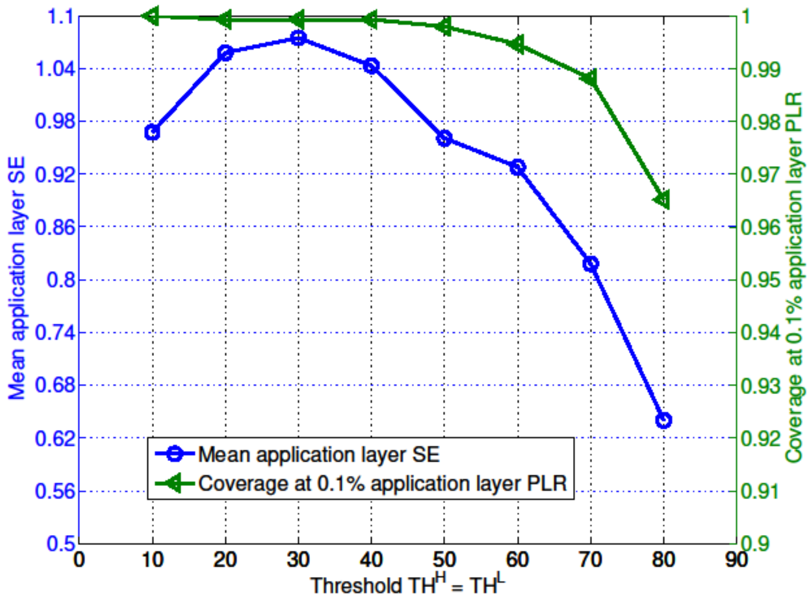

Figure 5.2-32: Further analysis of '2nd Layer EC + LA' (where both Layer 2 EC and LA are switched on) as a function of threshold $TH^H = TH^L$ values.

Throughput and block error rate evaluation against CQI

Simulation settings

To characterize the performance of the proposed NR broadcast solution in a full deployment scenario, system level simulations are conducted using the latest 3GPP Channel Model [31] and following the Urban Micro (UMi) street canyon open area scenario parameters presented therein [33]. Key parameters used are shown in Table 5.2-2 with any omitted parameters remaining identical to the calibration definitions in [31]. The MCS and CQI (Channel Quality Indicator) tables are set as specified in [32].

The cellular network layout is shown as in Figure 5.2-33, where the 285 UEs are randomly scattered around the configured 19 base stations, each serving 15 in average.

Parameter	Value
Cell layout	Hexagonal grid, 19 micro sites 3 sectors per site, full wraparound
User deployment	Random deployment, 5 users per sector, 285 users in total
Inter-site distance	200m
Indoor user ratio	80% indoor
Antenna	TR 36.897 (12° downtilt)
BW @ frequency	20MHz @ 2GHz
BS transmission power	44dBm
PTM resource allocation	Fixed frame allocation, period = 1, subframe map = 0111101111 (80%)

Table 5.2-2: Simulation parameter settings

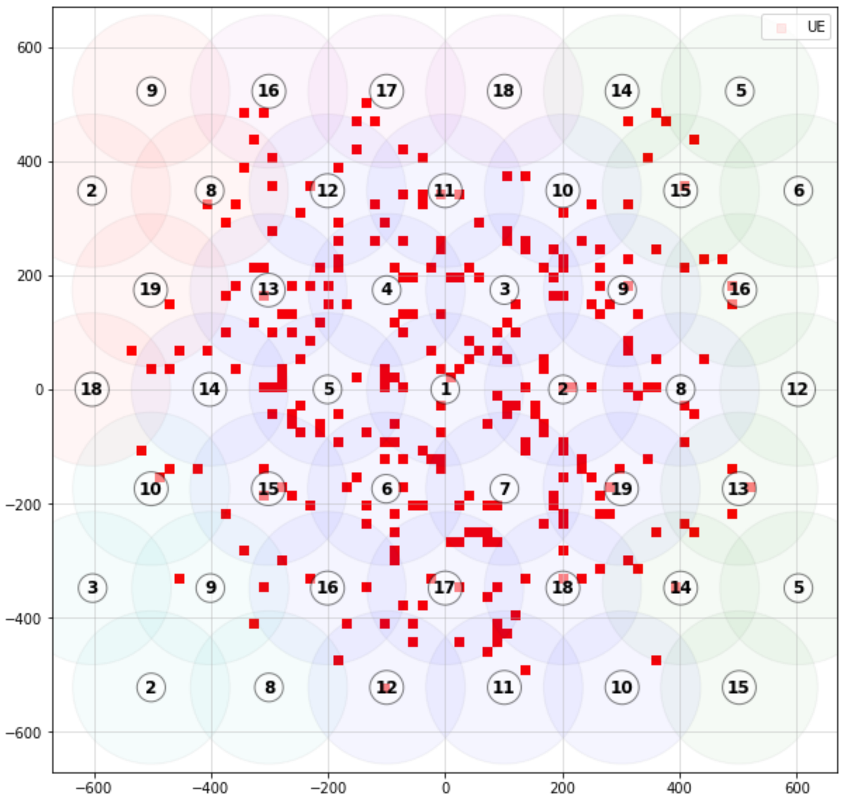

Figure 5.2-33: User distribution in simulation

Simulation results

Figure 5.2-34 presents the error rate of UEs correctly receving the 5G NR or 4G LTE-A PTM signals, alongside the mean, minimum and maximum user throughputs with increasing CQI.

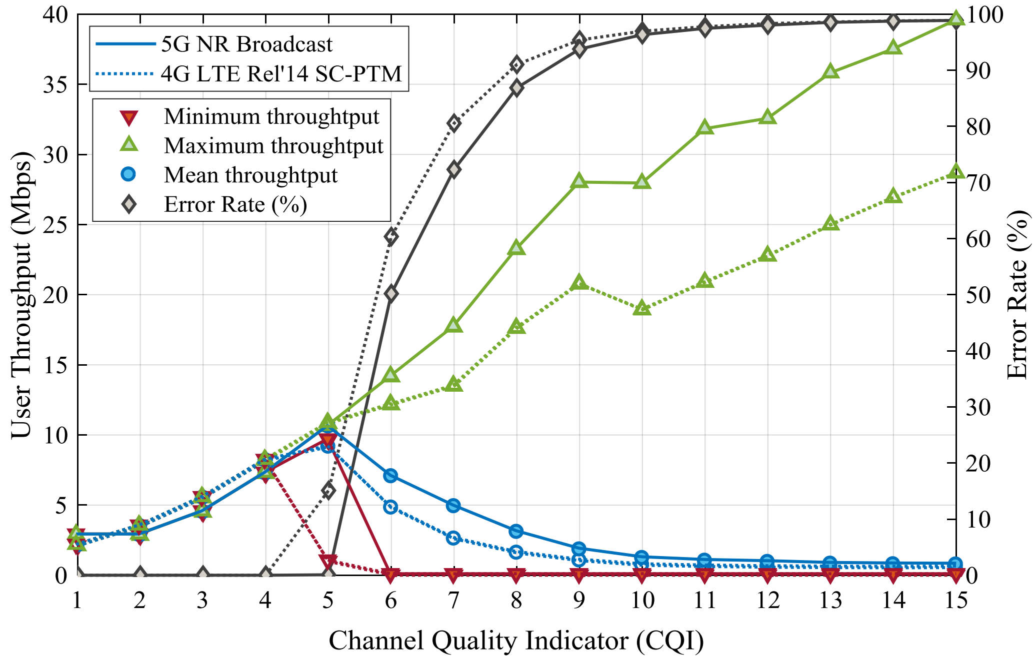

Figure 5.2-34: User throughputs in Mbps (left y-axis) and block error rate (right y-axis) per CQI transmitted with eMBMS and 5G NR broadcast

As it can be observed, the throughput and error rate increase with the CQI value, climbing close to 100% error at CQI 15, which means most data are lost as the average channel quality is far from as good as the assumed 15. However, for the few UEs which do have a good channel quality the maximum throughput can be achieved as high as nearly 40 Mbps for NR (shown with grey diamonds) and nearly 30 Mbps for LTE-A (shown with green up-pointing triangles). Of course, there are also UEs that cannot catch up with the highest rank and can only get a poor throughput as a minimum of nearly 0 (shown with red down-pointing triangles). It can be seen that 5G NR outperforms 4G LTE-A even in very limited 5G settings, where only SISO (Single Input Single Output) antenna configurations and the same bandwidth as for 4G LTE-A are used, thanks to the bigger TBS used in 5G, increased spectral efficiency, and also a new MCS table defined in 5G NR. It can be safely predicted that much more gain would be seen if using massive MIMO (Multiple Input Multiple Output), wider bandwidth on a higher frequency mmWave band with scalable numerology, and other 5G characteristics.

It is clear that for both 4G LTE-A and 5G NR there is a decisive cut off in CQI. For 4G LTE-A, between a CQI of 4 and 5 the error rate rises from 0% to over 15%. By a CQI of 6, the packet loss has climbed to over 60%. While in the case of 5G NR, the error rate rises to 8% and further to 50% at the CQI of 6. That means a drawback of the unified transmission against diversified channel quality of the all the broadcast receiver UEs, and would be considered unacceptable coverage in most cases. It is observed that for the same CQI values the throughput curves present a different story. Here for 4G LTE-A, the mean throughput for a CQI of 5 is 9.6 Mbps, a 1 Mbps improvement over a CQI of 4, while drops to 4.8 Mbps at the CQI of 6. Similarly on the side of 5G NR, the mean throughput for a CQI of 5 is 10.6 Mbps, a 3.3 Mbps improvement over the CQI of 4, and drops to 7 Mbps at the CQI of 6. This does imply that the granularity of CQI or MCS in broadcast is too limited.

NR-based PTM in contrast to 5G unicast

Simulation settings

The considered scenarios for performance evaluation are taken from 5G-Xcast deliverable D3.1 [3] which has performed benchmarking evaluation of LTE-A PTM by using ITU-R based environments. Herein, the scenarios include

- Urban 100% indoor: urban eMBB with 100% penetration of indoor UEs,
- Urban 100% outdoor: urban eMBB with 100% penetration of outdoor UEs,
- Rural 100% indoor: rural eMBB with 100% penetration of indoor UEs,
- Rural 100% outdoor: rural eMBB with 100% penetration of outdoor UEs, and
- Indoor office hotspot scenarios for eMBB use case.

Detailed parameters of the test environment considered for system-level simulations can be found in D3.1 [3] Section A.3.2.

Simulation results

Detailed analysis of system level simulations that compare NR-based PTM against 5G unicast is presented in Annex E.

The major observations from the analysis in Annex E are

- 5G unicast fully outperforms 5G PTM in case of lower number of UEs. Examples are urban 100% indoor for 10 – 15 UEs per cell; urban 100% outdoor for 10 -17 UEs per cell; and indoor office hotspot for 50 - 100 UEs in office.
- In some cases, the 5G unicast provide better average spectral efficiency than 5G PTM while the cell-edge performance (5-%ile user spectral efficiency) is lower for unicast than PTM. Examples are urban 100 % indoor for ~15 – 30 UEs per cell; urban 100% outdoor for ~17 - 30 UEs per cell; rural 100% indoor for 10 – 37 UEs per cell; rural 100% outdoor for 10 - 34 UEs per cell; and indoor office hotspot for 100 - 230 UEs in office.
- For very high penetration of UEs, the 5G PTM fully outperforms 5G unicast. Examples are urban 100% indoor and urban 100% outdoor for >~30 UEs per cell; rural 100 % indoor > ~38 UEs per cell; rural 100 % outdoor for > ~35 UEs per cell; indoor office hotspot scenario for > ~230 UEs in office.

6 Implementation guidelines for the deployment of 5G broadcast networks

The deployment of 5G-Xcast radio access is based on deployment guidelines in D3.3 [45].

7 IMT-2020 evaluation of NR

After successful calibration of the system level simulator in simplified reference scenarios (see Appendix A), system level simulations are performed for the IMT-2020 evaluation.

In the context of this project, system level simulations focus on downlink simulations in frequency range 1 (FR1), i.e. frequencies below or equal 6GHz. The considered usage scenario is eMBB with the three test environments Indoor Hotspot (InH), Dense Urban (UMa) and Rural (RMa).

There are four evaluation methodologies defined by ITU-R for system-level simulations in the course of IMT2020 evaluation [35]:

- Average spectral efficiency
- 5th percentile user spectral efficiency
- User experience data rate
- Area traffic capacity

According to the achieved values for the above mentioned evaluation metrics, it is concluded that for FR1 the 5G NR system specified by 3GPP outperforms the IMT-2020 requirements for eMBB given by ITU-R. For further details on the definition of the evaluation metrics, the parameter settings and the detailed simulation results see Appendix B.

8 Summary and conclusion

8.1 5G-Xcast RAT protocol and RRM design

The design target of 5G-Xcast RAT protocol and RRM includes resolving RAT protocol limitations of the current 3GPP's LTE-based broadcast / multicast systems that impose constraints on the RAT technical requirements in D2.1 [1] and D3.1 [3], as well as conducting studies on the RRM solutions that are expected to fulfil the functional requirement described in 3GPP's study item TR 38.913 [2]. The designed 5G-Xcast RRM solutions use 3GPP's NR as baseline for enhancement. The summary and conclusions from proposed 5G-Xcast RRM are described in Section 8.2 along with analysis of 5G-Xcast RAT protocol and RRM requirements.

8.2 Analysis of 5G-Xcast RAT protocol and RRM requirements

5G-Xcast project has considered four vertical market sectors that include media and entertainment, automotive, internet of things and public warning [1]. The basic principle of the 5G-Xcast project in various uses cases is the delivery of information to a large number of users or devices at the same time and in the same format. As such, a common RAT protocol and RRM framework is used as a design principle. Accordingly, the 5G-Xcast RAT protocol and RRM targeted resolving common and use-case specific challenges that impact requirements for the aforementioned use cases.

One of the key challenges of RRM for PTM that applies for all use cases is lack of feedback in conventional broadcast / multicast systems. Such limited support for feedback systems to assist the network to optimise the radio resources leads to challenges in terms of providing the required spectral efficiency and packet loss rates, which create constraint on requirements such as M&E1_R7, M&E1_R23, M&E1_R29, M&E1_R36, Auto1_R2. To this end, 5G-Xcast RRM techniques described in Section 4.3 address the challenge so that the aforementioned requirements are met.

The other RRM challenge common to various use cases is lack of flexible switching between PTP and PTM transmission schemes as well as mobility procedures between multicast areas that create challenges on service continuity, which in turn could constraint requirements such as M&E1_R24, Auto1_R1 and PW1_R12. 5G-Xcast RRM techniques described in Section 4.2.2 and Section 4.4.1 addresses the challenges so that the aforementioned requirements are met.

One of the use-case specific challenges in public warning use-case is limited flexibility on the trigger for Multimedia Broadcast and Multicast System. In many cases, it is the user who activates reception of multicast and broadcast content. But in case of public warning, users are not aware when such a warning message is going to be broadcasted; hence, this makes the UE to listen to the access network unnecessarily wasting UE's power which is relevant for such requirements as PW1_R5. Section 4.5 addresses this challenge so that the aforementioned requirement(s) are met.

Spectrum allocation for various use-cases is key to achieve flexible utilization of frequency bands to meet requirements such as M&E1_R38. A flexible spectrum allocation for various use cases is demonstrated in Section 4.6 fulfilling the aforementioned requirement.

The same security procedures applied to PTP transmissions should be applied in PTM ones to fulfil requirement M&E1_R20. For the terrestrial broadcast mode, if confidentiality is enforced, the relevant encryption keys should be made available to the users via other channels e.g. Unicast for non-ROM devices as demonstrated in D3.3

[45]. Optionally, for terrestrial broadcast with no unicast support for non-ROM, a theoretical approach that optimises physical layer RRM with consideration of security is described in Section 4.7 and it addresses aforementioned requirement.

In regards to expected requirement described in 3GPP TR 38.913 [2] for multicast / broadcast, the most relevant requirements from RAT protocol and RRM point of view include:

- The new RAT is expected to be **flexible** and **efficient** enough to support the requirements of existing services and new services. To this end, the RAT protocol and RRM design meets this requirement by contributions in
  - Section 4.2: Flexible resource allocation methods for 5G-Xcast
  - Section 4.3: Prospect of feedback schemes and FEC for PTM
  - Section 4.4: Efficient use of radio transmission methods
- The new RAT is expected to support efficient multiplexing of unicast and broadcast / multicast across, at least, time and frequency domains
  - D3.3 [45] and Section 4.2.1: Multiplexing among unicast and PTM transmission schemes
- The RAT is expected to support dynamic adjustment of broadcast / multicast areas based on user distribution or service requirements.
  - D3.3 [45] and Section 4.4.1: Protocol level analysis of dynamically defined multicast area

Extra information on the analysis of requirements can be found in Annex C.

8.3 Conclusion from coverage simulations

8.3.1 Analysis of 3GPP-based broadcast in substitution of DTT

By using real life scenario, the coverage simulation results in Section 5.2.1 have made a detailed analysis on proportions of areas that are not covered by conventional DTT broadcast, and on the distribution of a mobile network layer (LTE at 800 MHz) that can serve the areas not covered by DTT. The analysis focused on a scenario where mobile operators transmit video streams to users in areas where one or more DVB-T2 broadcast services are (temporary or permanently) not available. The major observation is that more than 36% of the involved cells are only marginally engaged since only 5% of less of their best server areas contains points below DTT threshold. Besides, more than 14% of the involved cells are almost totally engaged in DTT substitution as the percentage of their involved best server area is above 95%. The remaining 50% of the engaged cells have different percentages of their involved areas with a majority of them in the range between 5% and 35% of their areas. Corresponding analysis on the number of potential users (TV sets) showed that the number of potential users (TV sets) of a 5G mobile video streaming service in substitution of DTT varies significantly, from few users per cell to hundreds of users per cell.

8.3.2 Prospect of dynamic utilization of PTM and PTP

Using cell-neighbour relations and the number of users per cell at various TV transmission periods on realistic data, prospect of dynamic utilization of PTM or PTP is analysed by defining a methodology and applying it to a real life scenario. In the analysed coverage simulation in section 5.1.1 and 5.2.1, the following conclusions can be drawn.

- For the period of the maximum number of TV sets on, where large number of TV sets are switched on to consume the same content, it is shown that around

66% of the involved cells are suggested to use PTM transmission whereas the remaining 34% of the involved cells are suggested to use PTP transmissions.

- Considering the averages on number of TV sets on in the whole year of 2018, the proportion of involved cells where PTM is suggested goes down to almost 50% and PTP is suggested for remaining 50% of the involved cells.
- During the period where the least viewers watch TV, the proportion of cells where PTM is suggested is 39% which is at its minimum and the proportion of cells with PTP suggestion rises to maximum at 61% of the involved cells.

Hence, the percentages on suggested mix of PTM and PTP transmissions vary from a PTM 66% - PTP 34% mix to a PTM 39% -PTP 61% mix. In the analysed scenario, operator policies and actual requests of streaming in a live network could drive the (semi-)automatic switching between PTP and PTM transmissions within the above defined mixes.

8.4 Conclusion from system level simulations for PTM

8.4.1 Link adaptation for PTM

System-level evaluation of link adaptation for PTM is presented in Section 5.2.2. Herein, link adaptation analysis for PTM transmission with heuristic fixed MCS offsets as well as adaptive MCS via CQI report from a UE that has the worst radio link, SU-MIMO PMI and RI settings are considered. One major observation is that adaptive MCS selection, which operates at a faster time scale than higher layer FEC schemes, has limited benefit in regards to meeting coverage requirements. On the other hand, a fixed MCS setting along with cyclic PMI selection achieves >95% coverage with higher spectral efficiency as compared to fixed PMI setting mainly due to diversity benefits of cyclic PMI.

8.4.2 2nd Layer of FEC in RAN

System-level evaluation of the 2nd layer of FEC in comparison with conventional AL-FEC and “no AL-FEC” are presented in Section 5.2.2. The major finding is that by sacrificing around 20% overall spectral efficiency for redundant repair packets, the AL-FEC provides an improvement in the packet loss rate by around 35%, as compared to no AL-FEC. On the other hand, the 2nd layer of FEC avoids packet losses practically entirely, with around 3% sacrifice on average overall spectral efficiency compared to no AL-FEC, because transmissions of additional repair packets are triggered only with loss of packet, i.e., no regular redundant repair packets are used as in conventional AL-FEC. In regards to comparison of 2nd layer of FEC and AL-FEC in terms of average stalling frequency and average stalling period ratio for video streaming application, it is observed that these can be avoided practically entirely, while 2nd layer of FEC still operates at a considerably higher average spectral efficiency.

8.4.3 Cross-layer link-adaptation in co-ordination with higher layer EC

A practical mechanism of cross-layer link adaptation in coordination with higher layer EC schemes has been evaluated in Section 5.2.2 by using system-level simulation. The proposed scheme targeted improving the radio network efficiency while making the PTM transmission reliable. The findings of the SLS-based evaluation of the proposed scheme show that the application layer user spectral efficiency is considerably improved by the proposed scheme while maintaining the required QoE in terms of application layer packet loss rates.

8.4.4 Throughput and block error rate evaluation against CQI

Evaluation of 5G SC-PTM which uses NR numerology is compared against LTE-based 4G SC-PTM in Section 5.2.2. The major finding is that 5G SC-PTM outperforms 4G

SC-PTM even in very limited 5G settings, where only SISO (Single Input Single Output) and the same bandwidth as for 4G LTE-A are used, thanks to the bigger transport block size used in 5G, increased spectral efficiency, and also a new MCS table defined in NR.

8.4.5 NR-based PTM in contrast to 5G unicast

The system level simulation-based analysis that compares NR-based PTM against 5G unicast is shown in Section 5.2. The key observations include 5G unicast fully outperforms 5G PTM in case of lower number of UEs (example scenarios are urban 100% indoor for 10 – 15 UEs per cell; urban 100% outdoor for 10 -17 UEs per cell; and indoor office hotspot for 50 - 100 UEs in office). In some cases, the 5G unicast provide better average spectral efficiency than 5G PTM while the cell-edge performance (5-%ile user spectral efficiency) is lower than that of PTM for medium number of UE (example scenarios are urban 100 % indoor for ~15 – 30 UEs per cell; urban 100% outdoor for ~17 - 30 UEs per cell; rural 100% indoor for 10 – 37 UEs per cell; rural 100% outdoor for 10 - 34 UEs per cell; and indoor office hotspot for 100 - 230 UEs in office). However, for high penetration of UEs, the 5G PTM fully outperforms 5G unicast (example scenarios are urban 100% indoor and urban 100% outdoor for >~30 UEs per cell; rural 100 % indoor > ~38 UEs per cell; rural 100 % outdoor for > ~35 UEs per cell; indoor office hotspot scenario for > ~230 UEs in office).

8.5 Conclusion on system-level calibration for IMT-2020 evaluation

As a member of the 5G-PPP independent evaluation group, 5G-Xcast or in particular Nomor is responsible for the system level simulation. As a first step, a comparison against calibration results of 3GPP system-level simulators was conducted. The results are presented in Annex A. It is concluded that the presented results match very well with the 3GPP results, meaning that the system-level simulator is well calibrated against those used in 3GPP.

8.6 Conclusion on IMT-2020 evaluation of NR

In the context of this project, system level simulations for IMT-2020 evaluations have focused on downlink simulations in FR1 (frequencies below or equal 6GHz). The considered usage scenario is eMBB with the three test environments Indoor Hotspot (InH), Dense Urban (UMa) and Rural (RMa). From the evaluations, it has been concluded that for FR1 the 5G NR system specified by 3GPP outperforms the IMT-2020 requirements for eMBB given by ITU-R.

8.7 Conclusion on spectrum sharing in 5G-Xcast

Section 4.6 has presented a brief summary on the analysis of 5G-Xcast spectrum allocation options in different frequency bands for various PTM use cases. The use cases have been analysed against the spectrum bands they could use, then the spectrum bands have been analysed against the different allocation options, and the use cases have been analysed against the allocation options. Finally, all of these have been brought together in use case - spectrum band - allocation option - operator mapping.

A System-level simulator calibration for IMT-2020 evaluation of NR

A.1 Introduction

In 2012, ITU-R started to develop a vision of the international mobile telecommunication system for 2020 and beyond referred to as IMT-2020 [34]. To determine the international specification for 5G, which shall be presented in 2020, ITU-R has defined technical performance requirements in Report ITU-R M.2410-0 [35] and service and spectrum aspect requirements have been summarized in Report ITU-R M.2411-0 [36]. Furthermore, the ITU-R has specified evaluation guidelines in Report ITU-R M.2412-0 [29] to evaluate the candidate IMT-2020 radio interface technologies (RITs) or Set of RITs (SRIT) for different test environments.

Based on the schedule presented by ITU-R WP5D, proposals for IMT-2020 can be submitted from October 2017 to July 2019. The ongoing evaluation of the candidates will end in February 2020 [37].

3GPP defined a work plan for its submissions according to this timetable. At the beginning of 2018, the initial description was submitted [38]. It includes two submissions: Submission 1 is an SRIT composed by two RITs, namely NR and LTE, where NR is the term 3GPP used for the standard specified from Release 15 onwards. Submission 2 is an NR RIT.

An update, which contains the preliminary self-evaluation and link budget results and compliance templates in addition to the extended characteristics, was submitted in October 2018 [39]. The final submission is planned for July 2019 [40].

ITU-R has registered nine different Independent Evaluation Groups (IEG) [37], commissioned to verify the performance of candidate proposals for 5G. Proponents, such as 3GPP, are required to perform self-evaluation based on scenarios and constraints defined by the ITU-R in [29].

The 5G Infrastructure Public Private Partnership (5G-PPP), a cooperation between the European Commission and the European information and communication technology industry, or the 5G Infrastructure Association, representing the private side of 5G-PPP, has formed one of these registered IEG to evaluate 3GPP's proposal based on the IMT-2020 evaluation guidelines. This evaluation group mainly includes members of the EU funded phase-2 projects 5G-Xcast, 5G-MoNArch, One5G and 5G-Essence. Nomor is part of it and responsible for many of the system-level simulations, specifically those related to enhanced mobile broadband (eMBB).

In October 2018, 3GPP held a Workshop on 5G NR IMT-2020 evaluation in Brussels, Belgium. The workshop introduced the IEGs and the industry to the 5G mobile communication system developed by 3GPP. Additionally, the 3GPP submissions for IMT-2020 including the corresponding evaluations were explained and a short outlook was presented.

The first step of the evaluation process is to calibrate the system level simulator in simplified reference scenarios. Section A.2 presents the considered scenarios and the main calibration parameters including the configuration settings and calibration metrics. The calibration results of the system level simulator are compared against the 3GPP results in Section A.3.

A.2 Scenarios and calibration parameters

3GPP's calibration scenarios are largely based on the test environments defined by ITU-R in [29] which also specifies channel models, one of which in turn coincides with that defined by 3GPP in [31].

A.2.1 Test environments

For the IMT-2020 evaluation, the ITU-R defined different usage scenarios [29], namely enhanced mobile broadband (eMBB), massive machine type communications (mMTC) and ultra-reliable and low latency communications (URLLC), and combines each of them with one or several geographic environment(s) resulting in five different test environments, see Table A.2-1. These give the possibility to investigate the critical aspects in system design and performance.

Table A.2-1: Test environments defined by ITU

Scenario	Test Environments
eMBB	Indoor Hotspot - eMBB Dense Urban - eMBB Rural - eMBB
mMTC	Urban Macro - mMTC
URLLC	Urban Macro - URLLC

This document focuses on the three test environments related to the eMBB usage scenario.

A.2.2 Network layout

For the network layout no specific topography is taken into account, instead base stations are placed in regular grids [29].

For the Indoor Hotspot - eMBB test environment, 12 sites are placed at a height of 3 meter (m) with an inter-site distance of 20 m in a confined and isolated area of $120m \times 50m$, see Figure A.2-1. The scenario represents one floor of a building which has a height of 3 m with ceiling mounted base stations. Internal walls are modelled via the stochastic LOS probability model. In two variants of this scenario one site can be configured with one or three sectors or cells, respectively.

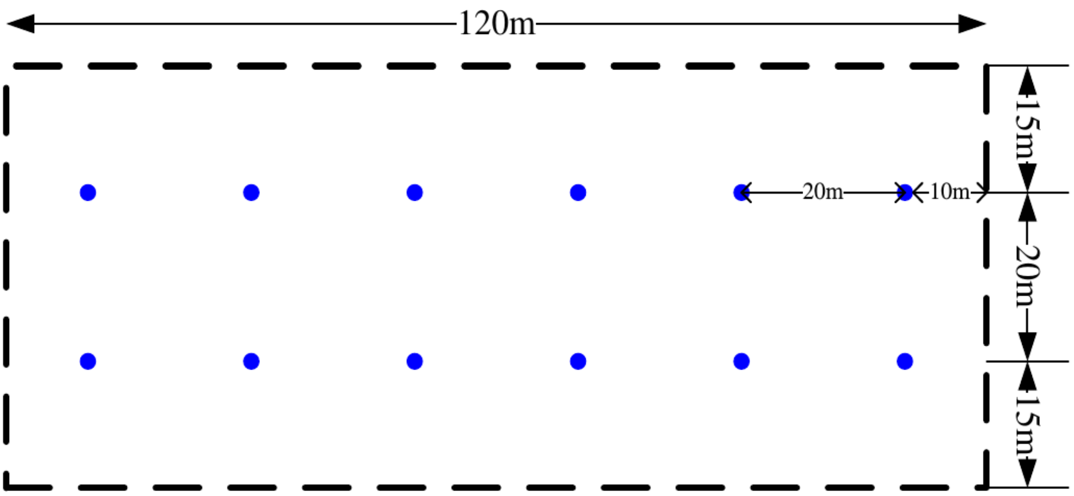

Figure A.2-1: Layout for Indoor Hotspot – eMBB [29].

The Dense Urban - eMBB test environment consists of a macro and a micro layer.

For the macro layer, a regular hexagonal layout is used, where each site has three sectors, see Figure A.2-2. In each macro cell area three micro sites are randomly dropped for the micro layer.

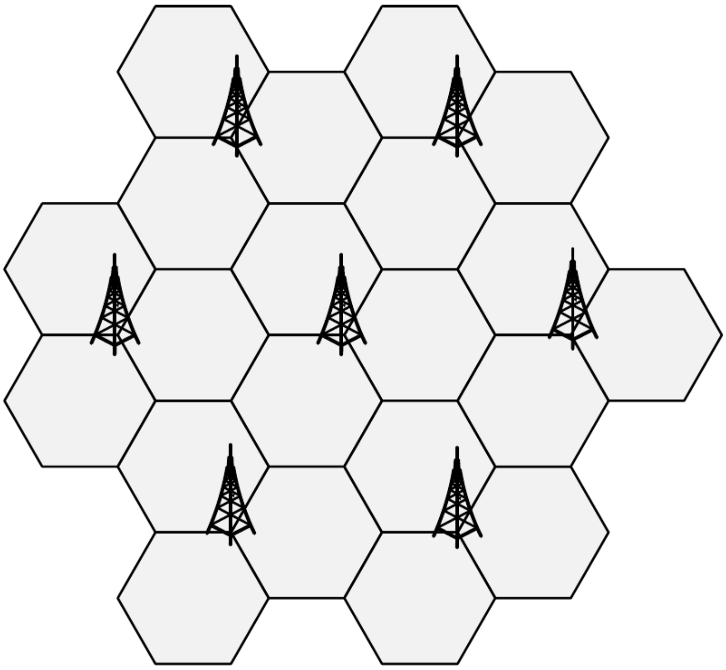

Figure A.2-2: Hexagonal site layout for Dense Urban – eMBB and Rural – eMBB [29].

For the purpose of calibration, 3GPP and therefore also herein only the macro layer is considered.

For the Rural - eMBB test environment the network deployment is the same as the macro layer of the Dense Urban - eMBB test environment, but differs in terms of inter-site distance and height of the base stations.

A.2.3 Parameter settings

In Table 5 of [29], the ITU-R defines evaluation configurations for each test environment. For several parameters as the number of antenna elements or the bandwidth, a range is given. 3GPP specified these parameters for its calibration within the framework of the self-evaluation. An overview of all parameters used is given in e.g. [41]. The following 3GPP parameter settings are applied for the calibration.

For each test environment different configurations are available. The considered scenarios with the characterizing configurations are summarized in Table A.2-2.

Table A.2-2: Scenario parameters with characterizing configuration.

Indoor Hotspot - eMBB		
	Configuration A	Configuration B
Carrier frequency	4 GHz	30 GHz
$Tx \times Rx$	32×4	64×32
GoB	-	✓

Dense Urban - eMBB		
	Configuration A	Configuration B
Carrier frequency	4 GHz	30 GHz
$Tx \times Rx$	128×4	256×32
GoB	✓	

Rural - eMBB		
	Configuration A	Configuration B
Carrier frequency	700 MHz	4 GHz
$Tx \times Rx$	64×2	128×4
GoB	Fixed down tilt	

Considering Indoor Hotspot - eMBB and Dense Urban - eMBB, carrier frequencies of 4 GHz and 30 GHz are used. Meaning two different frequency ranges are investigated; namely frequency range 1, frequencies below or equal to 6 GHz and frequency range 2, frequencies above 6 GHz. For the Rural - eMBB scenario, there are two configurations in frequency range 1, one with 700 MHz and one with 4 GHz carrier frequency.

In case of Indoor Hotspot – eMBB, Configuration A, 32 antenna elements are configured at the base station and 4 antenna elements at the UE. All antenna elements are controlled individually meaning there is a one-to-one mapping between transceiver units (TXRUs) and antenna elements.

The calibration of all Indoor Hotspot - eMBB scenarios are performed with one sector per site as well as with three sectors. As mentioned in Section A.2.2 the configuration can be selected by the proponent.

A Grid of Beam (GoB) with 8 or 12 different directions is applied at the gNB in the Indoor Hotspot - eMBB Configuration B scenario or in the two (A and B) configuration of the Dense Urban - eMBB scenarios, respectively, i.e., the antenna elements are grouped as disjoint sets into sub-array partitions served by different TXRUs. Within the TXRUs analogue beamforming is applied on the individual antenna elements, while for the combination of the different TXRUs digital precoding is used. In the Indoor Hotspot - eMBB Configuration B scenario the 64 antenna elements are grouped into 8 partitions each connected to a TXRU. Each partition has 4 columns and 2 rows of antenna elements. The TXRUs of the two Dense Urban - eMBB scenarios each feed partitions of 32 antenna elements arranged in 8 columns and 4 rows. While for Configuration A, 4 TXRUs are used, Configuration B uses 8 TXRUs.

At the UE, 4 antenna elements with a one-to-one mapping are configured for Configuration A both of Indoor Hotspot - eMBB and Dense Urban - eMBB. Considering the appropriate configurations of frequency range 2, GoB with 8 different directions is applied at the UE. 32 antenna elements are grouped into 4 partitions. Each partition has 4 columns and 2 rows of antenna elements. While for the gNB, the TXRUs or antenna elements are positioned such that the beams or patterns look all into the same direction, the partitions of the latter configurations are as separate panels positioned back-to-back to allow a reception of all different directions.

For Rural – eMBB, there is a fixed downtilt at the base station for all TXRUs. 8 antenna elements spaced in one column are fed by one TXRU. For Configuration A (carrier frequency of 700 MHz), there are 8 TXRUs; for Configuration B (carrier frequency of 4 GHz) 16 TXRUs, resulting in a total number of antenna elements of 64 or 128, respectively. On the UE side, 2 antenna elements are used for Configuration A, whereas 4 antenna elements are used for Configuration B.

At the gNB cross polarization with an orientation of +45° and -45° is applied. The orientation of the antenna elements at the UE is 0° and +90°.

For all simulations, a bandwidth of 10 MHz is applied and IMT channel model B [29] which corresponds to the 3GPP channel model for frequencies from 0.5 GHz to 100 GHz specified in TR 38.901 [31]. Further parameter settings can be found in [41].

A.2.4 Metrics for calibration

3GPP's calibration process is based on two metrics, namely Downlink Coupling Gain and Downlink Geometry.

The Downlink Coupling Gain includes the pathloss, the antenna gains and the average fast fading gains. Any processing gains at transmitter or receiver like beamforming or maximum ratio combining gain are excluded, except for analogue beamforming gains of the TXRUs where applicable.

The Downlink Geometry is the ratio of received signal power to the sum of interference and noise power where all signals are averaged individually over the used bandwidth. Like the Downlink Coupling Gain, it does not include any processing gain at transmitter or receiver except with analogue beamforming where applicable. As such the Downlink Geometry is a kind of wideband Signal to Interference plus Noise Ratio (SINR).

A.3 Calibration results

Here, Nomor’s system level simulator is calibrated against various simulators used in 3GPP, cf. [41]. The calibration results, regarding the metrics Downlink Coupling Gain and Downlink Geometry, are presented from Figure A.3-1 up to Figure A.3-16.

The results of the various 3GPP simulators are included in the figures tagged with legend entries ``3GPP #i'', the index i being that specified in [41].

The figures show a very good match of system level calibration results with the 3GPP results regarding Downlink Coupling Gain as well as Downlink Geometry. Only in Rural - eMBB, Configuration A (carrier frequency 700 MHz) our results indicate a slightly increased probability of the Downlink Geometry in the range below -3 dB, cf. Figure A.3-15.

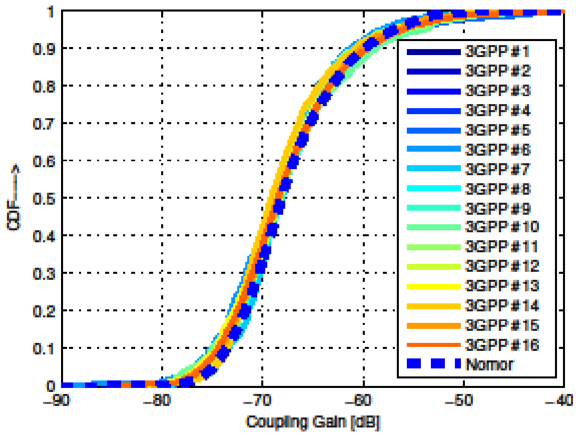

Figure A.3-1: Coupling Gain, Indoor Hotspot - eMBB, Configuration A, 1 sector

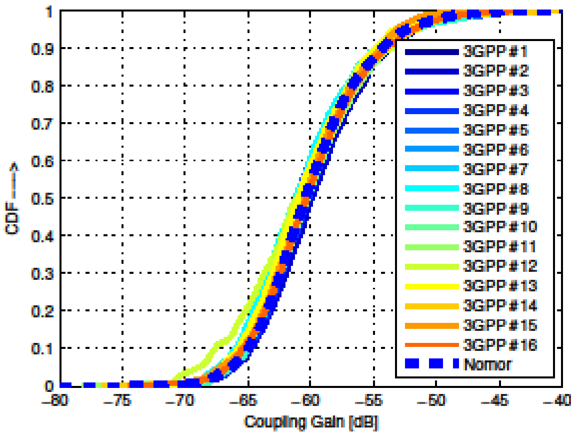

Figure A.3-2: Coupling Gain, Indoor Hotspot - eMBB, Configuration A, 3 sector

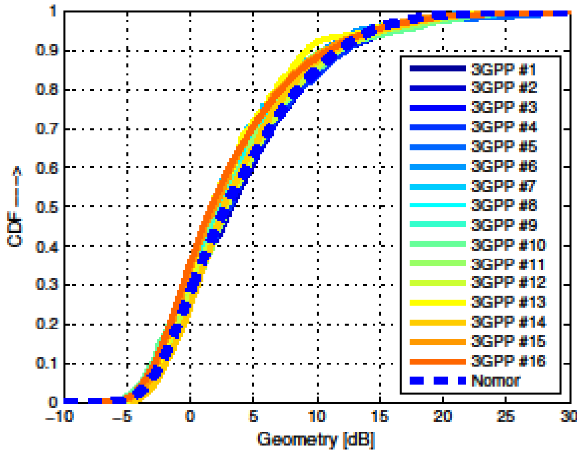

Figure A.3-3: Geometry, Indoor Hotspot - eMBB, Configuration A, 1 sector

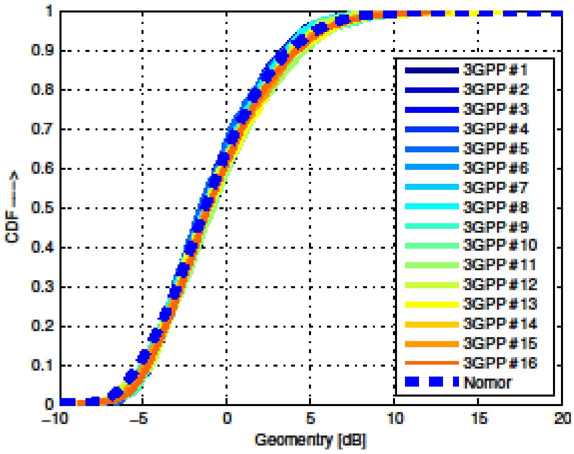

Figure A.3-4: Geometry, Indoor Hotspot - eMBB, Configuration A, 3 sector

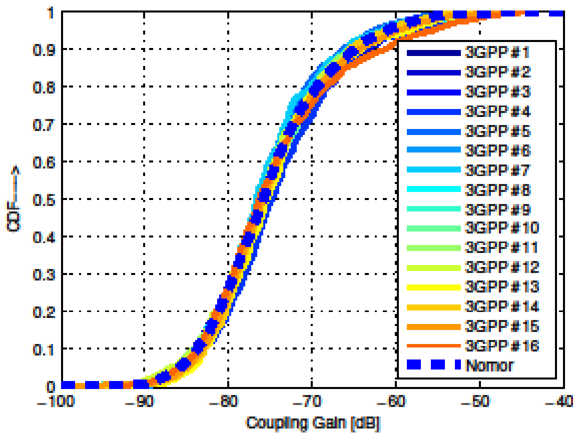

Figure A.3-5: Coupling Gain, Indoor Hotspot - eMBB, Configuration B, 1 sector

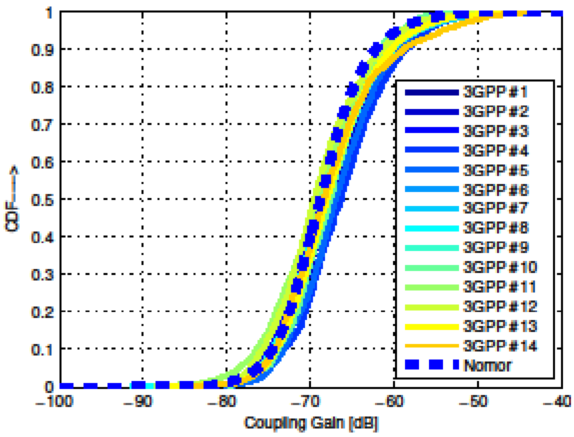

Figure A.3-6: Coupling Gain, Indoor Hotspot - eMBB, Configuration B, 3 sector

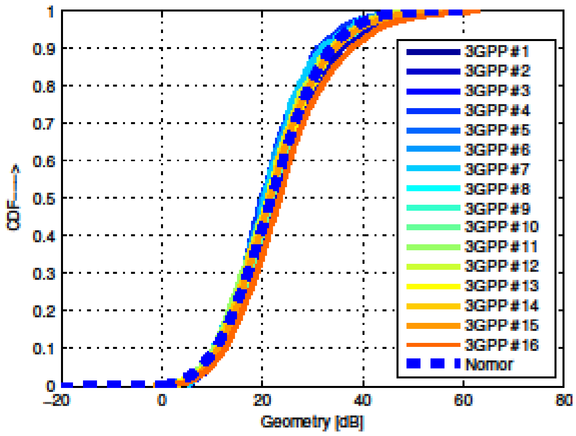

Figure A.3-7: Geometry, Indoor Hotspot - eMBB, Configuration B, 1 sector

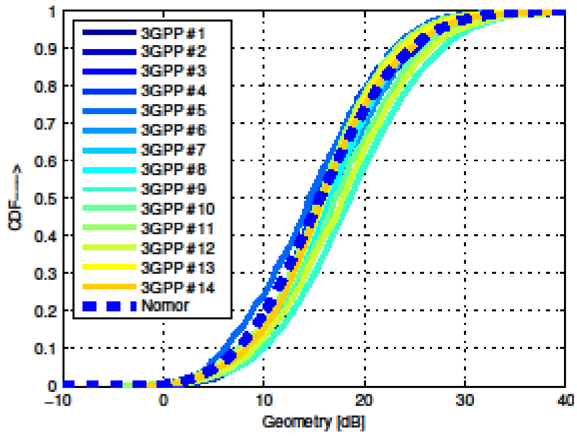

Figure A.3-8: Geometry, Indoor Hotspot - eMBB, Configuration B, 3 sector

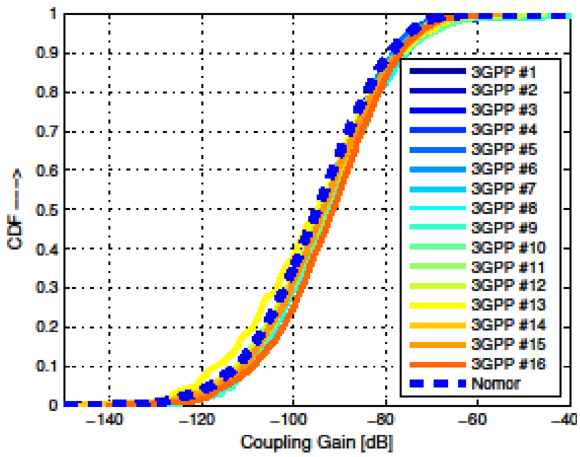

Figure A.3-9: Coupling Gain, Dense Urban - eMBB, Configuration A

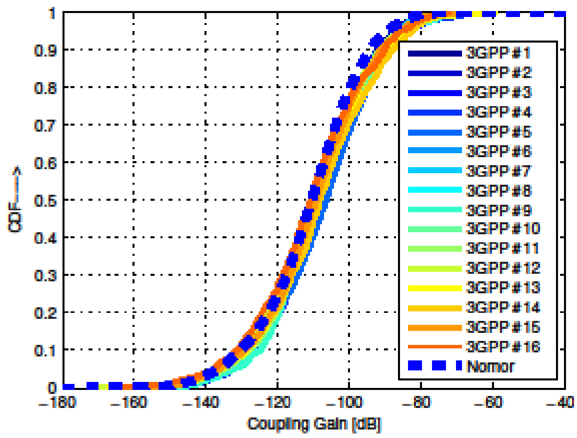

Figure A.3-10: Coupling Gain, Dense Urban - eMBB, Configuration B

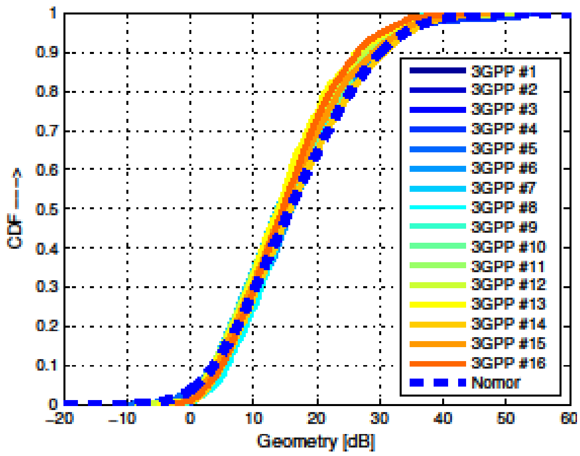

Figure A.3-11: Geometry, Dense Urban - eMBB, Configuration A

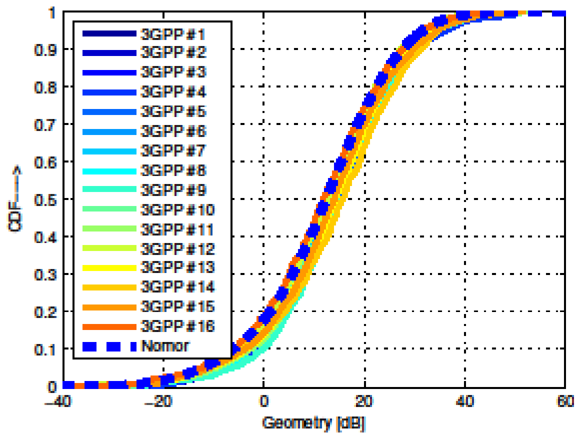

Figure A.3-12: Geometry, Dense Urban - eMBB, Configuration B

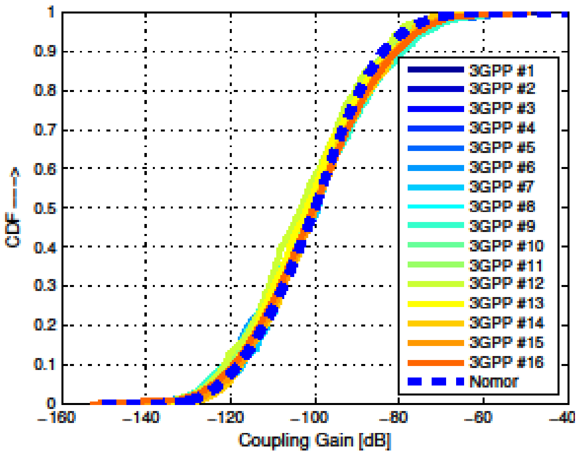

Figure A.3-13: Coupling Gain, Rural - eMBB, Configuration A

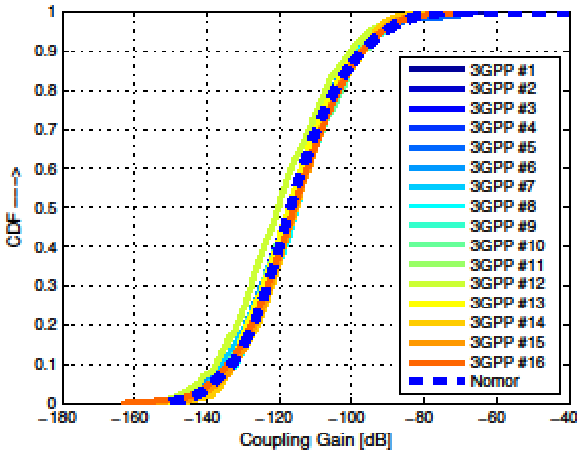

Figure A.3-14: Coupling Gain, Rural - eMBB, Configuration B

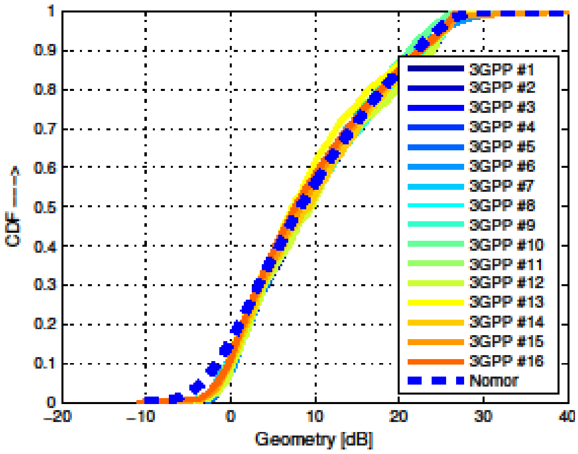

Figure A.3-15: Geometry, Rural - eMBB, Configuration A

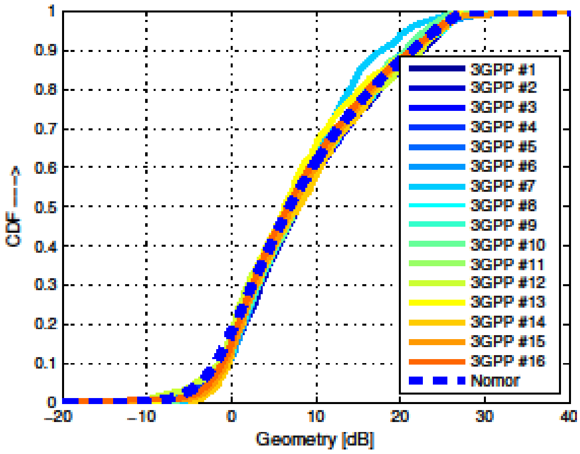

Figure A.3-16: Geometry, Rural - eMBB, Configuration B

B System-Level Simulations for IMT-2020 evaluation of NR

B.1 Evaluation methodology

The three different test environments of eMBB, namely Indoor Hotspot (InH), Dense Urban (UMa) and Rural (RMa), are evaluated via system-level simulations against four of the ITU-R key minimum technical performance requirements for IMT2020 which are explained in the following.

B.1.1 Average spectral efficiency

The average spectral efficiency is obtained by summing up the throughput of all users and dividing it by the effective bandwidth and the number of transmission reception points (TRxPs). The throughput $R_i(T)$ of user i is defined as the number of bits contained in the Service Data Units (SDUs) delivered to Layer 3 over a certain period of time T. Furthermore, the effective bandwidth BW is the operating bandwidth normalized appropriately by the ratio between UL and DL.

Considering a scenario with N users and M TRxPs where each TRxP transmits with effective bandwidthBW, the average spectral efficiency SE_{avg} is calculated by

$$SE_{avg} = \frac{\sum_{i=1}^{N} R_i(T)}{T \cdot BW \cdot M}. \quad (1)$$

The unity of metric average spectral efficiency is bit/s/Hz/TRxP [35].

B.1.2 5th percentile user spectral efficiency

For the normalized user throughput r_i of user i, the correctly received bits $R_i(T_i)$, meaning the bits contained in the SDUs delivered to Layer 3, are added up over a certain period of time T_i and divided by T_i as well as the effective channel bandwidth BW.

$$r_i = \frac{R_i(T_i)}{T_i \cdot BW} \quad (2)$$

Using the normalized user throughput of all users in a scenario and simulating many times the determined period of time, a Cumulative Distribution Function (CDF) can be created. The 5% point of this CDF is defined as the 5th percentile user spectral efficiency SE_{user} and is given in bit/s/Hz [35].

B.1.3 User experienced data rate

The user experienced data rate R_{user} is easily derived from the 5th percentile user spectral efficiency SE_{user}, by using equation (3) when one frequency band and one layer of transmission reception points (TRxPs) is applied. In case of carrier aggregation, the user experienced data rate is aggregated over the bands.

$$R_{user} = BW \cdot SE_{user} \quad (3)$$

In other words, the user experienced data rate is the 5% point of the CDF of the user throughput and is given in Mbit/s [35].

B.1.4 Area traffic capacity

In case one frequency band and one TRxP layer is applied, the area traffic capacity C_{area} can be derived from the achievable average spectral efficiency SE_{avg} as follows:

$$C_{area} = \rho \cdot BW \cdot SE_{avg} \qquad (4)$$

where ρ is the density of TRxPs per m^2. As done for the user experienced data rate, the area traffic capacity is summed over all frequency bands, as long as carrier aggregation is used. It is measured in Mbit/s/m^2 [35].

B.2 Parameter settings

In the context of this project, the performed IMT-2020 system level evaluations focus on downlink simulation in frequency range 1, i.e. frequencies below or equal 6GHz. In the parallel project 5G-MoNArch, which is also funded by the European Commission, system level evaluation for uplink in frequency range 1 is performed. Besides, downlink and uplink in frequency range 2, i.e. frequencies above 6GHz, are evaluated in terms of system level simulation in the context of 5G-MoNArch.

As mentioned in Appendix A, the focus of this project is on the eMBB usage scenario. The network layout of the three different test environments, InH, UMa and RMa, are explained in section A.2.2.

The main common system level simulation parameter settings are listed in Table B.2-1, the scenario specific antenna parameters are provided in Table B.2-2. These parameters are chosen according to the configurations applied during the self-evaluation of 3GPP towards IMT-2020, see documents of the folder "eMBB_SE.zip" which is attached to [41].

Note that for test environment InH two modes are applied, namely operating with one or three sectors per site.

Table B.2-1: System level simulation parameters settings for IMT-2020 evaluation

Parameters	Value
Carrier frequency	4GHz (700MHz for RMa, Config A)
Duplexing	TDD
System bandwidth	20MHz
Subcarrier spacing	15kHz
Frame structure	DSUUD
Transmission scheme	Closed MU-MIMO adaptation
MU dimension	up to 12 layers (up to 8 layers for RMa, Config A)
SRS transmission	precoded SRS for UE 2Tx ports
Channel model	3GPP TR 38.901 [31] (= IMT-2020 model B)

Table B.2-2: Scenario specific antenna parameters for IMT-2020 evaluation

	InH Config A, 1sector	InH Config A, 3sectors	UMa, Config A	RMa, Config A	RMa, Config B

gNB antenna configuration [M, N, P][10]	[4,4,2]	[8,16,2]	[8,8,2]	[8,4,2]	[8,8,2]
TXRUs at gNB	32 (1×1)	32 (4×2)	32 (4×1)	8 (8×1)	32 (4×1)
antenna element spacing at gNB(dH, dV)	(0.5,0.5)λ	(0.5,0.5)λ	(0.5,0.8)λ	(0.5,0.8)λ	(0.5,0.8)λ
UE antenna configuration [M, N, P][10]	[1,2,2]	[1,2,2]	[1,2,2]	[1,1,2]	[1,2,2]
TXRUs at UE	4 (1×1)	4 (1×1)	4 (1×1)	2 (1×1)	4 (1×1)
antenna element spacing at UE (dH, dV)	(0.5, –)λ	(0.5, –)λ	(0.5, –)λ	–	(0.5, –)λ

B.3 Simulation results

In the following subsections the KPI values evaluated by Nomor are discussed in comparison to the ITU-R requirements for IMT-2020 and the mean of the values submitted by different companies during 3GPP self-evaluation.

B.3.1 Indoor Hotspot

For InH average spectral efficiency SE_{avg}, 5th percentile user spectral efficiency SE_{user} and area traffic capacity C_{area} are considered as KPIs. Table B.3-1 shows, the ITU-R requirements for IMT-2020 are fulfilled in case of InH Config A (carrier frequency f_c= 4GHz) for both operation modes, 1 sector per site and 3 sectors per site, with respect to all three evaluation metrics. This is valid for Nomor's results as well as for the results given during 3GPP self-evaluation.

While for the evaluation of average spectral efficiency and 5th percentile user spectral efficiency, an frequency bandwidth of 20MHz is used, for area traffic capacity, system-level simulations are performed with a frequency bandwidth of 40MHz. The larger bandwidth provides a more efficient usage of bandwidth and a smaller overhead. Additionally, to achieve the ITU-R requirement of 10Mbit/s/m^2, carrier aggregation is applied. For the mode with one sector per site, Nomor uses 15 component carriers resulting in an aggregated bandwidth of 600MHz and an area traffic capacity of 10.31Mbit/s/m^2 . For the mode with three sectors per site, the density of TXRUs is three times larger and therefore 5 component carriers are sufficient to achieve the ITU-R requirement for area traffic capacity.

Table B.3-1: InH Config. A SLS results for IMT2020-evaluation

InH Config A		**1 sector/site**		**3 sectors/site**	
	ITU-R requirement	Nomor	3GPP self-evaluation	Nomor	3GPP self-evaluation
Avg. SE [bit/s/Hz/TRxP]	9	13.43	12.34	13.42	12.37
5%-tile UE SE [bit/s/Hz]	0.3	0.37	0.42	0.33	0.33
Area traffic capacity [Mbit/s/m^2] (used BW)	10	10.31 (600MHz)	10.38 (640MHz)	10.30 (200MHz)	11.47 (240MHz)

[10] M, N and P refer to the number of vertical, horizontal and polarization arrangement of antenna elements, respectively.

B.3.2 Dense Urban

Average spectral efficiency SE_{avg}, 5th percentile user spectral efficiency SE_{user} and user experienced data rate R_{user} are the evaluation metrics considered for test environment UMa. Considering the results of Nomor and the mean of the companies contributing to the 3GPP self-evaluation see Table B.3-2, we conclude that all ITU-R requirements for IMT-2020 are met by far for UMa Config. A (carrier frequency f_c = 4GHz).

As for area traffic capacity, a frequency bandwidth of 40MHz and additional carrier aggregation is applied to evaluate user experienced data rate. Nomor concluded that with 15 CCs each of 40MHz bandwidth a user experienced data rate of 106.88Mbit/s/m^2 is achieved. Considering the mean of the results submitted during 3GPP self-evaluation, 10 CCs are necessary to fulfil the ITU-R requirement of 100Mbit/s/m^2 for user experienced data rate.

Table B.3-2: UMa Config. A SLS results for IMT2020-evaluation

UMa, Config A	ITU-R requirement	Nomor	3GPP self-evaluation
Avg. SE [bit/s/Hz/TRxP]	7.8	14.77	14.95
5%-tile UE SE [bit/s/Hz]	0.225	0.28	0.42
User experienced data rate [Mbit/s/m^2] (used BW)	100	106.88 (600MHz)	106.75 (400MHz)

B.3.3 Rural

For evaluation of the two RMa configurations sets, average spectral efficiency SE_{avg} and 5th percentile user spectral efficiency SE_{user} are considered. Note that for Config A, where a carrier frequency of 700MHz is investigated, the UE consists of two TXRUs while for Config B, where a carrier frequency of 4GHz is simulated, there are four TXRUs at UE side, see Table B.2-2. A larger number of TXRUs at the UE means there is a higher degree of freedom for interference rejection combining (IRC).

Considering Table B.3-3, it is concluded that the 5G NR system specified by 3GPP outperforms ITU-R's IMT-2020 requirements for RMa. This is valid for both configuration sets and according to Nomor's evaluation as well as the 3GPP self-evaluation.

Table B.3-3: RMa SLS results for IMT2020-evaluation

RMa		**Config A**		**Config B**	
	ITU-R requirement	Nomor	3GPP self-evaluation	Nomor	3GPP self-evaluation
Avg. SE [bit/s/Hz/TRxP]	3.3	5.80	9.61	13.53	15.17
5%-tile UE SE [bit/s/Hz]	0.12	0.20	0.27	0.28	0.49

C Extra information on analysis of 5G-Xcast RAT protocol and RRM requirement

5G-Xcast project has considered four vertical market sectors that include media and entertainment, automotive, internet of things and public warning [1].

The first use case is media and entertainment. The media and entertainment use case includes hybrid broadcast service, virtual / augmented reality and remote live production. Herein, a wide range of application such as linear and on-demad video / audio content, streaming of social media content, linear and non-linear virtual / augmented reality content including live, and raw audio and video production feeds.

The second use case is automotive. Various V2X applications like road safety, various types of alerting, signage, mapping and autonomous driving would require information delivered from the Intelligent Transport System (ITS) infrastructure (such as ITS roadside units (RSUs), and sensors) to the vehicle. The delivery of information that would benefit multiple recipients concurrently could utilize a point-to-multipoint service. There are 39 requirements specified. Accordingly the V2X architecture and RAT protocols should meet the requirements where the V2X application server and the V-UE V2X application are involved.

In 3GPP Release 16 there are two modes for V2X Communication with MBMS reception specified: PC5 based and LTE-Uu based. The two modes can also operate simultaneously. That further develops into two implementation options: UE type RSU including a UE and the V2X logic, and gNB type RSU including an gNB, Local GW, and a V2X Application Server. In 5G-Xcast, to support the Auto 1 use case, gNB type of RSU is a preferred option as it can accommodate the requirements as the hub to coordinate and control the V2X applications. To minimise the latency on fronthaul, a gNB type of RSU should not use functional split i.e. no gNB-CU and gNB-DU implementation. The NR-Uu interface should be defined to support V1 between the V2X application on V-UE and the V2X application server on RSU, perhaps particularly for the Auto 1 network slice. Apparently Auto 1 should be operated in Mixed Mode with feedback link. The terrestrial broadcast mode cannot support all the applications in the Auto 1 use case.

The third use case considered in 5G-Xcast project is multimedia public warning alert (referred to as PW1 in D2.1 [[1]]). Herein, users are expected to be notified with alerts carrying multimedia and manifold information, which improves the effectiveness and reactivity of the users' responses. Cell granularity requirement for PWS (PWS_R4) and the same battery usage for PWS broadcast /multicast as that of unicast (PWS_R5) are some of the major requirements. The aforementioned requirements are addressed in D3.3 [45] and this document (D3.4).

The fourth use case is IoT. For IoT vertical market sectors, the requirements affecting the RAN are two as described in D3.3 [45]. On the one hand, uplink reporting available for confirmation of file delivery e.g. IoT devices confirming successful firmware update. This requires that the logical multicast / broadcast channels used for 5G-Xcast RAN allow for simple feedback (ACK on file delivery, CSI). This feature will be optional when launching a Xcast Broadcast transmission. On the other hand, IoT devices which are battery-constraint are supposed to have very low active mode time in RRC_CONNECTED state, so it is necessary to notify them when to wake up, so they can receive Software Updates or other critical signalling (IoT_R2). To satisfy this requirement, a Service Announcement transport channel is needed, not only for IoT devices but also useful for Receive-Only-Mode devices, while keeping the content transparent to 3GPP part of the network, e.g. distribution of the Electronic Programme Guide (EPG).

D Summary of state of the art in spectrum sharing

D.1 Discussion

As the amount of data in mobile networks increases, mechanisms for efficient spectrum use have been developed. Sharing spectrum between the users enables efficient utilization of valuable spectral resources. Spectrum sharing can be divided in exclusive spectrum use, i.e. no spectrum sharing, static sharing with radio licenses, dynamic sharing using electronic control like geolocation database or listen before talk equipment, and license-exempt public access. From a global perspective, practically all spectrum bands are shared. Regionally or per country, there can be exclusively allocated spectrum bands, but even then, more than 50 percent of spectrum is shared by different types of users. By far, the most common way of spectrum sharing is static sharing. Mostly but not always, radio communication using exclusive radio licenses is protected from harmful interference by the radio administration. License-exempt use is not interference protected, and dynamic spectrum sharing can be used to provide coordination for both interference protected and unprotected radio spectrum. Between licensing models and sharing types, different ways of coordination can be recognized. Radio licenses are the typical way of spectrum coordination for a radio administration. On certain bands, the radio licenses may be required but the mutual interference coordination is carried out by the industry. For example, when Programme Making and Special Events (PMSE) bands require a license, the coordination can be industry coordinated with the exception of the very large events. Listen before talk equipment can coordinate transmissions locally. One of the most common uncoordinated spectrum use for general public is Industrial Scientific and Medical (ISM) band, which is used, for example, by WiFi and Bluetooth.

There are two streams of dynamic spectrum management frameworks developed during this decade. One contains the centrally managed systems, including TV White Space (TVWS), Licensed Shared Access (LSA), and Citizens Broadband Radio Services (CBRS). The other one has dynamic spectrum sharing systems without a central coordination: Unlicensed LTE (LTE-U), Licensed Assisted Access (LAA), and MulteFire.

ETSI RRS [16] has a work item to study feasibility and technology for local high-quality wireless networks to access spectrum temporarily on a shared basis. The objective of the work is to identify how the current sharing frameworks like LSA (2.3 GHz) and CBRS (3.5 GHz) fit for this purpose. A comparison of CBRS and LSA for local temporary use can be found in [17]. The concurrent use of spectrum in 2.3 GHz frequency band and the related interference limits are defined in [18] and [19] and measurements and trials on sharing the spectrum in this band have been presented for example in [20] and [21].

With 5G, the clear new mobile spectrum trends are the need for wider bandwidths in the order of magnitude of 100 MHz and the use of the spectrum bands above 24 GHz. GSMA suggests that MNOs have interest also in unlicensed bands complementing the licensed spectrum. Due to current licenses in Europe, although widely harmonized bands and licensing conditions are strongly favoured, for example on 3.4 – 3.8 GHz 5G pioneer band, it is expected that the licensing conditions will differ from country to country.

A new spectrum opportunity is emerging by spectrum brokering. The FCC revisited rules for CBRS in 2016, and introduced the light-touch leasing process to enable secondary markets for the spectrum use rights held by PAL licensees [22]. Under the framework, no FCC oversight is required for partitioning and disaggregation of PAL licenses, and PAL licensees are free to lease any portion of their spectrum or license outside of their PPA. A study about spectrum broker for temporary licenses can be found in [23]. In addition to CBRS light-touch leasing, spectrum brokering will be useful for allocating spectrum for e.g. LSA “local and temporary high-quality networks”, Special Events, and R&D licenses.

According to Qualcomm [24] two paths for 5G development regarding spectrum sharing are foreseen — an evolutionary path and a revolutionary path as shown in Figure D.1-1. The evolutionary path includes unlicensed/shared spectrum used with a licensed anchor (similar to LAA) and a standalone operation in unlicensed/shared spectrum (similar to MulteFire), which both use LBT for coexistence. Key characteristic of the evolutionary path will be the ability to co-exist and share spectrum fairly with other technologies, such as LTE-U/LAA, MulteFire, and Wi-Fi, already deployed in unlicensed spectrum. Thus, it will be able to operate in existing unlicensed bands or in shared bands.

Figure D.1-1: 5G spectrum sharing development paths visioned by Qualcomm [24]

The ongoing 3GPP study provides a great opportunity also to explore new sharing paradigms targeting future shared/unlicensed spectrum that can deliver significant benefits in terms of increased spectral efficiencies, higher perceived user data speeds, and guaranteed bandwidth and QoS than is possible today. The revolutionary path holds the promise of enabling operators, including those with very limited or no existing licensed spectrum, to offer fiber-like 5G experiences within new shared or unlicensed bands.

The spectrum sharing can be classified in several dimensions. A few of the dimensions are discussed below. The basic licensing regimes impacting mobile access according to GSMA are Exclusive Licensed spectrum, Shared licensed spectrum, and Unlicensed spectrum [25], examples of them are LTE-A 800 MHz band 20 in Europe, 2.3 GHz band in the Netherlands, and ISM band on 2.4 GHz worldwide, respectively.

When the users have different priority levels, e.g. one user is the primary user and the other one is secondary, spectrum sharing is called vertical. When the users have the same priority level considering spectrum access, e.g. co-primary use, spectrum sharing is called horizontal. In many cases, spectrum is practically shared both vertically and horizontally, simultaneously. Examples of vertical, horizontal, and combined are 470-694 MHz band where TV is primary and PMSE is secondary, on the same band the PMSE users share the spectrum horizontally, and the band has simultaneous vertical and horizontal sharing.

Frequency coordination typically takes place on licensed bands, where users share the spectrum horizontally. Examples of manually coordinated bands are 2.3 GHz PMSE in most European countries. Electronic coordination includes the latest Dynamic Spectrum Access systems. An example of electronic coordination is 2.3 GHz PMSE use in the Netherlands, though it also requires manual involvement in the process. ISM band is uncoordinated or there is a very loose user based coordination.

Most manual coordination is managed by the radio administration, e.g. 3.5 GHz local licenses in Germany. The users coordinate the frequency use among themselves in normal situations for PMSE both on 470-694 MHz and 2.3 GHz bands. The regulator typically coordinates large events on those bands, but sometimes the coordination responsibility is assigned to an external frequency coordinator in large events, or one of the users, often the national broadcaster, coordinates the spectrum use.

Automated frequency coordination in horizontal sharing is often called coexistence management. Coexistence management has been standardized for IEEE 802.11 networks as 802.19.1 and for CBRS GAA users the coexistence management is studied in CBRS Alliance and in Wireless Innovation Forum. Automated Frequency Coordination (AFC) term was introduced for automatic frequency coordination of the coming 6 GHz band by FCC [26]. Reservation coordination in this context means a reservation system for the spectrum users similar to the Netherlands 2.3 GHz PMSE spectrum management system.

Unlicensed use in vertical spectrum sharing is difficult to implement using centralized management, but it has been deployed with spectrum sensing as Dynamic Frequency Selection (DFS) as a part of 5 GHz WiFi. Listen before talk is also based on sensing in a device and it is used for example in sharing between 802.11 and MulteFire.

Spectrum sharing currently touches most broadcasting services in the production phase. PMSE used in Outdoor Broadcasting (OB) and Electronic News Gathering (ENG) is an area where dynamic spectrum management can significantly increase the efficiency of spectrum use, due to the temporary and mobile use of spectrum by PMSE. The content distribution using broadcasting has so far been continuous in fixed location, possible sharing with PMSE on UHF band is easy to deploy with static licensing. On the other hand, the spectrum demand of broadcasting and multicasting has not been growing so dramatically that broadcasting or multicasting would be forced to look for spectrum as secondary user.

Increasing amount of video content is carried over mobile networks, especially in the countries that have subscription pricing without data gaps. Broadcast video content in the mobile networks is difficult to differentiate from general mobile broadband use, with the exception of the most popular live transmissions like football World Cup Final. In such events, broadcast and multicast reduce the load from RAN, reducing the need for additional spectrum capacity rather than create demand for dynamic spectrum access.

E System-level simulation on NR-based PTM in contrast to 5G unicast

This section demonstrates performance of 5G PTM, designed based on extension of NR to SC-PTM. The performance of 5G SC-PTM is compared against 5G unicast.

E.1 Scenarios

The considered scenarios for performance evaluation are taken from 5G-Xcast deliverable D3.1 [3] which has performed benchmarking evaluation of LTE-A PTM by using ITU-R based environments. Herein, the scenarios include

- Urban 100% indoor: urban eMBB with 100% penetration of indoor UEs,
- Urban 100% outdoor: urban eMBB with 100% penetration of outdoor UEs,
- Rural 100% indoor: rural eMBB with 100% penetration of indoor UEs,
- Rural 100% outdoor: rural eMBB with 100% penetration of outdoor UEs, and
- Indoor office hotspot scenarios for eMBB use case.

Detailed parameters of the test environment considered for system-level simulations can be found in D3.1 [3] Section A.3.2.

E.2 Result Analysis

Figure E.2-1, Figure E.2-3, Figure E.2-5, Figure E.2-7 and Figure E.2-9 show the spectral efficiency of 5G SC-PTM system for urban 100% indoor, urban 100% outdoor, rural 100% indoor, rural 100% outdoor and indoor office hotspot scenario, respectively for various MCS configuration. The corresponding CDF of RLC SDU loss rates are shown in Figure E.2-2, Figure E.2-4, Figure E.2-6, Figure E.2-8 and Figure E.2-10 for urban 100% indoor, urban 100% outdoor, rural 100% indoor, rural 100% outdoor and indoor office hotspot scenario, respectively. The major finding from the figures is that 5G SC-PTM is able to provide the expected QoE (~95% of UEs has RLC SDU loss rate below 0.1%) via configuration of QPSK, Rc = 0.19 in example scenarios urban 100% indoor, urban 100% outdoor, rural 100% indoor and indoor office hotspot scenarios. Sample configuration of QPSK, Rc = 0.19 or QPSK, Rc = 0.25 provides expected QoE performance in rural 100% outdoor scenario. The corresponding spectral efficiencies at MCS setting that fulfil the QoE requirements are considered the optimal user spectral efficiency for PTM.

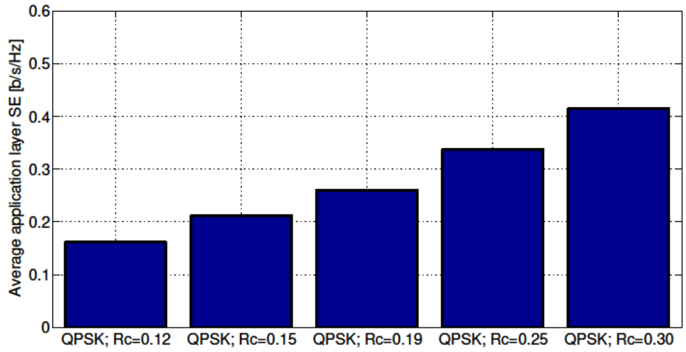

Figure E.2-1: PTM spectral efficiency (SE) for urban 100% indoor scenario.

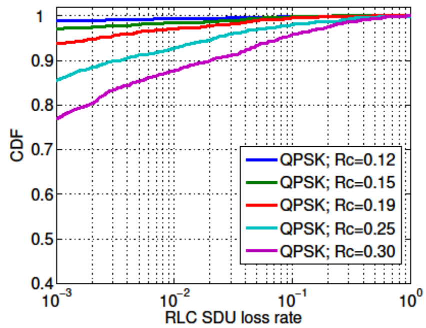

Figure E.2-2: CDF of RLC SDU loss rate for urban 100% indoor scenario.

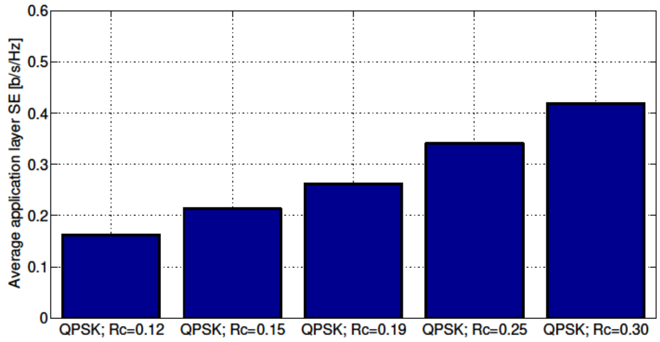

Figure E.2-3: PTM spectral efficiency (SE) for urban 100% outdoor scenario.

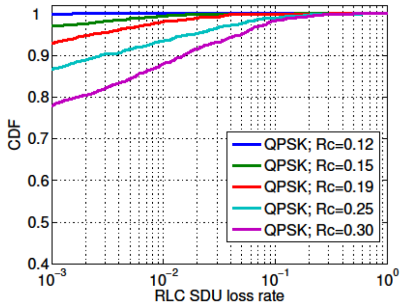

Figure E.2-4: CDF of RLC SDU loss rate for urban 100% outdoor scenario.

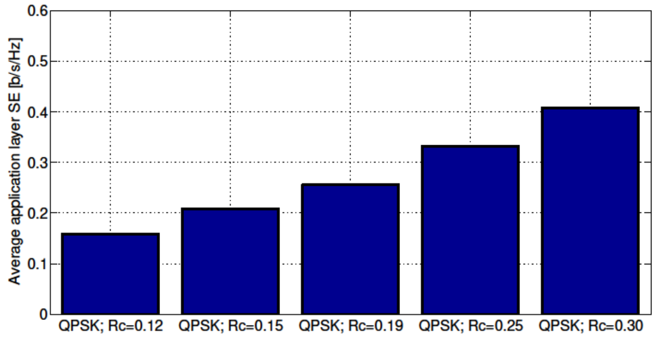

Figure E.2-5: PTM spectral efficiency (SE) for rural 100% indoor scenario.

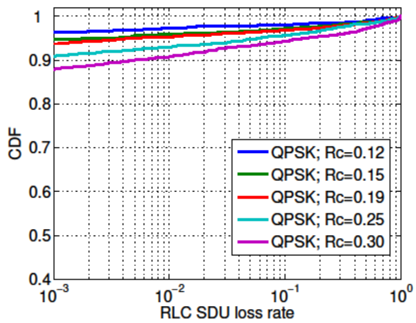

Figure E.2-6: CDF of RLC SDU loss rate for rural 100% indoor scenario.

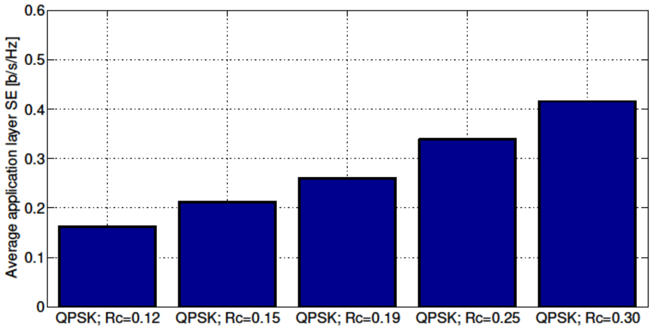

Figure E.2-7: PTM spectral efficiency (SE) for rural 100% outdoor scenario.

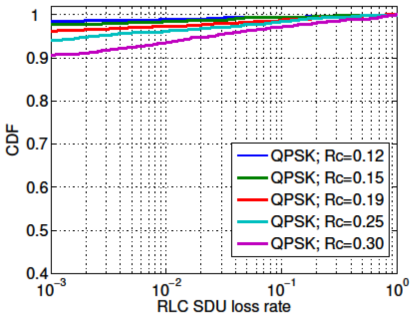

Figure E.2-8: CDF of RLC SDU loss rate for rural 100% outdoor scenario.

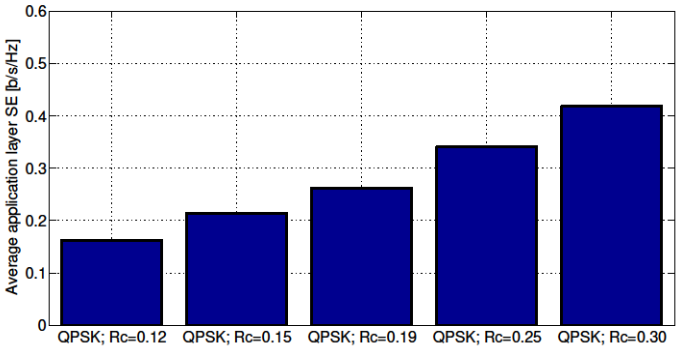

Figure E.2-9: PTM spectral efficiency (SE) for indoor office hotspot scenario.

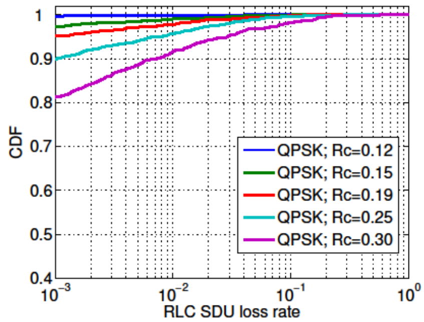

Figure E.2-10: CDF of RLC SDU loss rate for indoor office hotspot scenario.

The optimal PTM spectral efficiencies are compared with that of 5G unicast (Average and 5-%ile user spectral efficiencies) in Figure E.2-11, Figure E.2-12, Figure E.2-13, Figure E.2-14 and Figure E.2-15 for urban 100% indoor, urban 100% outdoor, rural 100% indoor, rural 100% outdoor and indoor office hotspot scenario, respectively for various numbers of UEs. Herein, the major observation is that

- 5G unicast fully outperforms 5G PTM in case of lower number of UEs. Examples are urban 100% indoor for 10 – 15 UEs per cell; urban 100% outdoor for 10 -17 UEs per cell; and indoor office hotspot for 50 - 100 UEs in office.
- In some cases, the 5G unicast provide better average spectral efficiency than 5G PTM while the cell-edge performance (5-%ile user spectral efficiency) is lower for unicast than PTM. Examples are urban 100 % indoor for ~15 – 30 UEs per cell; urban 100% outdoor for ~17 - 30 UEs per cell; rural 100% indoor for 10 – 37 UEs per cell; rural 100% outdoor for 10 - 34 UEs per cell; and indoor office hotspot for 100 - 230 UEs in office.
- For very high penetration of UEs, the 5G PTM fully outperforms 5G unicast. Examples are urban 100% indoor and urban 100% outdoor for >~30 UEs per cell; rural 100 % indoor > ~38 UEs per cell; rural 100 % outdoor for > ~35 UEs per cell; indoor office hotspot scenario for > ~230 UEs in office.

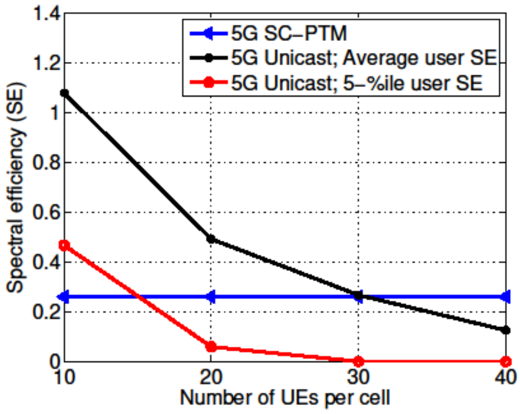

Figure E.2-11: Urban 100% indoor scenario.

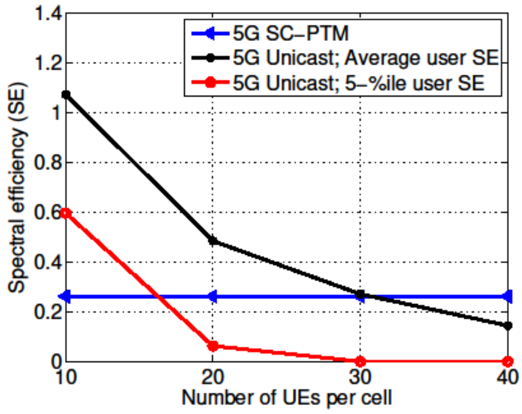

Figure E.2-12: Urban 100% outdoor scenario.

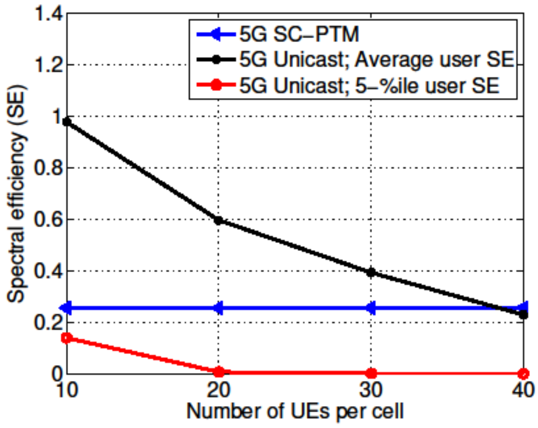

Figure E.2-13: Rural 100% indoor scenario.

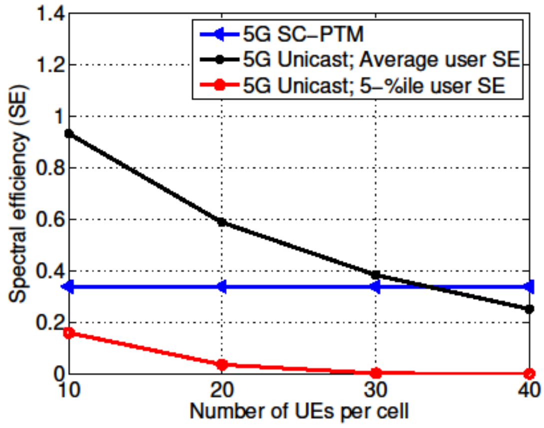

Figure E.2-14: Rural 100% outdoor scenario.

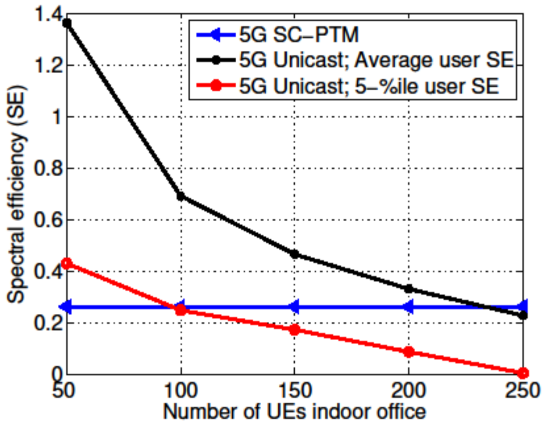

Figure E.2-15: Indoor office hotspot scenario.

References

[1] D. Ratkaj and A. Murphy, Eds., “Definition of Use Cases, Requirements and KPIs,” Deliverable D2.1, 5G-PPP 5G-Xcast project, October 2017.

[2] 3GPP TR 38.913 “Study on Scenarios and Requirements for Next Generation Access Technologies”, June 2018.

[3] D. Vargas and D. Mi, Eds., “LTE-Advanced Pro Broadcast Radio Access Network Benchmark,” Deliverable D3.1, 5G-PPP 5G-Xcast project, November 2017.

[4] 3GPP TS 38.212 v15.2.0; “NR; Multiplexing and Channel Coding (Release 15)” June 2018.

[5] 3GPP TS 26.346 v15.1.0, “Multimedia Broadcast Multicast Services (MBMS); Protocols and Codecs (Release 15)”, June 2018.

[6] 3GPP TR 36.890; “E-UTRA; Study on Single Cell Point-to-Multipoint Transmission for E-UTRA (Release 13)”, July 2015.

[7] 3GPP TS 36.300 v14.6.0; “Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); E-UTRAN; Overall description; Stage 2 (Rel. 14)”, April 2018.

[8] 3GPP TS 38.300 v15.2.0; “NR and NG-RAN Overall Description; Stage 2 (Release 15)”, June 2018.

[9] 3GPP TS 37.324 v15.0.0; “E-UTRA and NR; Service Data Adaptation Protocol (SDAP) Specification (Release 15)”, June 2018.

[10] 3GPP TS 36.323 v15.2.0; “NR; Packet Data Convergence (PDCP) Specifications (Release 15)”, June 2018.

[11] 3GPP TS 38.322 v15.2.0; “NR; Radio Link Control (RLC) Protocol Specifications (Release 15)”, June 2018.

[12] 3GPP TS 38.321 v15.2.0; “NR; Medium Access Control (MAC) Protocol Specifications (Release 15)”, June 2018.

[13] A. Awada, D. Navrátil, and M. Säilly “Study on Single-Cell Point-to-Multipoint Transmission for Public Safety Communications with eMBMS LTE Networks”, in IEEE WCNC, April 2016.

[14] T. Ho, M. Médard, R. Koetter, D.R. Karger, M. Effros, J. Shi and B. Leong “A Random Linear Network Coding Approach to Multicast”, IEEE Transactions on Information Theory, Oct 2016.

[15] M. Toemoeskoezi, F. Fitzek, D. Lucani, M. Pederson, and P. Seeling “On the Delay Characteristics for Point-to-Point Links using Random Linear Network Coding with on-the-fly Coding Capabilities”, in IEEE European Wireless Conference, May 2014.

[16] ETSI TR 103 588 V1.1.1 (2018-02). Reconfigurable Radio Systems (RRS); Feasibility study on temporary spectrum access for local high-quality wireless networks.

[17] Yrjola, Seppo & Kokkinen, Heikki. (2017). Licensed Shared Access evolution enables early access to 5G spectrum and novel use cases. EAI Endorsed Transactions on Wireless Spectrum. 3. 153463. 10.4108/eai.12-12-2017.153463.

[18] ECC, “Technical sharing solutions for the shared use of the 2300-2400 MHz band for WBB and PMSE.”, CEPT Report 58”, 2015.

[19] ETSI TS 103 235 V1.1.1 (2015-10). Reconfigurable Radio Systems (RRS); System architecture and high level procedures for operation of Licensed Shared Access (LSA) in the 2 300 MHz - 2 400 MHz band.

[20] D. Guiducci et al., "Sharing under licensed shared access in a live LTE network in the 2.3–2.4 GHz band end-to-end architecture and compliance results," 2017

IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), Piscataway, NJ, 2017, pp. 1-10.
[21] M. Matinmikko et al., "Field trial of Licensed Shared Access (LSA) with enhanced LTE resource optimisation and incumbent protection," 2015 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), Stockholm, 2015, pp. 263-264.
[22] FCC (2016) 16–55: The Second Report and Order and Order on Reconsideration finalizes rules for innovative Citizens Broadband Radio Service in the 3.5 GHz Band (3550–3700 MHz).
[23] Kokkinen T., Kokkinen H., Yrjölä S. (2018) Spectrum Broker Service for Micro-operator and CBRS Priority Access Licenses. In: Marques P., Radwan A., Mumtaz S., Noguet D., Rodriguez J., Gundlach M. (eds) Cognitive Radio Oriented Wireless Networks. CrownCom 2017. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 228. Springer, Cham.
[24] Wireless innovation —Dean Brenner. From LTE-U/LAA to 5G spectrum sharing. OnQ Blog.
[25] Laurent Bodusseau. Spectrum sharing.
[26] Paul Margie. Expanding Flexible Use in Mid-Band Spectrum between 3.7 and 24 GHz, GN Docket No. 17-183.
[27] V. Roca and B. Teibi, “Slding Window Random Linear Code (RLC) Forward Erasure Correction (FEC) Schemes for FECFRAME”, IETF Draft, march 2018.
[28] V. Roca, T. Tran, B. Teibi, C. Burdinant, and C. Thienot “Less Latency and better Protection with AL-FEC Sliding Window Codes: A robust Multimedia CBR Broadcast Case Study”, in IEEE WiMob, Oct 2017.
[29] ITU. Report M.2412-0 “Guidelines for Evaluation of Radio Interface Technologies for IMT-2020”, October 2017.
[30] T. Stockhammer, A. Shokrollahi, M. Watson, M. Luby, and T. Gasiba, “Application Layer Forward Error Correction for Mobile Multimedia Broadcasting”, Handbook of Mobile Broadcasting: DVB-H, DMB, ISDB-T and Media FLO, pages 239–280, 2008.
[31] 3GPP TR 38.901 v15.0.0; “Study on channel model for frequencies from 0.5 to 100 GHz (Release 15)”; June 2018.
[32] 3GPP TS 38.214 v15.3.0; “NR; Physical layer procedures for data (Release 15)”; September 2018.
[33] W. Guo, M. Fuentes, L. Christodoulou and B. Mouhouche, “Roads to Multimedia Broadcast Multicast Services in 5G New Radio”, IEEE BMSB, June 2018.
[34] ITU-R. ITU towards “IMT for 2020 and beyond”.
[35] ITU-R. Report M.2410-0 “Minimum Requirements related to Technical Performance for IMT-2020 Radio Interface(s)”. November 2017.
[36] ITU-R. Report M.2411-0 “Requirements, Evaluation Criteria and Submission Templates for the Development of IMT-2020”. November 2017.
[37] ITU-R. IMT-2020 Submission and Evaluation Process.
[38] 3GPP TSG SA. Draft LTI on 3GPP 5G Initial Description Template.
[39] 3GPP TSG SA. LTI on 3GPP 5G Preliminary Description Template and Self-Evaluation.
[40] Telecom Italia. Proposed way forward on IMT-2020 submission. 3GPP TSG RAN Meeting #77, RP-172098, 2017.
[41] Huawei. Summary of Calibration Results for IMT-2020 Self Evaluation. 3GPP TSG RAN Meeting #79, RP-180524, March 2018.
[42] Z. Chu, F. Zhou, P. Xiao, Z. Zhu, D. Mi, N. Al-Dhahir, and R. Tafazolli, “Resource Allocation for Secure Wireless Powered Integrated Multicast and Unicast Services with Full Duplex Self-Energy Recycling”, under review, IEEE Transactions on Wireless Communication, 2018.

[43] 3GPP TS 26.346: "Multimedia Broadcast / Multicast Service (MBMS); Protocols and codecs", September 2018.

[44] T. Tran, Ed., “Mobile Core Network”, Deliverable D4.1 v1.1, 5G-PPP 5G-Xcast project, June 2018.

[45] M. Säily, C. Barjau, Eds., “RAN Logical Architecture and Interfaces for 5G-Xcast”, Deliverable D3.3 v1.1, 5G-PPP 5G-Xcast project, February 2019.

[46] E. Garro, M. fuentes, J. Gimenez, J. Carcel, Eds, “Air Interface”, Deliverable D3.2 v1.0, 5G-PPP 5G-Xcast project, November 2018.

[47] 3GPP TS 36.331: " Radio Resource Control (RRC); Protocol specification".

[48] Y. Jung, J. Choi , “Hybrid ARQ Scheme with Autonomous Retransmission for Multicasting in Wireless Sensor Networks”, Sensors (Basel). 2017 Mar

[49] 3GPP 29.514, “5G System; Policy Authorization Service; Stage 3”, September 2018

[50] 3GPP 29.512, “5G System; Session Management Policy Control Service; Stage 3”, September 2018

[51] 3GPP 23.502, “Procedures for the 5G System”, September 2018

[52] 3GPP 29.518, “5G System; Access and Mobility Management Services; Stage 3”, September 2018

[53] F. Tesema, V. Pauli, “Layer 2 FEC in 5G Broadcast / Multicast Networks”, Nomor white paper, September 2018.

[54] L. Yu, C. Dietrich, V. Pauli, “ IMT- 2020 Evaluation: Calibration of Nomor’s System Simulator“, Nomor white paper, November 2018.

[55] EBU TECH 3348, “Frequency and Network Planning aspects of DVB-T2”, Geneva October 2014

[56] B. Mouhouche, L. Christodoulou and M. Fuentes, “Partial HARQ Retransmission For Broadcast in Fading Channels”, CS & IT 2017, October 2017.

[57] W. Guo and B. Mouhouche, “A Method to Tailor Broadcasting and Multicasting Transmission in 5G New Radio”, accepted by EuCNC 2019, June 2019.