Document Number**:** ICT-317669-METIS/D4.3

Project Name:
Mobile and wireless communications Enablers for the Twenty-twenty Information Society (METIS)

Deliverable D4.3

Final Report on Network-Level Solutions

Date of delivery: 01/03/2015
Start date of Project: 01/11/2012
Version: 1
Duration: 30 months

Deliverable D4.3
Final Report on Network-Level Solutions

Project Number:	ICT-317669
Project Name:	Mobile and Wireless Communications Enablers for the Twenty-twenty Information Society

Document Number:	ICT-317669-METIS/D4.3
Document Title:	Final Report on Network-Level Solutions
Editor(s):	Michał Maternia (NSN),
Authors:	Apostolos Kousaridas (NKUA) Osman Aydin, Stefan Valentin (ALU), Zhe Ren, Mladen Botsov (BMW), Tilak Rajesh Lakshmana, Yutao Sui, Wanlu Sun, Tommy Svensson (CTH), Emmanuel Ternon (DCM), Gabor Fodor, Nadia Brahmi, Neiva Lindqvist (ERICSSON), Renato L. G. Cavalcante, Emmanuel Pollakis, Miruna Raceala-Motoc, Slawomir Stanczak (HHI), Chan Zhou, Ömer Bulakci, Joseph Eichinger, Panagiotis Spapis, Hosein Nikopour, Kelvin Au (HWDU), Makis Stamatelatos, Sokratis Barmpounakis, (NKUA), Petteri Lundén, Athul Prasad (NOKIA), Reza Holakouei, Venkatkumar Venkatasubramanian Fernando Sanchez Moya (NSN), Marcin Rodziewicz, Pawel Sroka (PUT), Javier Lorca Hernando (TID), Ji Lianghai, Nandish Kuruvatti, Andreas Klein, Alexander Rauch (UKL), Daniel Calabuig, Sonia Gimenez Colas (UPVLC)
Dissemination Level:	PU
Contractual Date of Delivery:	01/03/2015
Security:	Public
Status:	Final
Version:	1
File Name:	METIS_D4.3.docx

Revision History

Revision	Date	Issued by	Description
1	28/02/2015	METIS	D4.3 release v1

Abstract

Research activities in METIS reported in this document focus on proposing solutions to the network-level challenges of future wireless communication networks. Thereby, a large variety of scenarios is considered and a set of technical concepts is proposed to serve the needs envisioned for the 2020 and beyond.

This document provides the final findings on several network-level aspects and groups of solutions that are considered essential for designing future 5G solutions. Specifically, it elaborates on:

- Interference management and resource allocation schemes
- Mobility management and robustness enhancements
- Context aware approaches
- D2D and V2X mechanisms
- Technology components focused on clustering
- Dynamic reconfiguration enablers

These novel network-level technology concepts are evaluated against requirements defined by METIS for future 5G systems. Moreover, functional enablers which can support the solutions mentioned above are proposed. We find that the network level solutions and technology components developed during the course of METIS complement the lower layer technology components and thereby effectively contribute to meeting 5G requirements and targets.

Keywords

Interference management, radio resource management, mobility management, 5G, multi-RAT, multi-layer, context awareness, nomadic nodes, clustering, dynamic activation/deactivation, D2D, V2X, V2V, interference identification, METIS

Executive Summary

This document examines the network-level enablers for 5G that were proposed throughout the duration of the METIS project. The specific technology components are grouped based on research challenges that need to be tackled for network technologies of 2020 and beyond. The prime objective of these studies is to develop and evaluate a set of novel network-level technology concepts that address the challenges foreseen in future scenarios with regard to interference, traffic and mobility management. We also propose several functional enablers which can support these solutions.

In the area of interference management, D4.3 provides several different concepts that aim at identification of interference sources. Particular aspects of radio resource management expected in 5G deployments are also analyzed and tackled. These aspects are: dynamic and flexible UL/DL slot allocation for time division duplexing mode, node densification, presence of moving relay nodes, provision of high data rate solutions to fast moving vehicles or connectivity of a large number of machine type devices. It is shown that for ultra dense deployments, centralized radio resource management can bring visible gains in terms of performance.

Network densification and heterogeneity of the network is also a great challenge for mobility management. This deliverable proposes several solutions to enhance mobility related aspects, such as detection of cells operating in millimetre waves region, handover optimization for moving relay nodes and D2D transmissions. Through the analysis of user device vs. network driven mobility it is confirmed that contemporary approaches (user driven active mode handovers and network driven mobility for users in idle) are a valid way forward..

Both mobility and resource management can be also enhanced by exploiting context awareness. In D4.3 several proposals exploiting supplementary information on user preferences have shown a visible improvement in user's quality of experience. An area which was studied extensively and which shown significant performance gains is the application of location related information (e.g., trajectory prediction, coverage/congestion/high capacity cell forecast, etc.). On top of evaluation of context awareness schemes, additional aspects of context awareness were addressed, e.g., gathering of appropriate data and related signalling exchange.

One of the aspects of 5G systems which was extensively investigated and discussed in METIS is D2D communication. Solutions described in this deliverable propose mechanisms for device discovery in- and out-of-network coverage, selection of appropriate mode of transmission (directly between involved devices or via radio access infrastructure) as well as algorithms for power control and SINR settings. It was proven that direct D2D operations can boost the network performance indicators (latency, power dissipation, spectral efficiency) with no or little impact to the ongoing cellular transmissions. Schemes that could be exploited in automotive safety applications are also proposed and evaluated.

Finally D4.3 describes particular enablers for the dynamic reconfiguration of the network. Several mechanisms, focused on activation and deactivation of appropriate access nodes in order to provide energy efficient network operations, are discussed. Different aspects of potential future 5G deployments are taken into account e.g., decoupling of control/user plane or exploitation of moving relays and nomadic nodes. Some algorithms for optimization of backhaul selection are also proposed. Another group of solutions covered in this document is mechanisms based on clustering. Schemes exploiting clustering of cells and devices are presented. Especially the latter ones, based on a cluster head operations, show a good potential to provide an efficient radio resource management for machine type communication.

Contents

List of Figures

List of Tables

List of Abbreviations

3GPP	3rd Generation Partnership Project
ABS	Almost Blank Subframes
ACK	Acknowledgement
APNet	Antennas, Propagation and Radio Networking
BLER	Block Error Rates
BRA	Balanced Random Allocation
BS	Base Station
CAPEX	Capital Expenditure
CARA	Context Aware Resource Allocation
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CFO	Carrier Frequency Offset
CH	Cluster Head
CO	Confidential
CoMP	Coordinated Multipoint
CPA	Cellular Protection Algorithm
CSG	Closed Subscriber Group
CSM	Code Space Multiplexing
CSIT	Channel State Information at Transmitter
CSPO	Cell Specific Parameter Optimization
DBA	Dynamic Bandwidth Allocation
DFT	Discrete Fourier Transform
DL	Downlink
DMRS	Demodulation Reference Symbol
DR	Dedicated Resources
DRX	Discontinuous Reception
DUE	Device-to-Device User Equipment
DWDM	Dense Wavelength Division Multiplexing
D2D	Device-to-Device
eICIC	Enhanced Inter Cell Interference Coordination
eNB	Enhanced NodeB
EPC	Evolved Packet Core
E2E	End-to-End
FDD	Frequency Division Duplex
FFT	Fast Fourier Transform
FRA	Future Radio Access
FRN	Fixed Relay Node
GAP	General Assignment Problem
GNSS	Global Navigation Satellite System
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile communications
GSM-R	Global System for Mobile communications – Railway
HARQ	Hybrid Automatic Repeat Request
HeNB	Home Enhanced NodeB
HO	Handover
HOM	Handover Margin
ICT	Information and Communications Technology
IEEE	Institute of Electrical and Electronics Engineers
IG	Information Gain
IIE	Interference Identification Entity
IMT	International Mobile Communications
IRC	Interference Rejection Combining
ITU	International Telecommunication Union
JCAC	Joint Cell Admission Control
JT	Joint Transmission
ITU-R	International Telecommunication Union-Radio
KPI	Key Performance Indicator
LAN	Local Area Network
LDRAS	Location Based Resource Allocation Scheme
LS	Least Square
LTE	Long Term Evolution
LTE-A	Long Term Evolution-Advanced
MA	Multiple Access
MAC	Medium Access Control
MBR	Minimum Bit Rate
MBSFN	Multi-Broadcast Single-Frequency Network
MC	Macro Cell
MH	Multi Hop
MIMO	Multiple-Input Multiple-Output
MMC	Massive Machine Communication
MMSE	Minimum Mean Square Error
MMW	Millimetre Waves
MPA	Massage Passage Algorithm
MRC	Maximum Ratio Combining
MRN	Moving Relay Node
MSE	Mean Square Error
MU-MIMO	Multi User Multiple Input Multiple Output
M2M	Machine to Machine
NN	Nomadic Node
OLPC	Open Loop Power Control
OP	Outage Probability
OPEX	Operational Expenditure
OPI	Overall Performance Indicator
PC	Power Control
PDMS	Proactive Delay Minimized Scheduling
PDR	Peer Discovery Resource
PDU	Packet Data Unit
PHY	Physical Layer
PRACH	Physical Random Access Channel

PSS	Primary Synchronization Signal
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RACH	Random Access Channel
RAN	Radio Access Network
RAT	Radio Access Technology
RB	Resource Block
RE	Resource Element
RF	Radio Frequency
RRM	Radio Resource Management
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSSI	Received Signal Strength Indicator
RX	Reception
SC	Small Cell
SCM	Spatial Channel Model
SDR	Software Defined Radio
SDN	Software Defined Network
SISO	Single Input Single Output
SM	Spatial Multiplexing
SON	Self Organizing Network
SOTA	State Of the Art

SP	Service Provider
SR	Shared Resources
SSPO	Street Specific Parameter Optimization
SSS	Secondary Synchronization Signal
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TRX	Transmitter
TTI	Transmission Time Interval
TX	Transmission
UDN	Ultra Dense Network
UE	User Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunications System
VUE	Vehicular User Equipment
VPL	Vehicular Penetration Loss
V2D	Vehicle-to-Device
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Anything
WCDMA	Wide Code Division Multiple Access
WLAN	Wireless Local Area Network

1 Introduction

The fifth generation of mobile cellular technologies (5G) has become one of the most widely addressed research areas in the IT worldwide. Looking back at the historical development of the telecommunication systems, the advent of the 5G seems inevitable, considering periodical and systematic introduction of new cellular generation: analogue solutions in the 80's (1G), transition to a digital domain and proliferation of the handheld devices in the 90's (2G), first broadband solutions in the very first decade of the new millennium (3G) and finally, a recently introduced high data rate commodities of Long Term Evolution-Advanced (LTE-A) also commonly referred to as 4G (WiMAX is not considered in this document and whenever the term 4G is used, we refer to technologies standardized by 3rd Generation Partnership Project (3GPP)). Seeing this periodical introduction of new cellular generations, it is almost certain that the fifth one will be introduced in the 2020's and beyond, but this generation shift is not only a historical necessity. Also International Telecommunication Union - Radiocommunication Sector (ITU-R) has recognised the need for 5G and is working to establish a process that will foster the development of this technology [ITUR14-M2320].

It is a unique opportunity for creation and standardization of a new technology that will be able to address the needs and expectations of services for the future network information society. These services identified by METIS are [MET14-D63]:

- **Extreme/advanced/flexible Mobile Broadband (xMBB):** provides high data rates and low-latency communications and improves Quality of Experience (QoE) through a more uniform experience over the coverage area, and graceful degradation of data rate and increase of latency as the number of users increases.
- **Massive Machine-Type Communications (M-MTC):** provides up- and down-scalable connectivity solutions for tens of billions of network-enabled devices.
- **Ultra-reliable/Critical MTC (U-MTC):** provides ultra-reliable low-latency communication links for network services with extreme requirements on availability, latency and reliability, e.g. V2X communication and industrial control applications.

At the moment of the creation of this document it is impossible to pin point the exact needs and requirements for these services and how they will exactly look like. It is however likely that 5G will have to cater for a large variety of use cases and demands. However, it is reasonable to assume that they can be generalized into several objectives such as:

- 1000 times higher mobile data volume per area,
- 10 to 100 times higher number of connected devices,
- 10 to 100 times higher typical user data rate,
- 10 times longer battery life for low power Massive Machine Communication (MMC) devices,
- 5 times reduced End-to-End (E2E) latency,

These goals should be achieved at a similar cost and energy consumption as today's networks [MET13-D11].

This document is one of the final deliverables of Mobile Enablers for Twenty-twenty Information Society (METIS) project, which has the ambition of setting the path toward the future 5G networks. It is the continuation of earlier METIS deliverables that deals with network level solutions for 5G: D4.1 [MET13-D41] and D4.2 [MET14-D42]. This series of deliverable focuses on providing solutions for future 5G multi-layer and multi-radio access technologies (multi-RAT), defined from the network level perspectives. Readers are also encouraged to familiarize with refined set of use cases and scenarios for 5G [MET15-D16], link level enablers

[MET15-D24], multi-antenna and multi-node coordination solutions [MET15-D33] as well as spectrum related considerations in [MET15-D54] and architectural proposals of [MET15-D64]. These documents, together with this one, will lay the foundation for the final METIS system level concept that will be presented in [MET15-D66].

1.1 Objective of the Document

The main objective of this document is to present network-level solutions for future mobile and wireless 5G communication systems, with the ambition of providing answers to the research challenges that could be summarized in following questions:

- What are the key enablers for dynamic network reconfiguration and why is it needed?
- How can we provide energy efficient network operations?
- Do we need new schemes for mobility management in the future and in which domains?
- Can we use context awareness to enhance user experience?
- What can we gain from D2D network-level solutions and how do they serve the 5G requirements?
- Do we need cell or device clustering in 5G?

Initial answers to these questions are provided using a set of technology components: network level solutions that are targeting at specific problems and challenges experienced in contemporary cellular networks and predicted for the future. Solutions to these problems are the result of the investigations carried out by leading European vendors, operators and academic partners.

1.2 Structure of the Document

Network-level technical solutions investigated in METIS and described in this document can be grouped into several research areas. The structure of this document reflects this grouping. Chapter 2 describes interference management and resource allocation schemes as well as techniques that aim at enhancing interference detection. Mobility management and robustness enhancements are described in Chapter 3 while context awareness approaches are discussed in Chapter 4. Technology components related to Device-to-Device (D2D) communication and Vehicle-to-Anything (V2X) are presented in Chapter 5. Chapter 6 covers dynamic reconfiguration enablers with a special focus on mechanisms targeting at energy efficiency for the radio access infrastructure and clustering.

Each chapter gives a motivation highlighting why these particular solutions are needed and after the description of the proposed 5G mechanisms, a short summary follows. The subchapter ‘Impact on the METIS 5G System Concept’ in each chapter relates the specific group of technology components to three main services identified by METIS for 5G systems: xMBB, M-MTC and U-MTC. The considered groups are also analyzed from the point of view of the METIS overall technical goals established at the beginning of the project. This analysis is captured in subchapter ‘Addressed METIS Goals’.

Finally, Chapter 7 provides an overall summary for the network-level research activities carried out in the METIS project. This chapter also contains the description of the identified gaps that were not specifically addressed in the project, but are important for the definition of future 5G systems.

Interested readers may also find an extended description of some network-level investigations and basic dependencies identified for considered solutions in the Annex part of this document.

2 Interference Management and Resource Allocation Schemes

2.1 Introduction

Smart resource allocation and proactive interference management are key enablers for efficient utilization of radio resources (such as power, frequency or time slots) available to the radio access network (RAN). This research area has been extensively studied, and there are numerous publications already available, that consider a variety of interference identification and resource allocation schemes (c.f. [And05], [EHB08], [GCC+10], [LHW+10], [LZW+11] or [SB12]). However, 5G solutions bring new challenges for radio resource management (RRM) that necessitate schemes and enablers different than the ones used in previous generations of cellular communication systems. Most important challenges and factors are:

- Shrinking cell sizes (much below 100 m) and a smaller number of users per cell. This results in more 'flashy' activity profile of the cell and bigger impact of intercell interference dynamics on the performance.
- Dynamic time division duplex (TDD) air interface proposed by METIS for ultra-dense network (UDN) in [MET15-D24]. This way of operations gives a potential for a dynamic adaptation to instantaneous traffic demands, but on the other hand can become a source of heavy interferences due to the decoupling of UL/DL slot allocation between neighbouring cells. For example UL transmission from UE1 in cell1 can interfere data reception of UE2 in adjacent cell2 configured for DL transmission.
- Increased heterogeneity of the 5G networks. Future solutions will not only comprise of the cells with different sizes and transmission power (i.e., macro, micro, pico or femto nodes), but will also include nomadic nodes.
- Wide spectrum of supported services starting from bandwidth hungry ultra-high definition (or similar) video services [ITUR14-BT2020] and ending on a very low bit rate, but energy efficient MMC operations.

The advent of a new generation of cellular systems creates a unique opportunity to embed solutions to the above mentioned issues in a clean slate design. Mechanisms described in this particular chapter are METIS proposals for interference management and resource allocation schemes. It also gives a description of mechanism that provides input to the RRM allowing estimation and identification of dynamic interference environment.

First four technology components described in this chapter focus on different methods that could be exploited for interference identification. The rest of the chapter covers solutions for interference management through appropriate allocation of radio resources.

2.2 Adaptive Projected Sub-gradient Method (APSM) for Interference Identification

Interference identification is the task of obtaining information about the expected interference at a location of interest. This knowledge is considered as a key aspect for improved quality-of-service (QoS) by better utilization of scarce wireless resources. The information about spatial and temporal interference patterns is highly desired to support proactive network reconfiguration algorithms and D2D communication. A basic ingredient to such interference patterns is the two-dimensional map of path-loss values. The proposed solution aims at reproducing such path-loss maps from sparse measurements arriving over time.

The time-varying long term channel gain matrix is estimated under the assumption of channel reciprocity and based on the information about the geographical position of wireless devices and channel measurements. Set-theoretic approaches are adopted where measurements and a priori knowledge (tuples of geolocation and path-loss information) are used to construct

closed convex sets. The current estimate of the channel gain matrix is found in the intersection of these sets. The sets are updated when new measurements arrive.

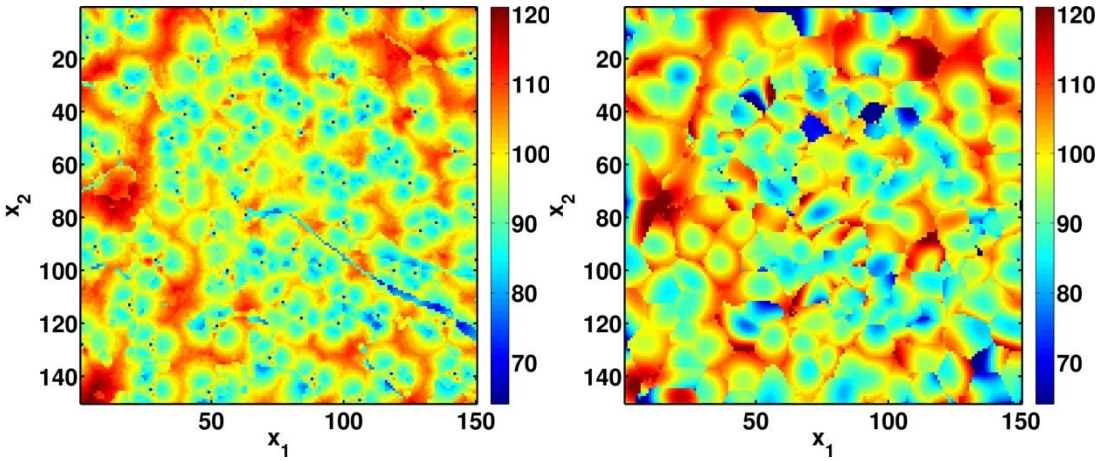


Figure 2-1. True path-loss map (left) and estimated path-loss map (right) reconstructed with the online-learning technique of APSM. (x_1,x_2) are coordinates of the map and colour coding corresponds to path-loss values in dB.

The application of path-loss estimation with the proposed APSM algorithm is evaluated in a realistic urban communication scenario. Figure 2-1 depicts true path-loss map (left) and the estimated path loss map (right) based on approximately 1000 measurements. A quantitative evaluation yields significant similarities between the estimated and the original path-loss map.

2.3 Minimum Mean Square Error (MMSE) Estimation for Interference Identification

We are interested in an estimate of the time-varying channel gain matrix. It is obtained by means of a statistical estimation approach that combines the measurements with (i) statistical knowledge of measurement uncertainty, and (ii) prior knowledge of spatial correlation of the interference links. We assume known positions of user equipments (UEs) and base stations (BSs) and known path loss exponents from which the a priori distribution of the channel gain matrix with a mean and a covariance matrix is derived.

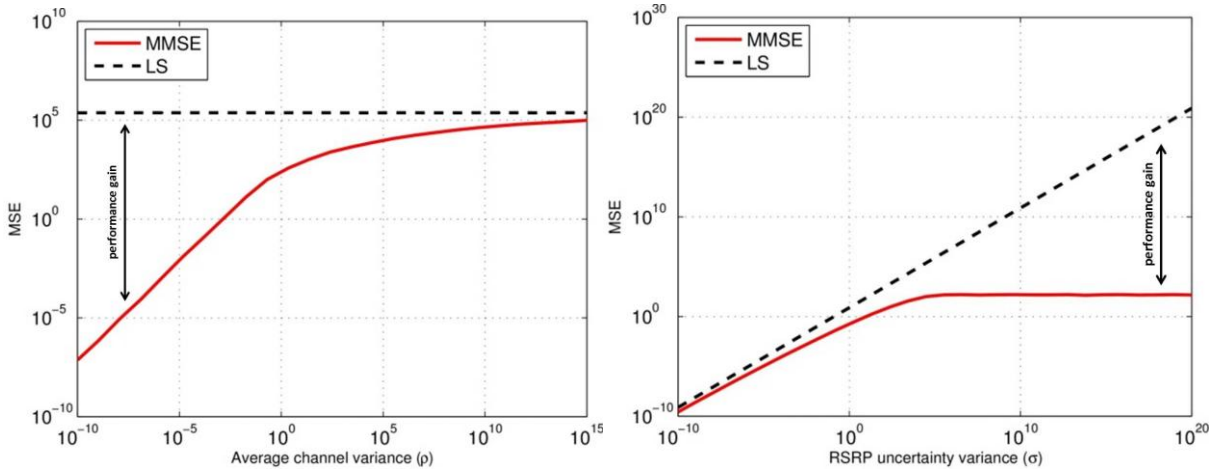


Figure 2-2. (a) Comparison of MMSE and LS performance for different channel variance ρ, (b) Comparison of MMSE and LS performance for different noise levels σ

Statistical knowledge about the channel gain vector and measurement uncertainty is exploited. Given some physical-layer measurements, an ideal linear model in which the prior distribution of the interference matrix and the uncertainty distribution is Gaussian in linear scale is derived. This model relates the measurements to the channel gain vector and

therefore can be used to derive an optimal linear MMSE (LMMSE) estimator for the channel gain vector. Since slow fading caused by shadowing is often assumed to have a log-normal distribution, a more realistic model in which the prior path-loss distribution is log-normal and the uncertainty distribution is Gaussian in dB scale is used. In this case, the model becomes non-linear, and therefore a closed-form "linearized" MMSE estimator, named linearized log-MMSE (LLMMSE), is derived to estimate the channel gain vector.

The results presented here show how the accuracy of interference estimation obtained from the proposed MMSE Estimator is affected by two system parameters, namely the Reference Signal Received Power (RSRP) uncertainty σ and the channel variance ρ. The performance of the MMSE is compared to the simple least squares (LS) estimator. The simulation results in Figure 2-2 show that the proposed MMSE estimator outperforms the LS estimator. The gains are large for high noise levels or when the channel variance ρ is small. The performance in low noise situations is similar to the LS performance as in such cases the solution of the MMSE estimator converges to the one of the LS estimator. Same behaviour is observed when the channel variance is high.

2.4 Interference Identification Using Multi Layer Inputs

The objective of this technology component is to provide a mechanism capable of identifying sources of interference in environments where topology information is available. Specifically, the introduction of a centralized entity, the Interference Identification Entity (IIE),is proposed that exploits available knowledge of LTE-A-like network components, namely the UEs, the (home) eNBs ((H)eNBs), and the HeNB that cause interference (aggressors). This knowledge, among the others, includes signal strength indicators (RSRPs, reference signal received quality (RSRQ)), uplink (UL) and downlink (DL) operating frequencies and fractional frequency reuse and the exact position of UEs. The added value of IIE is that it examines uplink and downlink, cellular and D2D communication distinctively, in order to indicate aggressor nodes in a fine grained manner.

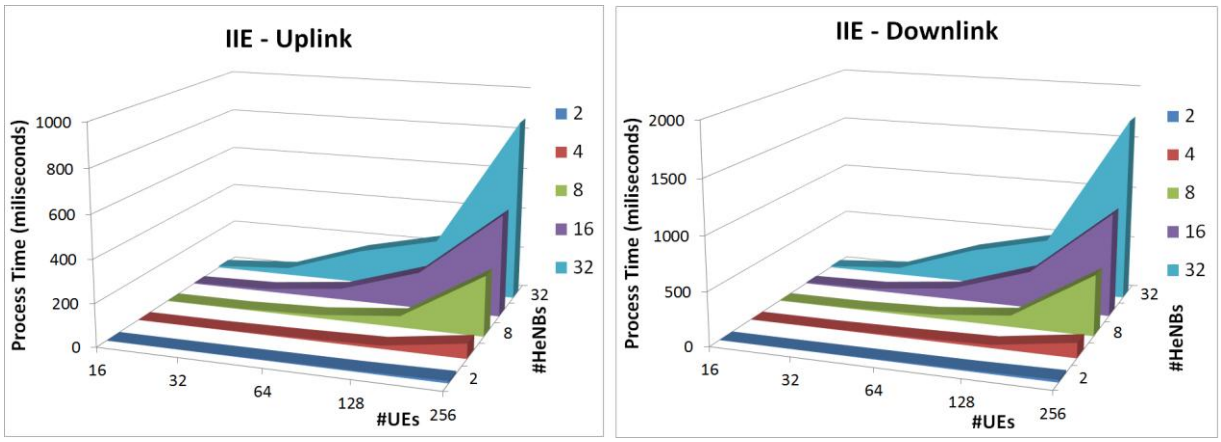


Figure 2-3. Simulation results: Caption (a) IIE Uplink computation overhead (b) IIE Downlink computation overhead

In order to evaluate the performance of IIE, a realistic business case scenario of a shopping mall is examined (in our case comprised of a single floor). We have scaled the number of LTE femtocells deployed in the topology and separated UL and DL cellular communication. Experimental results depicted in Figure 2-3 have shown that IIE computational overhead in downlink is ~x2 higher compared to uplink. In downlink, time delay is affected by the number of (H)eNBs and their associated UEs. In uplink, time delay is affected by the number of (H)eNBs and minor fluctuations appear in case of rise of associated UEs. In case of heavy

load conditions (32 HeNBs and 256 UEs), time complexity does not exceed 1.7 sec approximately [LLK+11].

Existing solutions refer mainly to inter-cell interference management schemes, where eNBs exchange messages between each other in order to communicate interference indicators and coordinate frequency reuse in a distributed manner. Even though the proposed solution of IIE is a centralized one, it is feasible and deployable in current 3GPP standards due to the fact that it exploits indicators already provided in commercial systems.

2.5 Multi-Kernel Based Method for Interference Identification

Considered solution uses a multi-kernel based method to identify the interference pattern (in terms of spatio temporal path loss information) in a wireless communication environment. More precise, we design an adaptive kernel-based algorithm that is able to reconstruct path-loss maps based on path-loss measurements arriving over time. The main advantage of this approach is its low complexity due to its sparsity-awareness, its online adaptivity and the good performance in presence of measurement errors.

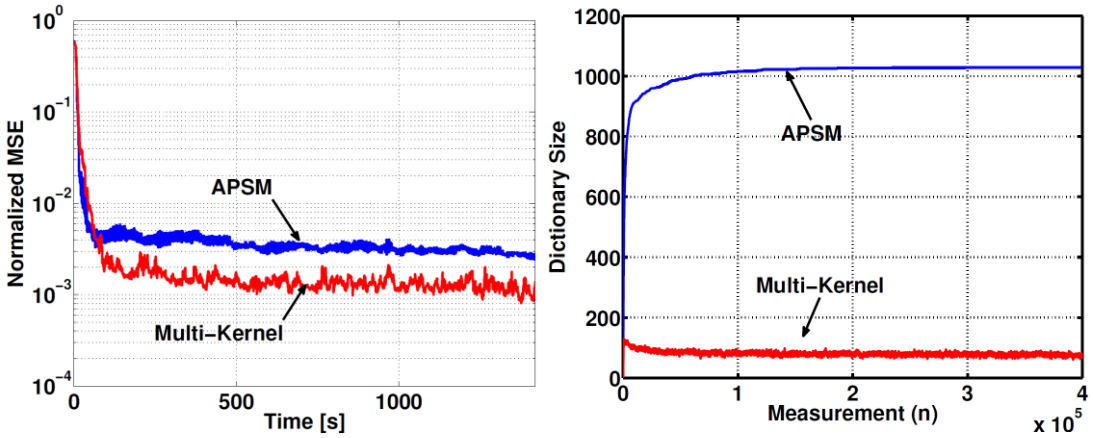


Figure 2-4. Caption (a) comparison of MSE performance (b) comparison of dictionary size

To evaluate the accuracy of the path-loss map estimation we study the evolution of the Mean Square Error (MSE) over time for the proposed multi-kernel based method and the Adaptive Projected Subgradient Method (APSM) from Section 2.2. The MSE is computed based on the difference between the estimated path-loss at pixels and the true value for the path-loss. Figure 2-4 shows a fast decay of the MSE for both methods. The APSM method has better initial performance as it achieves a better MSE in shorter time than the multi-kernel method. But the multi-kernel method outperforms APSM only after about 100s, achieving lower MSE. Furthermore, the multi-kernel method has a much smaller dictionary size (number of measurements used to adjust weights of kernels) compared to the dictionary of APSM (see Figure 2-4). The dictionary size of roughly 10% of the APSM's dictionary has a positive effect on the computational complexity of the multi-kernel method allowing for an online implementation in the real systems.

2.6 Coordinated Fast Uplink and Downlink Resource Allocation in UDN

Densification of network nodes, along with flexible UL/DL TDD air interface proposed by METIS for 5G UDN, pose new challenges for RRM in future network deployments. Above mentioned factors are taken into account in performance assessment of different levels of centralization: from completely decentralized (similar to 4G LTE-A operations of standalone eNBs) to fully centralized approach that could be realized, e.g., using Framework for Control/User Plane Design With Over-The-Air Signalling For UDN (cf. Section 2.8).

Figure 2-5 depicts evaluation results of random dense deployment of small cell access nodes. Radio link packet delay is chosen as a performance metric, because low latency is a desired

5G capability that helps among the others to reduce baseband costs (especially for buffering) and enable e.g. a Tactile Internet [Fet14]. Performance of 4G LTE-A air interface (fixed UL/DL slot allocation, decentralized RRM) is confronted with 5G UDN air interface (flexible UL/DL) of decentralized and centralized RRM.

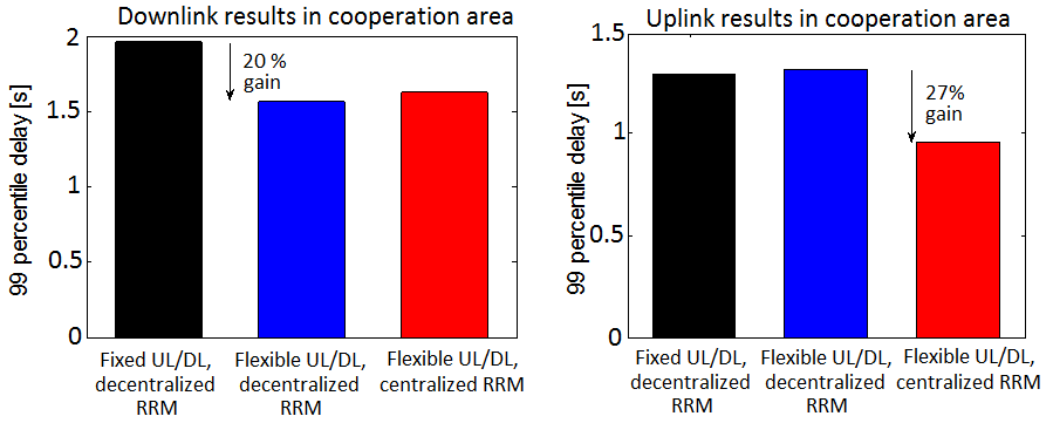


Figure 2-5. Packet delay of 99th percentile cumulative distribution function (CDF) for different modes of RRM operations. Gains of flexible UL/DL air interface over fixed UL/DL split and gains of centralized RRM over decentralized RRM.

Based on the simulation results carried out in indoor deployment of METIS Test Case 2: Dense Urban Information Society [MET13-D11], there are significant performance improvement in case of centralized scheduling (25% and 20% packet delay reduction in UL and DL respectively, comparing to 4G LTE-A operations). Gains are especially visible in a unplanned deployment of small cells. Still, centralized approaches require higher signalling overhead for reporting Channel State Information (CSI) and scheduling decisions to/from central entity. They also introduce more delays caused by scheduling decision procedure.

2.7 Out-of-band Advanced Block Scheduling in Heterogeneous Networks

This technology component proposes a coordinated advanced block scheduling for interference avoidance, using out of band (via TDD) signalling for coordinating Frequency Division Duplex (FDD) communications among the small cells and the macro layer, as well as providing a synchronization reference for the small cells. The LTE-A TDD air interface is used to send subframe scheduling information from the macro stations towards their neighbouring small cells, instead of setting individual X2-like links. Therefore, the macro layer can reach a wider number of small cells, the usage of the radio subframes is more efficient, and the global interference level is reduced. Details about radio resources to be used in the next frames are thus provided by the scheduler of the macro BS over a separate TDD link. Interference between small cells is not addressed (only between macro and small cells).

Traditional frame usage coordination procedures are time-domain based, employing semi-static patterns of unused subframes exchanged through inter-cell interfaces (such as X2 [3GPP12-36423]). These patterns are however infrequently updated, while air interface traffic can be widely variable in short time scales; in addition, signalling through X2 is prone to a significant delay and other impairments. Therefore radio resources coordination between macro cells and small cells cannot be optimally performed.

The advantages of this solution are: enhanced inter-cell coordination; re-use of spare resources by small cells in a low to medium load conditions; good coverage of coordination information if a smaller bandwidth is employed for the TDD link compared to that in FDD; and over-the-air synchronization of small cells with no need to upgrade backhaul lines. The cost would be the need to implement separate TDD transmitters and receivers at the macro cell and small cells, respectively.

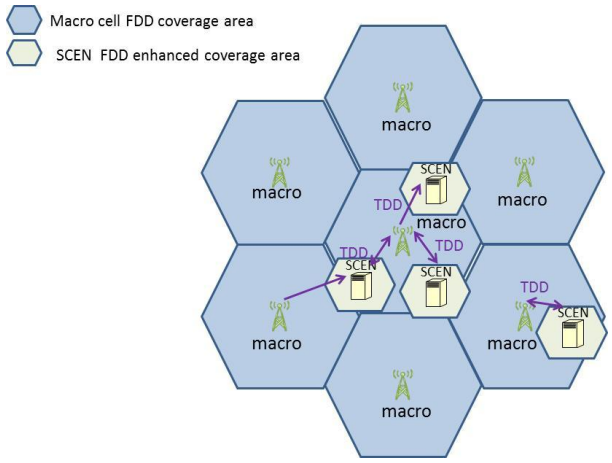


Figure 2-6. Coordinated advanced block scheduling using TDD showing both macro cells and small cell entities (SCEN)

2.8 Framework for Control/User Plane Design with Over-the-Air Signalling for UDN

Proposed technology components using centralized RRM focuses on the optimization of the overall system performance and efficiency through solutions encompassing both cellular transmission (Section 3.5) and network-facilitated D2D (Section. 6.5). Future UDN deployments foreseen for 5G networks are very likely to have backhaul connection of different quality, which may limit performance of some centralized RRM schemes that require reliable signalling exchange between involved nodes. As a remedy, a decoupling of control and user plane is proposed where a central entity (e.g. macro node) use over-the-air signalling to coordinate operations of small cells within its coverage.

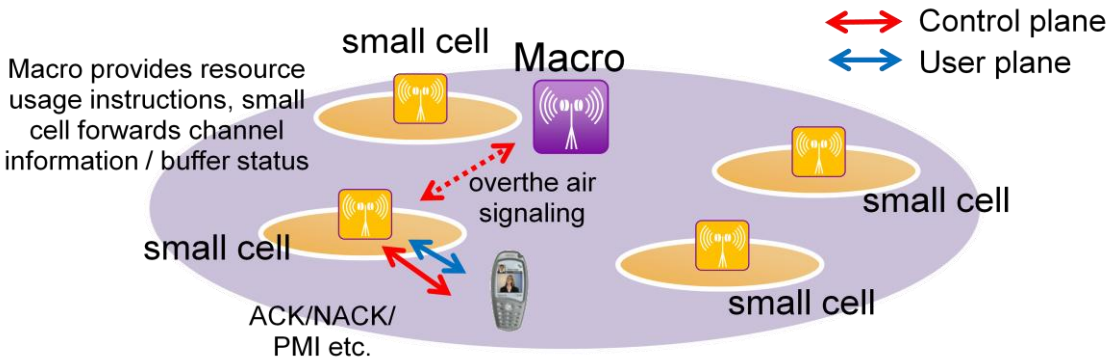


Figure 2-7. Decoupling of control/user plane in HetNet scenario

In the approach depicted in Figure 2-7, UEs are reporting CSI and buffer state information to connected small cell. This information is further forwarded to a central macro station that takes RRM decision based on information gathered from coordinated small cells. Communication between macro and small cell is done using over-the-air signalling that allows providing very low latency and high reliability if the backhaul connection is not capable to do so. Separation of control/user plane in this novel approach can be realized in two variants. In the first one, the control plane is completely handled between macro and UE except physical layer related signalling. The second variant assumes that the control plane towards the UE is handled by small cell, whereas the macro cell provides instructions to small cells on how to use available resources (as shown in Figure 2-7).

2.9 CSI-based Coordination Scheme for Macro or Small Cells with Non-coherent JT CoMP

The low delay spread in small cells gives rise to a large coherence bandwidth, and the question addressed here is how the frequency resources should be allocated to the users prone to interference, keeping in mind that CSI at Transmitter (CSIT) could be a burden in UDN to realize a Joint Transmission Coordinated Multipoint (JT-CoMP). In this regard, we consider a non-coherent JT-CoMP in the downlink with continuous data transmission, where the CSI is available only at the receiver (user). This is an open-loop system without any feedback. We also consider a closed loop system with one-bit feedback due to Hybrid Automatic Repeat Request (HARQ) with incremental redundancy. We investigate the performance in terms of the long term throughput to ascertain what should be the initial transmission rate, r_m, where the users expect continuous data transmission from the BSs (full buffer and equal priority for the users). Also, the performance in terms of outage probability under slow and fast fading channel conditions is evaluated.

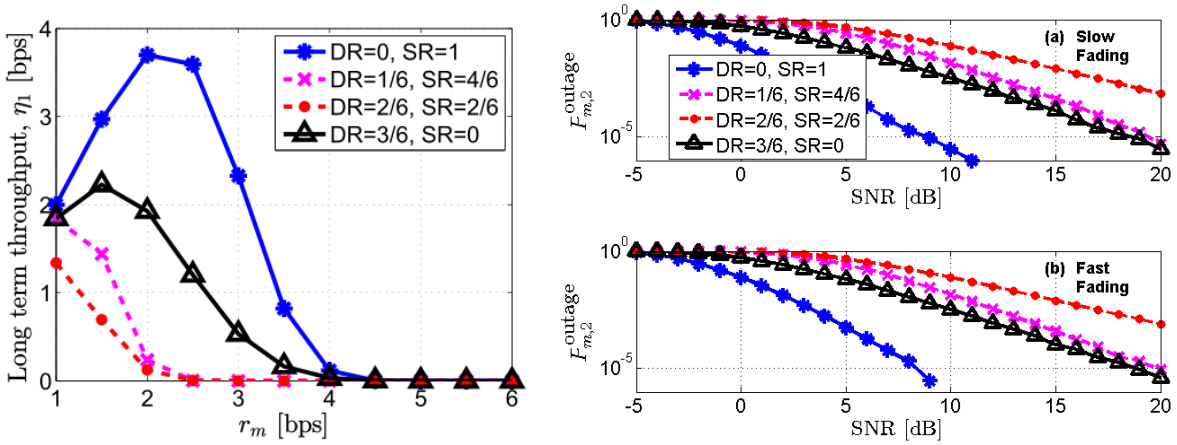


Figure 2-8. (Left) The long term throughput versus initial transmission rate (no retransmission) of the *m*th UE. (Right) Outage probability versus SNR. The ratio in which the dedicated resource (DR) and the shared resource (SR) are distributed for the *m*th UE is captured in the legend. For example, DR = 0, SR = 1, all the frequency resources are shared for the *m*th UE.

In our earlier work [MET14-D42], we had quantified the long term throughput with respect to signal-to-noise ratio (SNR), where the scheduling using shared resources (SR) between users perform better at low SNR, while dedicated resources (DR) perform better at high SNR. Also, the average delay with one-bit feedback is low with a lower initial transmission rate, while having an aggressive transmission rate requires the SNR to be high, so that the average delay can be kept low.

In the results presented in Figure 2-8 (left), for an SNR of 0 dB, the long term throughput is evaluated for various initial transmission rates in an open loop system. It can be observed that sharing the frequency resources is the best option at low SNR, which complements the results presented in [MET14-D42]. Also, in Figure 2-8 (right) for a closed loop system with 1 retransmission (totally 2 transmissions), sharing the frequency resources provides the least outage probability for both slow and fast fading conditions. The fast fading conditions perform better with more diversity.

Further considerations can be found in Section A.1.2.

2.10 Time-Sharing Approach to Interference Mitigation Using Resource Auctioning and Regret-Matching Learning

The objective of this approach is to mitigate the inter-cell interference in a heterogeneous wireless system, with both macro and small cells present. The proposed approach is a distributed one, where each BS performs a long-term resource allocation optimization process separately, however, with the neighbouring BSs exchanging information on their decisions and experienced gains (or interference information). The decisions taken by BSs are in form of so called strategies, which define the use of resources in time and frequency and the maximum allowed transmit power for different parts of available spectrum. Therefore, this interference mitigation mechanism is based on efficient usage of resources in time and frequency by adjustment of BS transmit power.

The optimization process is based on a game theoretic approach where each BS learns the regrets (drawbacks) of use of available strategies and aims to minimize its average regret over time by application of the iterative regret-matching procedure [CHP+09]. The regret values are related to so called payoffs (utilities) achieved with each strategy, which can be defined e.g. based on achieved throughput values. En example of such approach is the Vickrey-Clarke-Groves [CHP+09] resource auctioning, where payoffs depend on achieved throughput, but also on throughput loss of other BSs.

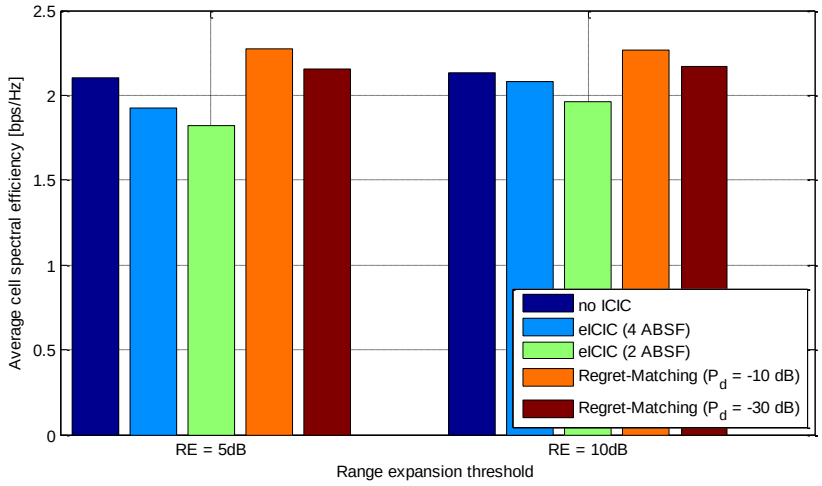


Figure 2-9. Average cell spectral efficiency

The proposed approach has been compared with LTE-A enhanced inter cell interference coordination (eICIC) mechanism based on achieved throughput. Two configurations of the suggested method have been investigated with different values of the transmit power modifier P_d, that indicates the power reduction factor for selected sub-bands for different strategies. Results depicted in Figure 2-9 show, that depending on the used configuration of the proposed approach, the described mechanism provides up to 8% gain in achieved average cell spectral efficiency and up to 40% gain in 5th percentile user throughput comparing to pure LTE-Advanced (without interference mitigation). On the other hand, the achieved 5th percentile user throughput for the proposed approach is slightly lower than for LTE-Advanced eICIC scheme, however, the average cell spectral efficiency is even over 15% higher than in case of LTE-A eICIC. More results are available in Section A.1.3.

2.11 X2-based Distributed Interference Management in Femtocells (DIM-X2)

This approach aims at improving cell-edge users' data rate as well as reducing power consumption in UDN. Coordination messages are exchanged among cells, which allow application of a fractional reuse of resources in a distributed and adaptive fashion. Cells are

grouped into clusters in order to schedule, in each cell of a cluster, certain resources of preferential use over the rest of the cells in the same cluster. Users allocated to these preferential resources must be selected carefully, based on network measurements. Resources used with preference in one cell are blocked in neighbour cells within the cluster, thus reducing not only intra-cluster interference, but also power consumption in the network.

The amount of information required to be exchanged among the cells is reduced and delay-tolerant. Besides, the clustering technique is independent of the interference management scheme and thus can be properly fitted to any network topology.

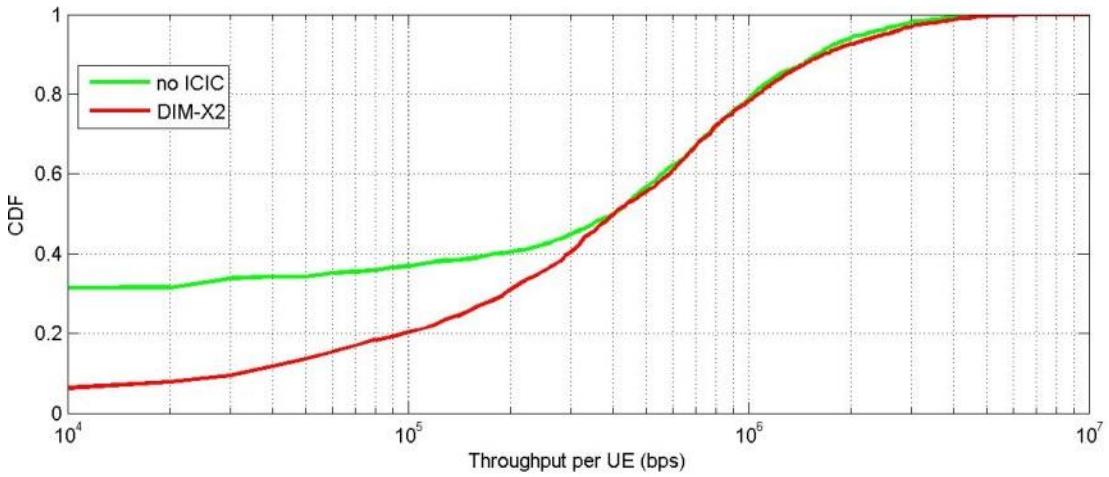


Figure 2-10. CDF of the mean throughput per user. Cell-edge users improve significantly their experienced throughput because they are protected against interference thanks to the use of preferential resources, while cell-centre users are not protected against interference and thus, their experienced throughput is not affected

Figure 2-10 shows how the proposed approach improves cell-edge user's throughput while maintaining the throughput achieved by the rest of users. When applied to a heterogeneous network, total throughput gain is 12% and power consumption is reduced by 10%.

2.12 Self-Organization of Neighbouring Femtocell Clusters

The objective of this technology component is to provide a method to give a preferential access of a terminal in a domestic scenario to a neighbour femtocell operating in Closed Subscriber Group (CSG) mode, when the neighbour signal is higher than the signal from the own femtocell. This solution proposes a neighbour-friendly hybrid access mechanism, where only the devices in the CSG list of the own BS can connect to a neighbour femtocell under control of the network. The complete process is illustrated in Figure 2-11(a). An over-the-air synchronization mechanism is also proposed that leverages on beacon signals broadcasted by synchronized macro cells, for inter-cell interference coordination purposes. Such approach allows avoiding upgrades of the backhaul network. The beacon is transmitted on unused resources, therefore suffering only from thermal noise, and enjoys enhanced detection characteristics.

Current state of the art (SOTA) 3GPP schemes preclude connecting to closed access femtocells, even if their signal is higher than that of the serving node, thus creating a blockage area in shared frequency deployments. For SOTA femtocell synchronization, Network Listening Mode suffers from inter-cell interference caused by the collisions between synchronization signals, and hence poorer coverage than the proposed solution.

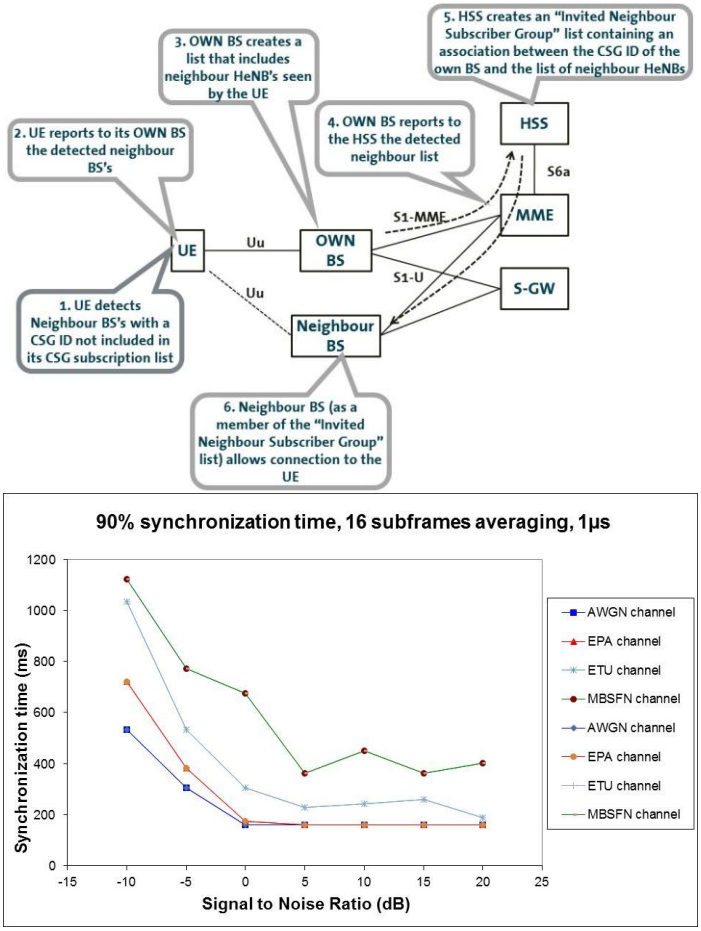


Figure 2-11. (a) Invited neighbour subscriber group list generation, showing the detailed procedure for neighbour-friendly connection to a femtocell;
(b) 90% synchronization time for 16 subframes averaging and 1 µs synchronization

Figure 2-11(b) shows the 90th percentile of the synchronization time for a macro-femto maximum time offset of ±1µs, and it is apparent that typical radio channels require synchronization times below 200 ms with SNR ≥ 20 dB, except multi-broadcast single-frequency network (MBSFN). For ±3 µs the 90% synchronization time is 160 ms (for 16 subframes averaging window) for all radio channels with SNR ≥ 0 dB, except MBSFN. MBSFN channels would require a different beacon detection strategy due to the lower coherence bandwidth that precludes non-coherent detection. SNR distributions of the beacon signal obtained in simulations of a real Madrid city centre for indoor locations, confirm that 80% indoor points enjoy a beacon's SNR higher than 18 dB, and all indoor points are above 10 dB, thus yielding significantly lower synchronization times.

2.13 Coordinated Resource Usage in Virtual Cells

In cellular wireless communication systems mobile terminals (UE) are moving through the coverage area of several cells. Thereby, QoS provided by different cells depends on factors like the instantaneous channel quality between the UE and the BSs or on the induced interference. Hence, with a fixed code rate the QoS at a UE is changing while moving through the cells and symbol retransmissions occur more often. The idea of this approach is to exploit the broadcast property of the wireless channel combined with fountain codes [MK05]. Fountain codes have the characteristic that the receiver needs to receive any N packets sent by the transmitter to decode the original file of size K. Hereby, N needs to be only slightly larger than K and the order of the received packets is not important. The following communication scheme is proposed. In the uplink direction a UE broadcasts fountain encoded

symbols which are potentially received by multiple BSs. These BSs form a virtual cell to jointly collect and decode the symbols. Similarly, in the downlink multiple BSs from the virtual cell cluster transmit different fountain encoded symbols to the UE which is able combine the symbols received from several BSs and encode the full message.

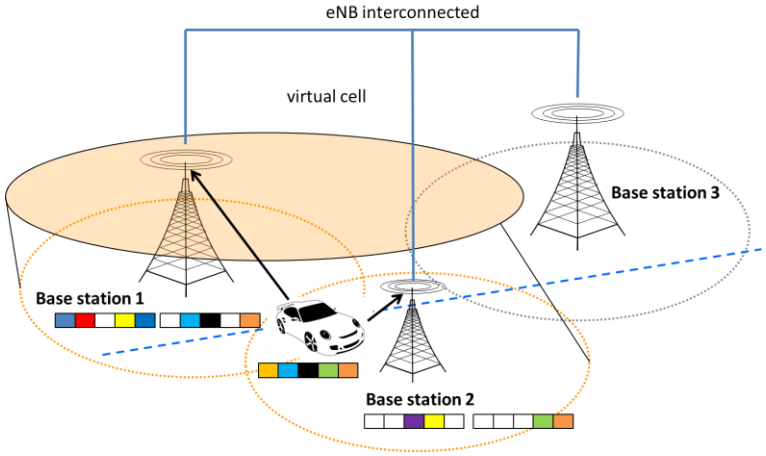


Figure 2-12. Concept of coordinated resource usage in virtual cells using fountain codes

The approach has two major advantages. Firstly, the use of fountain codes removes the necessity of sending acknowledgements (ACKs) for every received encoded symbol and eliminating the need for packet retransmissions. The second improvement originates from the spatial gain present in virtual cells which comes with the reception of encoded symbols at different cells. The concept is illustrated in Figure 2-12 for the uplink case. The mobile user transmits fountain encoded symbols (depicted as coloured blocks) to multiple base stations forming a virtual cell. Once any base station is able to decode the full message (five coloured blocks in the example) it broadcasts an ACK and all other base stations stop decoding.

2.14 Scalable Solution for MMC with SCMA

Sparse code multiple access (SCMA) [MET15-D24] is a non-orthogonal multiple access scheme where coded bits are mapped to multi-dimensional sparse codewords from SCMA codebooks. One application scenario of SCMA is a grant-free uplink contention based multiple access where each SCMA layer represents a user. A layer is spread across the entire time-frequency resources of a contention region. Each layer is characterized by a specific SCMA codebook. The system can be overloaded where the number of multiplexed layers is more than a spreading factor.

The grant-free contention based multiple access scheme with SCMA offers the following advantages: 1) scalability to provide massive connectivity and coverage, 2) low signalling overhead, 3) energy efficiency, 4) low transmission latency, 5) high reliability and 6) moderate detection complexity.

SCMA is scalable to support different levels of connectivity and coverage. This is enabled by changing SCMA parameters such as the sparseness of codewords (number of non-zero elements) and spreading factor. A near-optimal multi-user detection performance can be obtained with a moderate complexity based on message passing algorithm (MPA). Low signaling overhead and latency is achieved by contention-based transmission and blind detection without the dynamic request/grant procedure. An SCMA blind receiver performs joint user activities and signal detection. Overall, this allows the grant-free SCMA mechanism to provide high reliability based on traffic requirements.

The KPI evaluated by system level simulation is the number of devices supported. Results demonstrate that a contention-based SCMA system can support up to 3.95 times more devices than LTE R11 system under 3GPP case 1 scenario at an average packet drop rate of 1% (see Section A.1.8 for detailed assumptions).

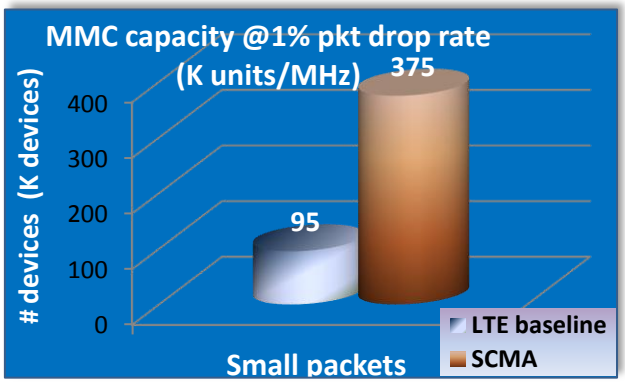
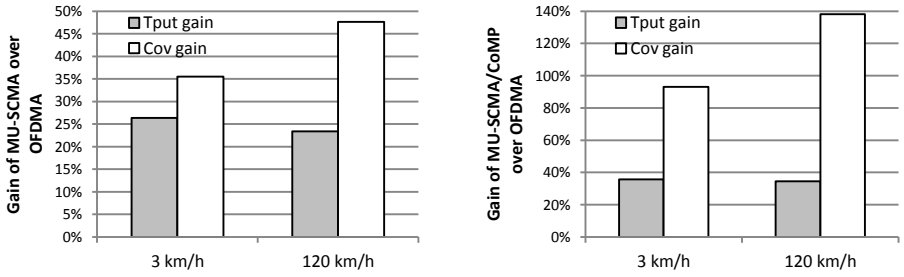


Figure 2-13. Number of devices supported (in thousands)

2.15 Downlink Multi-User SCMA for Mobility-Robust and High-Data Rate Moving Network Mobility (MN-M)

The objective of this technology component is to provide a mechanism to increase the capacity and availability of a network even at the presence of high speed moving users. SCMA is a new modulation and a non-orthogonal multiple access scheme in which coded bits are mapped to multi-dimensional sparse codewords. Codewords of multiple layers or users are overlaid in code and power domains and carried over shared OFDMA time-frequency resources. In multi-user SCMA (MU-SCMA), users are multiplexed over SCMA layers with the minimum channel knowledge to increase the overall throughput and coverage of the network with a low sensitivity to user mobility and variation of fading channel over time. In addition, interference is mitigated through inter-cell SCMA layer coordination, power sharing, and an efficient multi-user detection at terminals using the MPA with a near optimal performance and moderate complexity.



(a) (b)

Figure 2-14. System-level simulation results, (a) gain of MU-SCMA over OFDMA, (b) gain of MU-SCMA plus multi-cell coordination (CoMP) over OFDMA

User multiplexing improves the overall throughput of a downlink wireless network. Existing solutions exploit spatial domain to multiplex users either over antennas of one or multiple cells (CoMP). In these solutions, user multiplexing is enabled by multi-user MIMO (MU-MIMO) or joint transmit MU-MIMO CoMP. Despite the large potential performance gain, these techniques are only applicable to low mobility scenarios due to their sensitivity to channel aging. Even for low mobility users, the sensitivity to quantization error of channel state information (CSI) and the cost of feedback overhead limit the practical gain of these schemes. The proposed DL MU-SCMA scheme is an open-loop multiplexing scheme where different code domain layers are assigned to different users without the need of full CSI knowledge of the co-paired users. Compared to MU-MIMO, this system is more robust against dynamic channel variations in high speed scenarios. The proposed scheme benefits from low feedback overhead and less sensitivity to channel variations to support high speed users with improved capacity and coverage gain. Furthermore, SCMA supports CoMP for seamless handover for

high speed users with reduced end-to-end delay and load balancing through inter-transmit point (TP) layer/power sharing.

The system-level simulation results of Figure 2-14 show the relative cell average throughput and user 5% tile coverage gain (cell edge throughput) of MU-SCMA multiplexing and MU-SCMA plus CoMP over OFDMA. According to these results, SCMA multiplexing improves the average cell throughput and cell-edge throughput. Furthermore, the relative gain of MU-SCMA over OFDMA maintains for high speed scenario where the closed-loop multiple access schemes such as MU-MIMO fail due to the channel aging effect and high rate of CSI feedback signaling. Therefore, SCMA multiplexing and CoMP techniques can provide high data rate and mobility-robust services for moving networks (MN).

Further details can be found in Section A.1.9.

2.16 Interference Management for MNs in Ultra-Dense Urban Scenarios

In previous METIS studies, the benefits have been demonstrated of using half-duplex MRNs to serve VUE inside well-insulated vehicles in noise limited scenarios and scenarios with limited co-channel interference. In later studies in METIS, the concept of MRNs is extended to full-duplex MNs deployed on vehicles. Both MRNs and MNs communicate with the macro BSs via the air interface, and form their own cells inside public transportation vehicles to serve the VUEs on board. However, the backhaul links and access links of MRNs are assumed to operate in the same frequency band, and therefore the MRNs can only work in a half-duplex fashion. In the vision of METIS, in 5G systems more bandwidth can be freed up for small cell deployment. Thus, the use of full-duplex MNs with the backhaul and access links on different frequency bands becomes feasible in the 5G systems. In order to further understand the impacts of deploying MNs in densely deployed scenarios, further investigations have been conducted. In [MET14-D63], the preliminary results of using full-duplex MNs identified that the problem that limits the performance at the VUEs in a densely deployed urban scenarios is not only the Vehicular Penetration Loss (VPL), but also the complicated inter-cell interference. Therefore, in this contribution, we have investigated and applied different interference management schemes for both macro and micro cells to alleviate the problem caused by inter-cell interference. It is shown that by deploying a new type of node, i.e., the MNs, can help to improve the performance at the VUE significantly, compared to serving the VUE directly from the BS. The impact on the regular outdoor UEs are very limited.

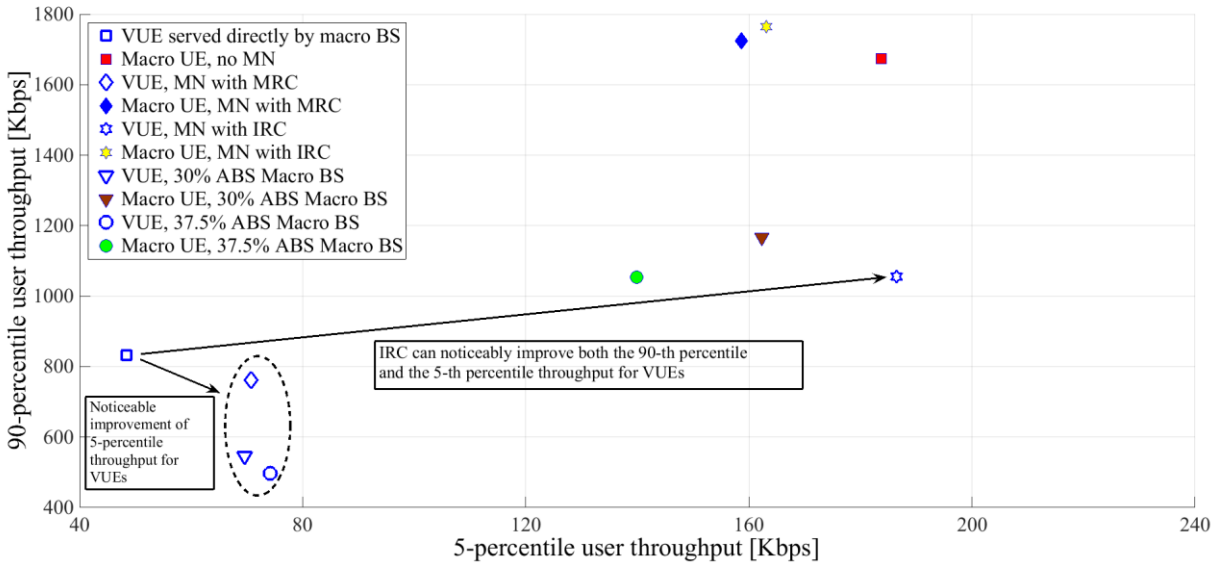


Figure 2-15. The 5-percentile vs. 90-percentile throughput of VUE and macro UE when the VPL is at 30 dB

Figure 2-15 shows the 5-percentile versus 90-percentile throughput of VUE and macro UE by using different interference coordination or interference cancellation schemes. As we can see from the figure, if we use interference rejection combining (IRC) at the backhaul links of the MNs, it significantly improves the performance at the VUEs. Other schemes can improve the 5-percentile VUE throughput, either by to improving the desired signal strength (e.g., by using maximum ratio combining (MRC) at the backhaul of the MNs), or reduced the interference (e.g., to configure almost blank subframes (ABSs) at the macro cells). Furthermore, from the results we can see that for the macro UEs, there is no significant performance change if IRC or MRC is used by the MNs; however, as the use of ABSs consumes resources at the macro cells, this scheme will have some impact on the macro UE performance. The corresponding performance for the micro UEs are given in Section A.1.7.

We have demonstrated that, by the use of MNs to serve vehicular UEs in public transportation vehicles, the QoE of the vehicular UEs can be improved, and the impact on regular outdoor UEs are very limited. Thus, there is a clear system gain by deploying MNs also in an interference limited ultra-dense urban scenario. Therefore, the use of MNs is a promising solution in the METIS system to better meet the needs for vehicular users [SGS15].

2.17 Support of Low-Cost MMC Devices Under Poor Coverage and High Frequency Offset Conditions

Low-cost MMC devices can suffer from both coverage and frequency stability impairments. The presence of high carrier frequency offsets (CFO) in low-quality local oscillators, combined with the use of low transmit powers for reduced battery consumption, can cause significant issues in bad coverage scenarios caused by poor reception. We propose to reserve a subset of MBSFN subframes for dedicated MMC traffic, in such a way that cells are grouped into MMC coordination areas where resources are shared and MMC traffic is broadcast to (and received from) MMC devices under control of an MMC controller node. Both UL and DL coverage can be enhanced and resource management of MMC traffic can be made independent from the other types of traffic, and tailored to actual MMC traffic needs. UL coverage and CFO support in discrete Fourier transform (DFT)-spread-OFDM (DFT-s-OFDM) modulation can additionally be improved for MMC by applying a repetition pattern prior to the DFT that boosts the spectral density of the subcarriers. The solution improves coverage without compromising frequency diversity, as would happen with standard bandwidth reduction schemes in scenarios with low channel coherence bandwidth (especially indoors). The maximum supported CFO is also increased by a factor equal to the greatest common divisor of the repetition factor and the symbol length.

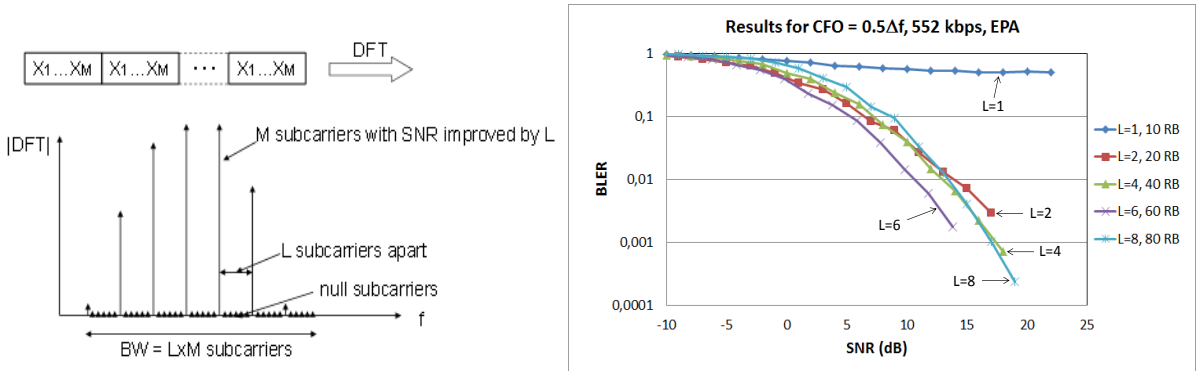


Figure 2-16. Improved operation for DFT-s-OFDM: (a) Baseband repetition pattern prior to the DFT; (b) BLER vs. SNR simulation results for CFO=0.5Δ*f*, showing the improvements when increasing the repetition factor L and hence the occupied bandwidth (in RBs)

Current SOTA 3GPP solutions for MMC improvements rely on simple packet retransmissions and bandwidth reductions at the cost of reduced frequency diversity, not allowing CFO values equal to or above than half the subcarrier width Δf. Simulation results for LTE with the proposed scheme (Figure 2-16) show how it can significantly improve the block error rate (BLER) characteristics in the UL even under large CFO conditions. The allowed repetition factors can be reported to the BS upon initial access by means of dedicated control signalling, and actual values can be selected according to the device capabilities.

2.18 Impact on the METIS 5G System Concept

Investigated technology components address all fundamental METIS services. Flexible air interface based on dynamic UL/DL TDD proposed by METIS in [MET15-D24] was proven to efficiently cater for dynamic traffic demands, even in a non-planned UDN deployment. Decoupling of control and user plane seems an interesting option to pursue. Additionally to bringing the benefits of improved mobility robustness [IKT12] and fostering fast dynamic activation and deactivation of access nodes (cf. Section 6.3), it may also enable efficient centralized RRM mechanisms (e.g. macro could play a role of coordinator for small cells within its range). In order to provide a reliable signalling connection for such a scheme, signalling over the air could be used. If centralization is not possible, a number of distributed RRM mechanisms proposed in this chapter could be used instead.

Efficient RRM and proactive interference management mechanism presented in this chapter focus mostly on xMBB services. Solutions achieving better performance in comparison to SOTA approaches were proposed, and although in stable traffic conditions and regular cell distribution their gains are limited, they are the enablers for unplanned cell deployments and are able to provide a control over highly heterogeneous environment.

Apart from xMBB solutions some mechanisms for M-MTC (cf. Section 2.14) and U-MTC (cf. Section 2.17) are also presented, based both on OFDMA and SCMA.

A new air interface modulation based on SCMA is an interesting alternative for OFDM, especially for provisioning of high data rate connection to the fast moving vehicles, when MIMO closed loop schemes fail due to channel aging effects [YYK+14]. QoE of users in fast moving vehicles can be also enhanced by utilization of MRNs as discussed in Section 2.16.

A particular aspect of future networks that still requires further research is mechanisms in mmW deployment, that were not investigated in this study.

2.19 Addressed METIS Goals

Considered mechanisms allow better utilization of available radio resources. METIS objectives described in this chapter focus on reaching 1000x higher traffic volume 10-100x higher data rates and 10-100x higher number of connected devices at a reasonable cost.

2.20 Conclusions

Several interference detection and estimation methods were proposed, that could be used in the 5G systems. Out of them, Kernel Based approach has the highest potential with comparable complexity as other schemes.

In order to reach the highest performance, especially in a unplanned ultra dense deployment operating in flexible UL/DL TDD mode, centralization of RRM is needed (cf. packet delay statistics in Section 2.6). Gains of centralization are mostly visible for users exposed to the most disadvantageous interference conditions (e.g. caused by a mismatch of the TDD slot allocation in neighbouring cells). One proposal for more centralized resource assignment is macro playing a coordinator role for the small cells that are within its coverage. Signalling over the air between access nodes can be used for the exchange of latency-critical control information, if the backhaul infrastructure cannot guarantee this latency. The macro air

interface can be also used to provide synchronization signals for small cells within the coverage of a macro for inter-cell coordination purposes. Such a solution creates the possibility to alleviate the requirements for the backhaul and as a result it may reduce its cost. Heterogeneous deployments naturally offer the possibility to coordinate resources by broadcasting macro scheduling decisions in an implicitly master role, so that small cells can avoid interference by scheduling resources not already used by the macros.

For the inter-UDN coordinated spectrum sharing RRM schemes readers are referred to [MET14-D53].

Apart from above mentioned centralized solutions, several options for distributed and fully decentralized RRM mechanisms were proved to improve SOTA solutions. An improvement in cell-edge user's throughput and a significant reduction in power consumption can be achieved using a simple coordination scheme based on the exchange of low load messages among nodes. Neighbour-friendly cell camping in UDN (cf. Section 2.12), together with over the air synchronization of small cells through enhanced beacons, can enable distributed ICIC algorithms based on the clusters naturally formed by the neighbourhood relationships.

Particularly problematic interference constellations can be experienced when ultra dense networks are complemented with moving relay nodes. Moving relay nodes are typically serving users in vehicles, who are not only limited with vehicular penetration losses, but to a large extent also affected by harsh interference environment. Hence one of the most promising remedies are also interference rejection combining (IRC) based receivers [TBM+13]. Other studies in the METIS project investigate feasibility of changing the mode of operation from a moving relay node to a nomadic base station, depending on the location of the vehicle. Such approach allows trading of in-vehicle coverage for coverage/capacity extension for users outdoors [MET14-D53]. For the vehicles or users that are actually on the move, two novel approaches were proved to be useful. The first one is based on the utilization of fountain codes and clustering of virtual cells. Such approach allows us to exploit broadcast properties of wireless channels. The second approach uses the multiple access technique of SCMA which allows high data rate open loop transmission for users that are moving faster than 100 kmph. Transmissions coded with SCMA are decoded using MPA receivers and this approach can be also used to provide access for a large numbers of machine type devices. In this scenario, a grant-free access is enabled by multi-dimensional SCMA codewords spread across all time-frequency resources. SCMA with the grant-free contention based multiple access is able to provide scalability for massive connectivity and coverage, low latency and significantly reduced signalling overhead for energy efficiency, as well as the high reliability and moderate detection complexity.

Operations for MMC can be also enhanced by DFT-spread-OFDM modulation which increases maximum supported CFO and improve transmission for low cost devices. Coverage enhancements in low-cost MMC usually require bandwidth reductions, but this comes at a loss in frequency selectivity that impairs performance. A simple bandwidth increase that result from changes in DFT-s-OFDM can boost performance and allow higher CFO values compared to SOTA [Lor15].

3 Mobility management and robustness enhancements

3.1 Introduction

One of the biggest research questions that need to be answered on a network-level is the mobility management approach for the beyond 2020 cellular systems. The need for novel mobility schemes is driven by network densification or novel transmission modes such as direct D2D, but also related to higher frequency regimes. This last factor may significantly affect the contemporary mobility management schemes as shadowing effects in mmW range are much more severe comparing to the centimetre waves. In consequence handovers may occur more often and the cell borders may be much sharper than what we experience in modern networks. The combination of mmW and techniques like massive MIMO may additionally challenge procedures for mmW cells detection as the search for narrow-beamed signals may be particularly cumbersome for UEs. The increased heterogeneity of the network, already described in previous chapters, can also alter the approach toward mobility management [GJ03], [FET03], [CMK+08], [BHF+09], [FK12]

All of these aspects are used as a starting point of mobility related METIS investigations. Considered solutions range from smart and device-aided approaches, to sophisticated network-based mechanisms that could be used in the future for self organizing networks (SONs).The sections below describe network-level mobility management enhancements that were proposed in the METIS project. It needs to be noted that there is a large group of solutions that improves mobility management using context information. These solutions (among the others) are described in Chapter 4.

3.2 Resource and Power Efficient Service to Layer Mapping and Connectivity in UDN

Currently, the UE is unaware of the capacity available in the network, while the network is unaware of the total amount of data that UE would transmit or receive, on a longer timescale. To address this issue, in this technology component, we propose introducing a middleware in the UE that is responsible for initiating non-urgent transfers of a large amount of data (such as software updates, backups, photo uploads etc.). The middleware needs to be aware of the currently available and predicted future connection opportunities, as well as the characteristics and QoS requirements of the user traffic, to know when to best initiate a task. For example, user device operating system could indicate the middleware that it is having 500 MB software update that it would need to download within the next 2 hours. If UE is moving and connected to the macro cell layer, the middleware could decide to postpone the download and wait for a better connection opportunity (e.g. small cell) and that the UE becomes stationary.

The connectivity management framework for ultra-dense deployments is as shown in Figure 3-1.(a). Here user preferences, service requirements and device status is taken into account to create service mapping with certain applications using the macro cell for its connectivity requirements and other (low-priority) applications using only small cells for connectivity. There could be three approaches for this – (a) purely UE autonomous action based on prediction whether a better connection opportunity will become available in time, (b) with information exchange between the UE and the network, or (c) the network assists the UE by indicating when more capacity is available.

In LTE-A networks, such long term or even short time scale delaying of traffic scheduling is not possible. When there is downlink data awaiting for delivery at the UE, either new bearers are setup or UE requests creating of new bearers for sending or receiving the data. Thus, the baseline assumption here is that currently deployed networks do not have service to layer mapping capabilities.

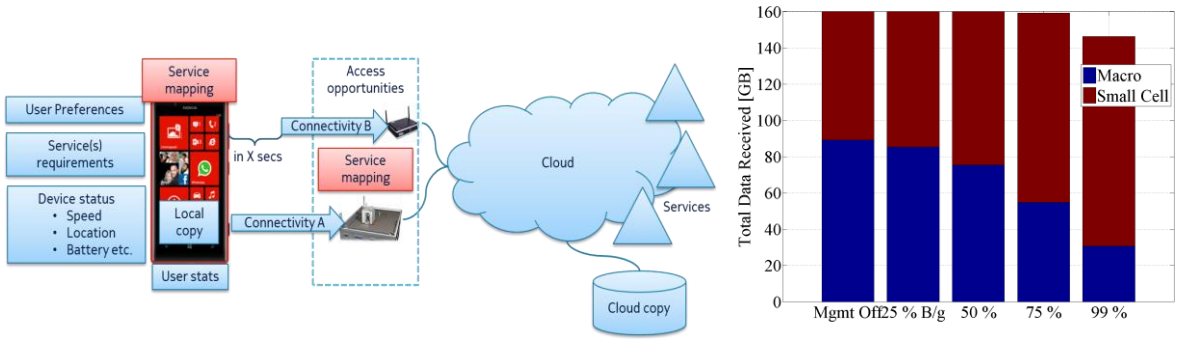


Figure 3-1. (a) UDN connectivity management framework (b) Total downlink receive data from macro and small cells, with and without network management

Evaluation of this technology component in comparison without the network management which is considered as the current baseline is as shown in Figure 3-1.(b), with different percentages of the background low-priority traffic (B/g) as compared to the total traffic in the network. The evaluations were done in METIS test case 2 [MET13-D11], and from the results we can observe that significant data offloading from the macro cell layer can be achieved by deprioritizing such traffic to be scheduled only through the small cell layer, and the additional capacity can be used for delivering other high priority traffic. More details of the simulations is mentioned in the Section A.2.1, further results can be seen in [YWL+14] and [MET14-D42].

3.3 User Anchored Multi-RAT Self-managed Load

The objective of this technology component is to provide a device-centric approach where the UE acts as an anchor node for load balancing in both idle and connected modes. This technique is especially suitable for multi-layer and multi-RAT networks, in which BSs broadcast cell load indicators through an appropriate broadcast control channel. These indications can then be decoded by UEs as part of their normal measurement procedures or estimated via signal analysis, and taken into account in both idle and connected modes. Devices in idle mode can exploit the cell load indications of multiple neighbour cells as additional inputs to cell selection and reselection algorithms (both intra-frequency and inter-frequency), without intervention from the network. UEs in connected mode can also exploit the cell load indications by reporting to the serving BS the neighbour cells' load indications for handover decisions, in such a way that the candidate cells with excessive cell loads are disregarded whenever possible.

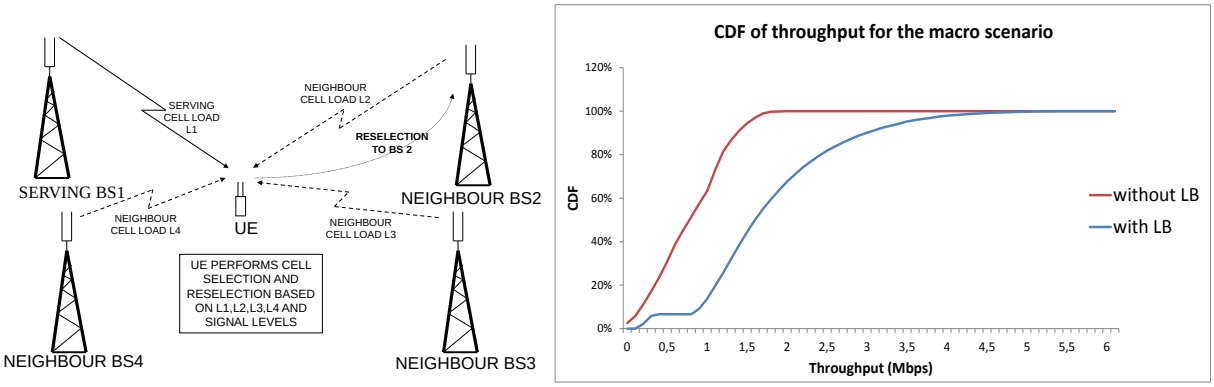


Figure 3-2. Simulation results: (a) Simplified operation in idle mode; (b) CDF of cell throughput in connected mode for the macro scenario showing the improvement of user-anchored load balancing

In contrast with SOTA 3GPP solutions, where a complex control signalling involves multiple core network nodes, there is no need to upgrade any network interfaces in this solution, thus fitting in mixed-RAT scenarios. While network upgrades can be extremely costly when they

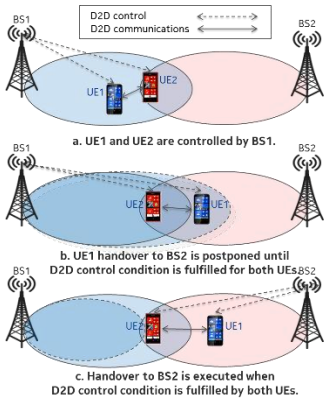
involve interoperability between different RATs and/or manufacturers, changes in devices can benefit from the users' habit of renewing handset in few years' time (or even less than a year). This can be exploited by device makers so as to upgrade their processing and incorporate such load balancing decisions. Simulation results for LTE in connected mode (Figure 3-2) show how the proposed UE-centric load balancing mechanism can significantly enhance the throughput CDF with minimum network upgrades. Average throughput improvements of 229% in macro scenarios and 210% in macro/micro scenarios can be obtained in realistic co-channel Madrid layouts under random user allocation. Cell-edge performance (10th percentile) is also improved by 560% in purely macro scenarios and 334% in macro/micro scenarios.

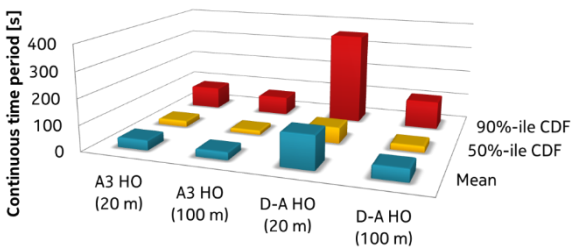
3.4 D2D Handover Schemes for Mobility Management

D2D provides a low power, high data rate, and low latency communications between end-users compared to the cellular services in which the network infrastructure is involved. However, reliable low latency communication becomes challenging in case of mobility. For instance, at the cell-edge, D2D UEs may often be handed over to a neighbouring cell BS at different times. This can make it harder to maintain the ongoing delay-sensitive D2D services if more than one BS is involved in D2D control unless the BSs have an ideal backhaul connection.

To address these problems we propose two complementary solutions. D2D-aware handover management (illustrated in Figure 3-3(a)), where the HO of UE1 is delayed until also UE2 can be handed over to BS2. This delaying of the HO (or making it earlier) to achieve synchronized HO for the D2D pair is done provided that the link quality allows this (i.e. neither of the UEs should experience a connection drop).

D2D-triggered HO aims as well at keeping a group of D2D UEs under the same BS. Therefore, when there is a new device joining a D2D group, it is controlled by the same cell or BS already controlling the D2D UEs of the group.





	A3 HO (20 m)	A3 HO (100 m)	D-A HO (20 m)	D-A HO (100 m)
Mean	38.2	29.5	128.1	47.9
50%-ile CDF	23.8	16.1	61.7	24
90%-ile CDF	86.7	72.1	356.3	109.8

Figure 3-3. (a) D2D-aware handover (b) Simulation results showing the benefit in terms of increased continuous connected time of a D2D pair in the same cell

The simulation results in Figure 3-3 (b) show that the continuous time of stay under a same cell (this cell can change, but both D2D UEs are in the same cell) can be increased significantly with D2D-aware handover (D-A HO in the figure) compared to independent HO for each UE (A3 HO, i.e. is done individually for each UE based on A3 event [3GPP14-36331]). This gain is 235% and 62% for 20 m and 100 m maximum distance between the D2D pair UEs, respectively. This can be done without compromising the reliability of the HO. This is because the HO is done regardless of the D2D pair UE when the link starts to get too weak.

This solution reduces the network signalling overhead and improves the D2D E2E latency by maximizing the time period when D2D UEs are under the control of the same cell [YLV+14].

3.5 Small Cell Mobility Enhancements in Multicarrier and mmW Small Cell Network - Small Cell Discovery in mmW Small Cell Networks

The propagation characteristics of frequencies above 6 GHz bring additional challenge to small cell discovery. This is most notably because of heavier signal deterioration in case of obstructed environment. As a result, mmW communication is mostly available when there is a short-range line of sight (LOS) link between the UE and BS.

We consider a scenario illustrated in Figure 3.4, where the mmW band BSs are serving also a low-frequency carrier (below 6 GHz). We propose detecting whether there is a LOS link between UE and BS based on measurements of the low frequency signal and controlling the measurements on mmW band based on this. Detection of whether the link is LOS is based on received signal characteristics (channel impulse response, root mean square delay spread, received signal strength indicators) on the frequency carrier where the UE is primarily connected. At BS the aggregated statistics on received signal characteristics can be used for switching on/off the mmW transceiver of the BS.

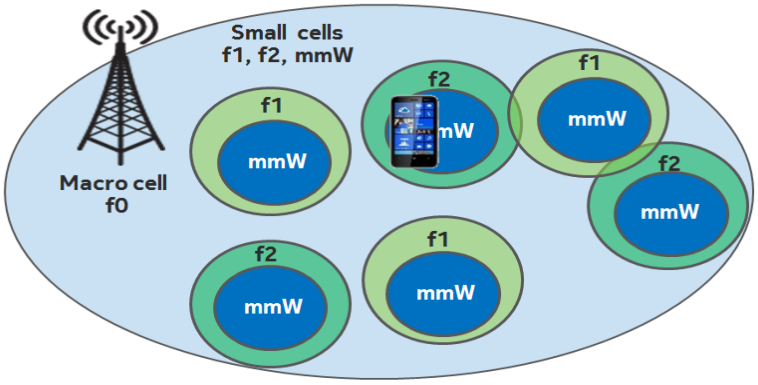


Figure 3-4. Small cell discovery in mmW networks

The current releases of LTE-A do not support mmW, but the baseline assumption would be likely that the UE would need to directly measure carriers on mmW band to find cells there. Compared to that, the main benefits of this solution are in terms of improved energy efficiency. The UE saves energy as it can avoid some unnecessary measurements of the mmW band (or can measure other carriers more frequently with the same power consumption). The network can save energy by powering on the mmW transceiver only when there are UEs within range.

3.6 Small Cell Mobility Enhancements in Multicarrier and mmW Small Cell network - Network Assisted Small Cell Discovery

Finding small cells efficiently is challenging in a dense network deployment. This is especially the case when the small cells are deployed on multiple carriers, as autonomous searching by the UE would mean frequent measurements on multiple carriers. There is a trade-off between the UE power efficiency (how much energy is consumed by measurements) and the delay in finding the small cells.

We propose a solution where UE obtains assistance information from the network to target its search of small cells. This information consists of radio fingerprints of the coverage carrier (e.g., RSRP of neighbouring macro cells) that correspond to a small cell location or a HO region as illustrated in Figure 3-5 (a). As part of the normal operation on macro cell carrier, UE performs neighbour cell measurements and compares the obtained neighbour cells RSRP measurements to the fingerprints received from the network. When there is a fingerprint match, UE reports this to the network indicating which fingerprint matched. In response the network can configure targeted measurements (on a specific carrier) for finding the small cell.

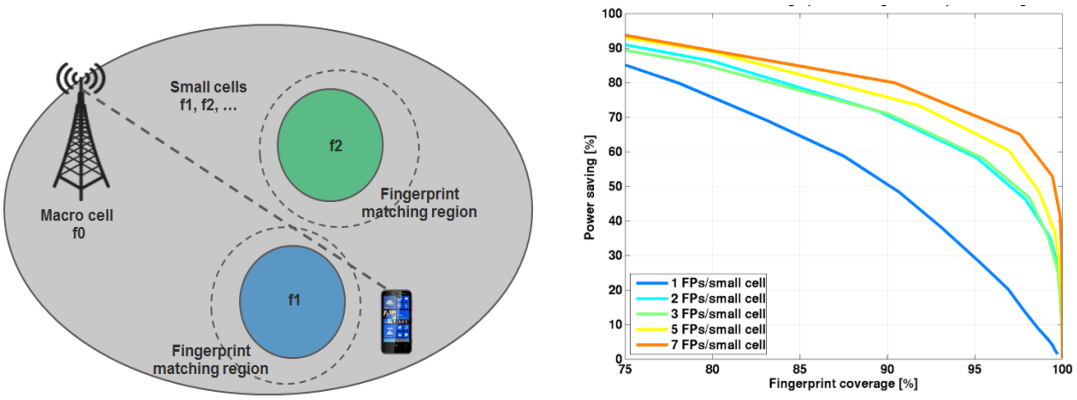


Figure 3-5. (a) Network assisted small cell discovery (b) Simulation results showing trade-off between inter-frequency measurement power saving and fingerprint (FP) coverage (percentage of area with a fingerprint match out of total small cell coverage)

In a deployment as in Figure 3-5 (a), LTE-A UE would need to be configured to search for inter-frequency small cells continuously, to avoid missing any offloading opportunities. This would cause a significant power consumption burden for a UE (in the order of 10% additional power consumption (see Section A.2.2 for more details), relatively even more if discontinuous reception (DRX) is configured). As seen from Figure 3-5(b) The network assisted small cell discovery based on macro cell radio fingerprints can reduce this by 60-70% with a reasonable initial signalling overhead (sending the fingerprint data to the UE). Additionally, if there are small cells deployed on several different carriers, employing the present solution, UE measurements can be directed towards the correct carrier and thus the latency of offloading can be reduced. More details can be found in Section A.2.2.

3.7 UE vs Network Driven Handover

From the point of view of 5G system design it is important to check two handover design options: driven by UE or by network infrastructure. Performed analysis is based on LTE-A numbering for dense deployment scenarios and evaluation results are captured in Table 3-1

Table 3-1. Analysis of handover interruption time for UE autonomous and Network controlled handovers

KPI	Network/ inter-site	UE/ inter-site
Interruption time UL	1.25 ms	6.5 ms*
Interruption DL	4.5 ms	6.0 ms*
HO reaction time**	11.25 ms	0.5 ms
Packet forwarding time	11.75 ms	15.75 ms*
Total handover time	27.5 ms	21.25 ms
Max extra packet latency	5.0 ms***	6.5 ms*

* = impacted by long BSI reading time, ** = time between detection of channel degradation and receiving a HO command/sending “Bye” message. Network HO impacted by long HO decision making time, *** = equal to X2 latency

For the network controlled handover smaller interruption time is expected for active users (i.e. with data in the buffer). Additionally this scheme fits better to the cloud architecture, allows

easier enforcing of network policy, admission control before handover and contention-free RACH. To minimize the handover interruption time we suggest RACH-less handover and avoiding system information reading.
UE autonomous handover allows for a shorter handover reaction time, i.e., imply less handover failures due to long measurements and X2-like transmissions, enables more frequent consecutive handover for high mobility UEs. To minimize the interruption time we suggest faster information reading and 1-step RACH.

3.8 Handover Optimization for Moving Relay Nodes

This contribution shows that by deploying a new type of node, i.e., the moving relay nodes (MRN), performance of the vehicular user equipment (VUE) can be improved significantly, when compared to serving the VUE directly from the BS. This investigation focuses on the mobility aspect of MRNs deployed into the existing system. The power outage probability (OP) performance of a VUE inside a well-isolated vehicle is studied. The communication with VUE is affected by the vehicular penetration loss (VPL) and a limited amount of co-channel interference. A generalized framework to optimize the handover parameters of the system that minimize the end-to-end OP of the VUE is proposed. Using a practical propagation channel we compare the performance at the VUE served by a half-duplex MRN with the baseline direct single-hop BS to VUE transmission, as well as a half-duplex fixed relay node (FRN) assisted transmission, where the position and HO parameters are optimized to minimize the average end-to-end OP at the VUE.

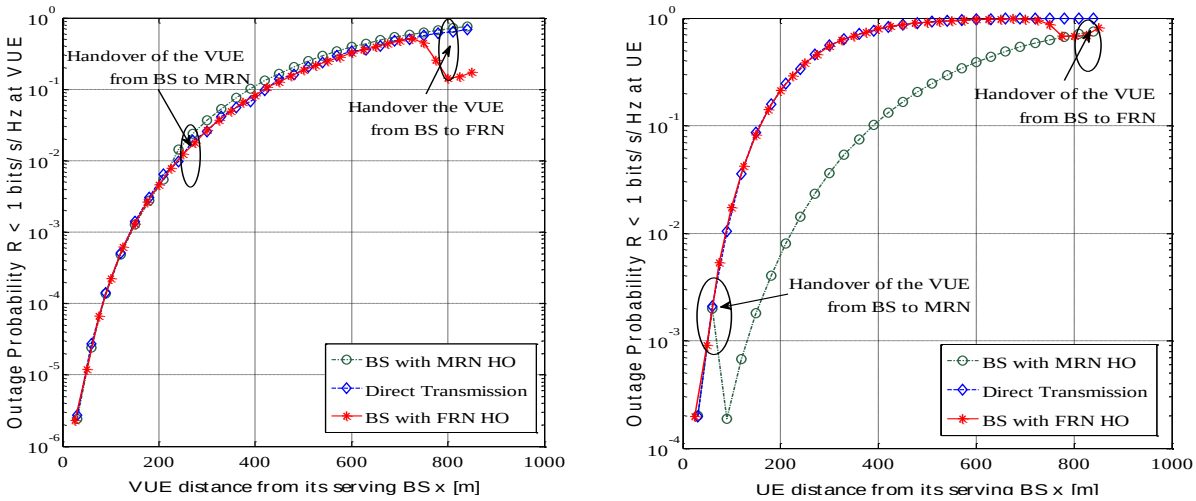


Figure 3-6. Outage probability at the VUE when a) VPL = 10 dB and b) VPL = 30 dB

Performance assessments are shown in Figure 3-6. A minimum rate requirement R is set at the VUE. If the received rate is less than R, an OP is declared. The detailed simulation setup is given in Annex A3.1. As we can see from the results, if the VPL is low, i.e., VPL is at 10 dB, the performance of MRN assisted transmission is almost as good as the direct hop BS-to-VUE transmission. The advantage of using MRN reveals itself when the VPL increases. At a 30 dB VPL, there is a clear advantage to serve the VUE by using the MRN assisted transmission, as when the VUE is handed over to the MRN, the OP at the VUE is significantly lowered. Furthermore, using the FRN to serve the VUE only, helps to lower the OP at the VUE when the vehicle is close to the FRN at low VPL.

3.9 Impact on the METIS 5G System Concept

Described mechanisms focus on xMBB solutions, although mentioned D2D mobility concept can be mapped also to U-MTC. It was shown, that in order to fully exploit heterogeneous network environment, more information is needed (e.g., cell load information, radio fingerprints, etc.) for handling measurements, handover or scheduling decisions for users on

the move. A prime example here is the case where a delay tolerant traffic (e.g. software update) is deliberately postponed, until an unloaded cell (or high capacity cell) is reached. A method exploiting low (sub 6 GHz) measurements for triggering the mmW cells search is also proposed.

3.10 Addressed METIS Goals

Technology components described in this chapter are not directly addressing METIS goals. Still, they are a potentially vital component of future deployments since they allow for an even distribution of the traffic in the heterogeneous networks and efficient attachment to the most suitable access nodes. Therefore they contribute to the goal of reaching x1000 higher traffic volume at the cost comparable to contemporary deployments.

3.11 Conclusions

The mechanisms described in this section aim at providing seamless experience for moving users in 2020 and beyond. Our investigations suggest that performing a handover using network infrastructure, similarly as it is done in LTE-A standard, is still the most feasible. However there are a number of enhancements that could be foreseen for this procedure. In case of ongoing direct D2D operations it is desirable that involved devices stay under the control of the same access node.Load balancing in multi-RAT environments can be scaled with the number of layers and/or RATs in UDN if the user device represents the anchor point for load balancing decisions. Interoperability can thus be guaranteed without costly exchange of control information between nodes in a multi-RAT, multi-vendor environment. There are also clear benefits in utilizing information about the device environment. Some prime examples that were identified are cell load information and availability of high capacity cells in the neighbourhood. This information (in combination with context awareness described in the next chapter) can be also used for smart large-scale operations such as massive software upgrade or release of popular content. Increased complexity at the network side can be also offset with reduced power consumption of devices, as proven in realistic deployment using radio finger prints.

4 Context Awareness Approaches

4.1 Introduction

Context awareness is one of the promising emerging techniques which can impact the overall user experience. By context information we understand '(…) *any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and application themselves'* [Dey00]. Some of the examples of the previous research done can be found in [BCG09], [MDS+12], [PBS+12] or [PKW+12]. Studies carried out in the METIS project exploit data such as location information (and consequently user's trajectory), user preferences or the capabilities of individual devices. Numerical evaluation of proposed solutions reveal a potential of context awareness to provide a smart resource allocation, improve the mobility management and help the network to tailor the user experience according to the demand and expectation. Acquisition, modelling, distribution and final decision making are the four fundamental functionalities of a context-aware system, therefore, the approaches described below cover also potential enablers to process and provide context information.

Future 5G system is expected to be an umbrella for new as well as legacy standards such as LTE-A or WiFi. Already today we can observe that in many areas it is already possible to access from one place different solutions defined by multiple standards and different technologies are highly overlapping (separated usually in frequency domain). This poses a new challenge of appropriately distributing devices/services to the different layers/networks in order to fully utilize the potential and advantages of each solution.

Context awareness creates the possibility to dynamically adapt to the application needs in order to provide seamless delivery of right information to the right recipients. It can be also used to enhance the performance of already existing networks, especially in the area of mobility related performance indicators.

First three solutions described in this chapter focus on enablers for context awareness. Remaining approaches show the benefits of exploitation of context awareness.

4.2 Optimized Distribution Scheme for Context Information

The exploitation of context information enables the efficient network operation. On the other hand it may require additional signalling to transfer the information from the points where it is available to the points where it may be exploited. Context may concern dynamic inputs or static ones, available in the RAN or the evolved packet core (EPC). Additionally, context information is usually acquired, processed and exploited at different network entities which are placed at different positions and layers of the mobile radio network, and it has to be exchanged via several interfaces.

Table 4-1. Feature Extraction with the use of Information Gain

FeatureSelection (X, L^1, k)	***Explanation*** ***X: The table of observations*** ***L^1: The class labels*** ***k: the number of features to retain***
for i=1:1:columns(X**)**	For all features of X, calculate their IG given the set of labels.
IG_i = IG($X_{:,i}$**,** L^1**)**	
Y,indexes **= quicksort(**IG**)**	Sort the IG values in descending order and retain the indexes of the k highest values. Remove all other columns from X and return the remaining array.
return $X_{:,indexes[1..k]}$	

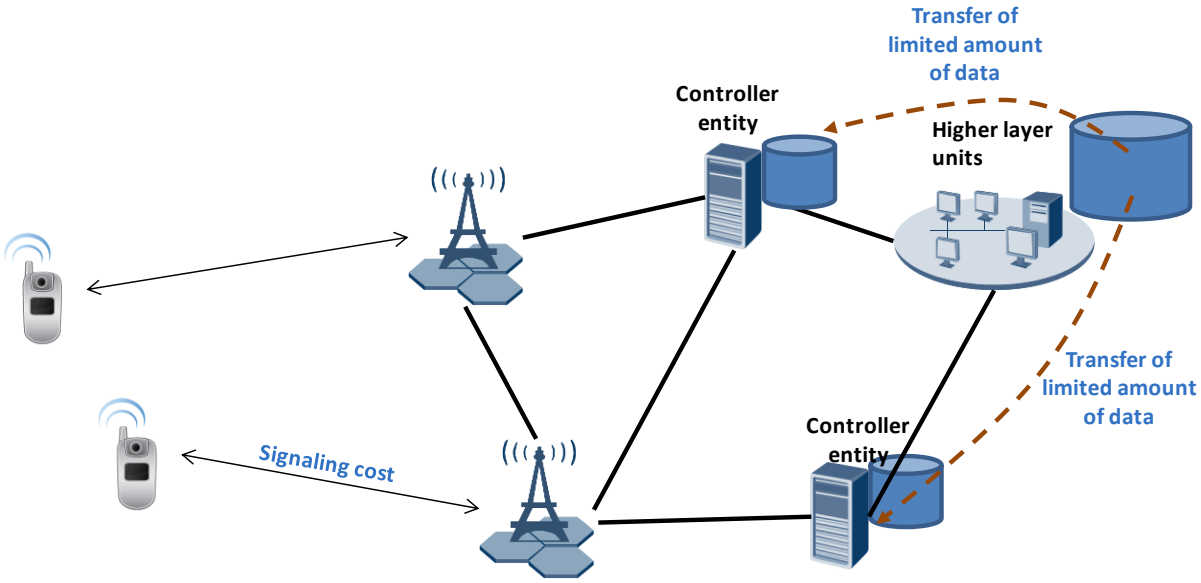


Figure 4-1. Optimized context information distribution scheme for using Controller Entities

In this technology component we propose the application of information gain (IG) statistical tool Table 4-1 in the static information available in the EPC (e.g., in the Home Subscriber Server) so as to avoid transferring the parameters that do not provide extra information regarding the users. The IG measures the amount of information about the class prediction, by considering only one input/feature. In other words, by using the IG, we try to evaluate what is the effect of each input in the differentiation of the users/devices, and we keep only the inputs that play significant role in the differentiation of the users/devices (i.e., Feature Extraction). Additionally, considering that there is a vast amount of available information in the EPC, we propose the introduction of the Controller Entities (i.e., entities positioned closer to the decision making points, for enhanced static context distribution – this enables the rare communication with the HSS, the charging functions, etc.). Then, as shown in Figure 4-1, only the required information fields that differentiate the users significantly are transferred to the Controller Entities.

4.3 Context Information Building Using Data Mining Techniques

The objective of this technology component is to provide a mechanism capable of exploiting user behavioral information towards improving radio resource management operations like cell (re)selection and handover. Existing 3GPP radio access solutions do not exploit the user habits and identified behavior at the user equipment side, which are processed using data mining schemes for the sake of more efficient resource allocation.The proposed mechanism relies on the principle that the users tend to have similar behaviors in specific locations during specific time-periods [WSP+09]. Essentially, given historic observations and a specific time period we attempt to predict user behavior using a-priori known locations, services and their durations. The goal is to exploit user habits for better user to RAT/layer mapping and increase user capacity and experienced throughput.

In order to assess the validity and viability of the approach we run a number of experiments in a simulated environment comprising a single floor shopping mall with dimensions 100 m x 200 m having various LTE femto and macro cells deployed. Moreover, we assumed the existence of different user behaviour profiles (morning/afternoon shift employees and visitors). The assessment was carried out employing throughput, delay, packet loss and number of handovers as KPIs. Each experimental scenario was run twice (with and without the prediction engine) and extracted KPI values were compared. Overall results (Table 4-2) demonstrate the added value of the proposed mechanism (increase in throughput, reduction in delay and packet loss, reduction in the number of handovers).

Table 4-2. Context enabled scenarios are compared against baseline scenarios (e.g. KPI Context Enabled Scenario/ KPI Baseline Scenario)

KPI	1 macro cell				2 macro cells	
	10 femtos, 40UEs	10femtos, 80 UEs	20 femtos, 40 UEs	20 femtos, 80 UEs	20 femtos, 40 UEs	20 femtos, 80 UEs
UL Throughput (Kbps)	+38%	+11.2%	+15%	+10.28%	1.7%	-1.8%
DL Throughput (Kbps)	+13%	+3.9%	+5.4%	+2.56%	1.4%	-0.3%
UL Delay (s)	-38%	-17.6%	-60.8%	+24.36%	-53.2%	-32.9%
DL Delay (s)	-50%	-54.7%	-56.6%	-16.29%	-48.1%	-28.7%
UL Packet Loss ratio	-53%	-33.5%	-55.7%	-25.21%	-3.8%	4.4%
DL Packet Loss ratio	-22%	-11.83%	-25.1%	-6.19%	-22.2%	-1.9%
Handovers	-8.7%	-5.94%	-5.8%	-5.52%	-2.8%	-1.9%

4.4 Signalling for Trajectory Prediction

This TeC is applicable in scenarios in which a large number of users travelling together (e.g. in public transportation) form a data intensive moving user group. Exploiting movement data and context information of diurnal user movements (public transportation, vehicular users, etc.) allows for predicting user cell transitions and lays the basis for designing efficient resource reservation or smart resource mapping schemes. The objective of this technology component is to evaluate the signalling involved in trajectory prediction and proactive triggering of suitable RRM mechanisms. The next cell prediction requires parameters such as position (e.g., GPS coordinates), velocity of user group, geometry of the UE with respect to neighbouring cells signalled to the BS at regular intervals. Further, context messages indicating the arrival of moving user groups are required to be signalled from serving BS to the predicted next cell to proactively trigger suitable RRM schemes. The positioning algorithms considered are Assisted Global Navigation Satellite Systems (A-GNSS) and Uplink Time Difference of Arrival (UTDOA) [3GPP14-36355].

3GPP solution makes use of signal strength measurements for identifying a target BS and handover execution. However, there is no mechanism for predicting cell transitions well in advance and send context messages, which indicate the arrival of data intensive moving user group, to the predicted next cells and for proactively triggering suitable RRM schemes. The messages required to enable context-aware RRM by trajectory prediction and their sizes are depicted in Figure 4-2(a). The “Geometry (dB)” value comprises geometry values toward six closest BSs, thus geometry based approach results in least message sizes. The signalling overhead involved in positioning are depicted in Figure 4-2 (b). A-GNSS requires more signalling than UTDOA method, but yields higher accuracy [3GPP14-36355].

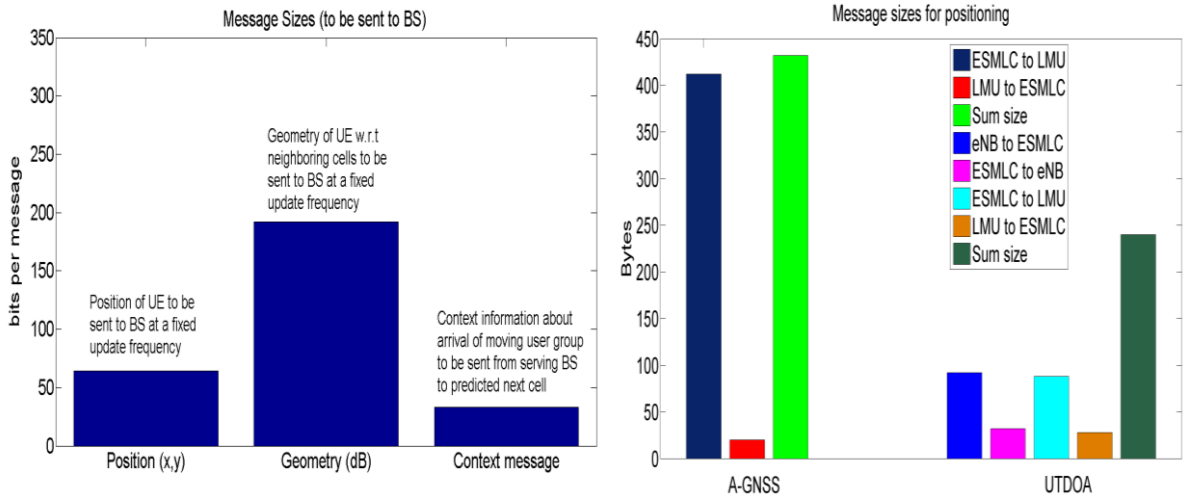


Figure 4-2. Simulation results: (a) Sizes of individual messages to be sent to BS (b) Messages required by positioning methods (cf. [3GPP14-36355])

4.5 Context Awareness Through Prediction of Next Cell

This TeC addresses scenarios in which a large number of users travelling together (e.g., public transportation) form a data-intensive moving user group. The objective of this technology component is to exploit context information for predicting cell transitions of such moving user groups. Based on this context information suitable RRM schemes, e.g., Load Balancing (LB), will be proactively triggered to accommodate the expected user traffic demands in the predicted cell. The expected target cell can be obtained by collecting parameters, such as Global Positioning System (GPS) coordinates, velocity of user group, distances to neighbouring BSs, and geometry measurements with respect to neighbouring BSs. Exploiting movement data and context information of diurnal user movements (public transportation, vehicular users, etc.) allows for predicting user cell transitions and lays the basis for designing efficient resource reservation or smart resource mapping schemes. The signalling overhead associated with this mechanism is discussed in Section 4.4.

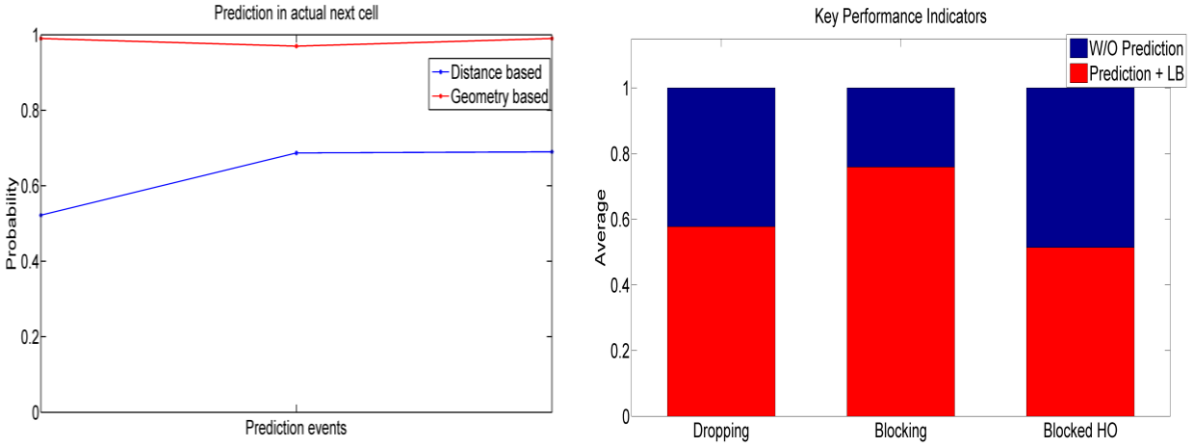


Figure 4-3. Simulation results: (a) Performance of prediction schemes (b) Performance improvements w.r.t. Key Performance Indicators

Existing solutions for next cell prediction are based on complex schemes involving machine learning [MPM12], neural networks [LH05], route clustering [Laa05] etc. Complexity of such schemes can render them unable to proactively trigger context-aware RRM in timely fashion. 3GPP solution makes use of signal strength measurements for identifying a target BS and handover execution. However, there is no mechanism for predicting next cell well in advance

and send a context message to the predicted next cell for indicating the arrival of data intensive moving user group and for proactively triggering a suitable RRM scheme. Further, context information about diurnal behaviour of user groups improves the accuracy of prediction of cell transitions. In particular, geometry based prediction of user trajectories yields superior performance compared to distance based prediction (cf. Figure 4-3(a)). The proposed solution shows reductions in dropping of users (around 40%), blocking of users (around 20%) and blocked handovers (around 50%) [Kle15], illustrated in Figure 4-3(b).

4.6 User Oriented Context-Aware Vertical Handover

The objective of this technology component is to provide a mechanism capable of selecting the optimal RAT to connect to in an ultra-dense radio environment. Existing solutions are mainly using simple Received Signal Strength (RSS)-related schemes. In addition, they have not been deployed in commercial systems due to technical issues that needed to be tackled before the adoption of integrated solutions. The described solution, taking advantage of the evolution of recent 3GPP standards offers a holistic business case, feasible and deployable.

COmpAsS (COntext-Aware RAT Selection) mechanism triggers handover decision based on context-related information, which is built by collecting parameters such as the mobility (velocity) of the UE, the traffic load of the macro/femto LTE-A BSs or (Wi-Fi Access Point), as well as the backhaul load of the network and session-related context information (e.g. how sensitive a specific service/session is to network latency). Selection of appropriate layer or access node is defined on a per session/application basis.

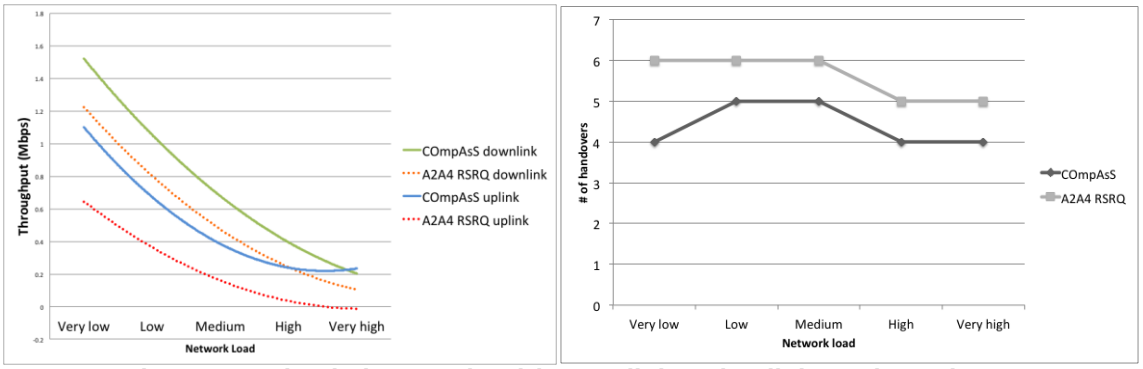


Figure 4-4. Simulation results: (a) Downlink and uplink UE throughput, (b) Overall number of handovers

In order to evaluate the performance of the proposed mechanism we simulate a realistic business case scenario of a shopping mall comprising of 3 floors, and 20 shops per floor (with LTE femtocells and Wi-Fi APs installed). The KPIs, which were selected (i.e., UE throughput, latency, packet loss and number of handovers), were juxtaposed to the well-established LTE handover algorithm A2A4 RSRQ [RTR07]. Regarding the throughput (Figure 4-4(a)), latency and packet loss results, both in downlink and uplink cases, it is showed that the proposed scheme outperforms the legacy LTE solution in all KPIs by 10-20%. In addition, the number of handovers is minimised in case of the proposed scheme ((Figure 4-4(b)), resulting in reduction of signalling, as well as in the minimization of the experienced delays due to the reduction of overall number of handovers.

4.7 Handover Optimization Using Street-Specific Context Information

This technology component describes the mechanism and the algorithms for handover optimization using the location context information. Many modern vehicles are already equipped with advanced cellular communication terminals (e.g., based on Universal Mobile Communication System (UMTS) and LTE). In order to enable reliable cooperative driver assistance services that improve traffic safety and efficiency in the future, a seamless mobility support for vehicular terminals is required. As the traditional LTE handover optimization only aims at choosing Cell Individual Offset or Handover Margin (HOM) and exploiting the speed

profile by adjusting the handover Time-To-Trigger (TTT), the driving direction and radio propagation properties are not addressed. In contrast to conventional user terminals that move unpredictably, vehicular terminals move along a defined path (i.e., streets) and within a certain speed range. Moreover, the position and trajectory of vehicles are usually well known due to on-board GPS and navigation systems. Thus, handover decisions can be improved based on this additional information.

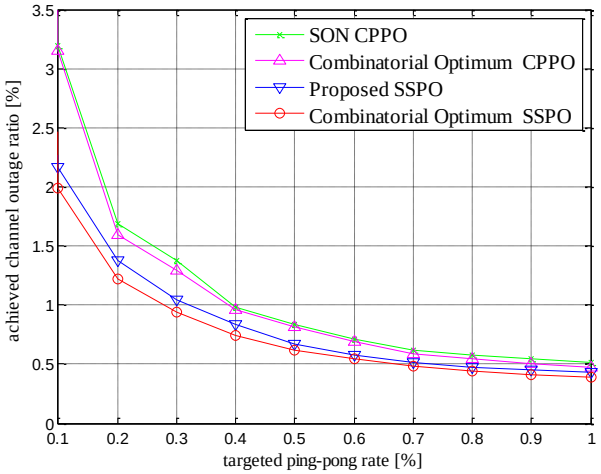


Figure 4-5. Performance evaluation of the street-specific parameter optimization (SSPO)

Two implementation variants can be considered, depending on the location of the decisive unit for the handover. The first possibility is to employ a central unit which manages the handover parameters for each street. The vehicles send periodically the location information (e.g., the street ID) to the eNB which is connected to the central unit. Then, based on an optimization algorithm, the corresponding handover parameters (HOM, TTT) will be sent to the vehicles. Another variant is to keep a database of handover parameters at every vehicle. Based on the location information, the vehicles select the parameter for optimization. However, the data base needs to be updated by the central unit in case a different optimization objective is targeted or a different traffic profile is observed.

Figure 4-5 displays the performance evaluation based on the algorithm proposed in [RFL+13]. Channel outage ratio, which is defined as the ratio of duration with low SINR (<-5dB) to the total time in the handover region, is used to measure the handover performance. Two algorithms, including the proposed street-specific parameter optimization (SSPO) and SON cell-specific parameter optimization (CSPO) are evaluated. In general, the SSPO outperforms CSPO, especially when the target ping-pong rate is very low. Furthermore, the SSPO achieves a similar performance as the combinatorial optima by brute-forcing. For more details on algorithms and results please refer to [RFL+13].

4.8 Context-aware mobility handover optimization using Fuzzy Q-Learning

Mobile networks typically form dynamic structures, where continuously new sites are deployed, capacity extensions are made, and system parameters are adapted to local conditions. In the past, mobile network operators spent a lot of efforts on manually tuning site-specific handover parameters. In order to reduce the degree of human intervention in network optimization processes, optimal handover parameter settings, which yield an appropriate trade-off, e.g., with respect to connection drops, handover failures, and ping-pong handovers, have to be identified and enforced autonomously.

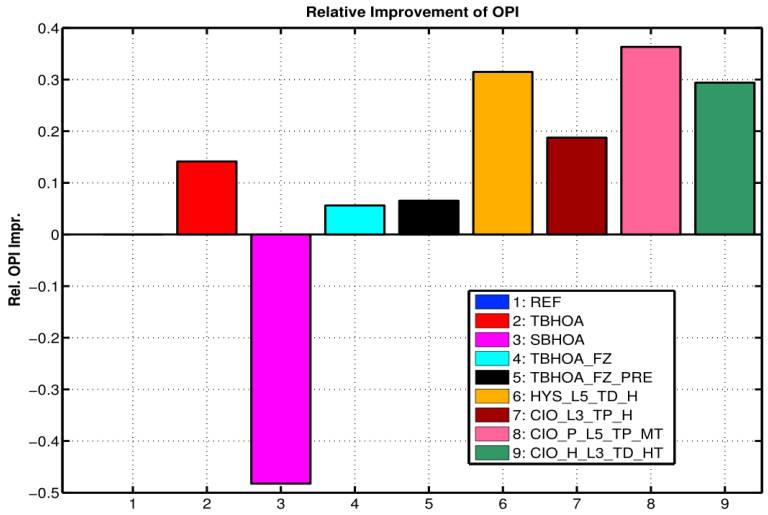


Figure 4-6. Relative OPI improvements per optimization scheme (fixed adaptation strategies are outperformed by FQL based solutions [Kle15])

Therefore, the developed solution implements a Fuzzy Q-learning (FQL) based self-learning and self-tuning mechanism [Kle15] that aims at optimizing the robustness of mobility support according to locally observed conditions and that in the long-term is able to reduce operational expenditure (OPEX) and improve QoE.

Figure 4-6 depicts the performance of various handover parameter optimization schemes with respect to an Overall Performance Indicator (OPI), which accounts for connection drops, handover failures, ping-pong handovers, and user satisfaction. In contrast to a 3GPP legacy system or optimization approaches relying on fixed adaptation strategies (#2-#5), most significant improvements of 32%, 19%, 36%, and 29%, respectively, are achieved by different variants of the developed self-tuning solution. In particular, scheme 8 (*FQL (CIO_s & CIO_t)*) is particularly suitable in case of direction-oriented user mobility, since it adapts handover parameters in a cell pair-specific manner. More details can be found in Section A.3.2 and [Kle15].

4.9 Smart Mobility and Resource Allocation Using Context Information

Mobile media streaming suffers from the time-variant data rate of cellular networks. In coverage holes and overloaded cells, the RAN cannot fulfil the rate requirement of the streaming service for the required time. As a consequence, the playout buffer of the streaming application at the UE runs empty and the media stream stalls. These service outages are the main road blocks for high quality audio and video streaming in cellular networks today. To avoid such buffer underruns, we introduce a new system for wireless resource allocation in cellular networks. Operating in the core network, this system assigns minimum bitrate (MBR) constraints to the RAN schedulers using the concept of guaranteed bit rat bearers as standardized in [3GPP14-23203]. By satisfying these requirements, the RAN scheduler then minimizes probability of a media stall, while maximizing spectral efficiency.

The novelties compared to conventional MBR bearer control lie in our anticipatory and context-aware adaptation algorithms. Being aware of streaming parameters such as encoding rate and maximum buffer size, the context-aware resource allocation (CARA) engine in the core network uses a local buffer model to estimate the current buffer state at the UE. Based on a prediction of the UE's upcoming channel state, the CARA engine then selects the MBR such that the overall duration of stalls is minimized.

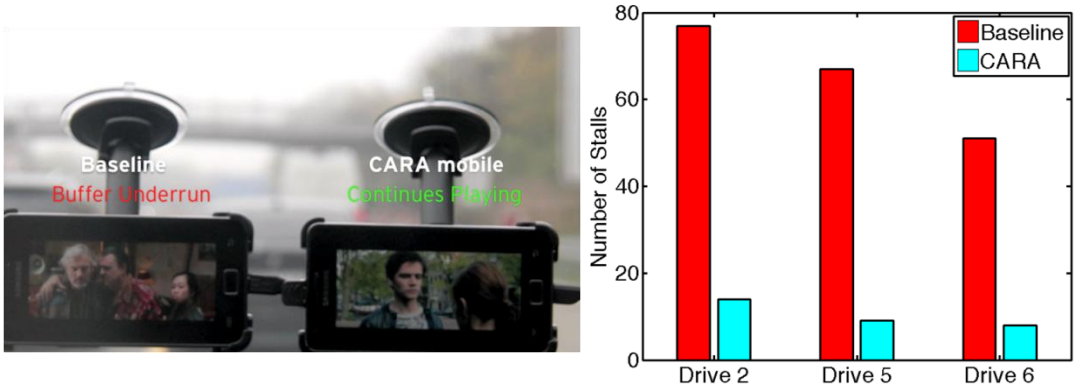


Figure 4-7. Results and setup of the CARA test drives showing a six-fold reduction of the number of stalls for a high-definition live video streaming in a real cellular network

By allocating the MBR requests according to the predicted channel gain, the system highly improves the user's QoE, while even maximizing the average spectral efficiency. This results from the fact that streaming requests are served at high channel gain, while at low channel gain the streaming application satisfies its MBR constraint only from the buffer. This enables multiuser diversity gains by allocating the channel to more UE's with high channel state. Simulation results show spectral efficiency gains of 20%. Measurement results obtained from test drives in diverse scenarios consistently show a more than six-fold reduction of the number of stalls. These outstanding results make all the difference between acceptable and unacceptable video quality on the road.

4.10 Long-term context-aware scheduling for UDN

The proposed method addresses the problem of minimizing end to end packet delay by time-unit-based packet scheduling to meet individual packet deadline exploiting knowledge of deadlines and channel information (CSI) from next *n* upcoming time units. The number of dropped packets is considered as a system performance indicator. The proposed scheme is named Proactive Delay-Minimized Scheduling (PDMS).

In Figure 4-8 the proposed performance of PDMS scheduler is shown at system load of 65% and compared against a non deadline-aware scheduler (cf. Section A.3.3). When scheduler is aware of long-term CSI, 30% reduction in dropped packets can be observed. At the same load, awareness of long-term CSI variations of a particular user cluster (context users), lead to less system packet drop (3%) which translates to almost 15% less dropped packets for the context users only. It can be observed that PDMS scheduling gains are lower, the higher the system load is. Apart from higher load, lower diversity gain because of using fewer resources, contributes to degradation of performance. More detailed evaluation and description of PDMS scheduler can be found in Section A.3.3.

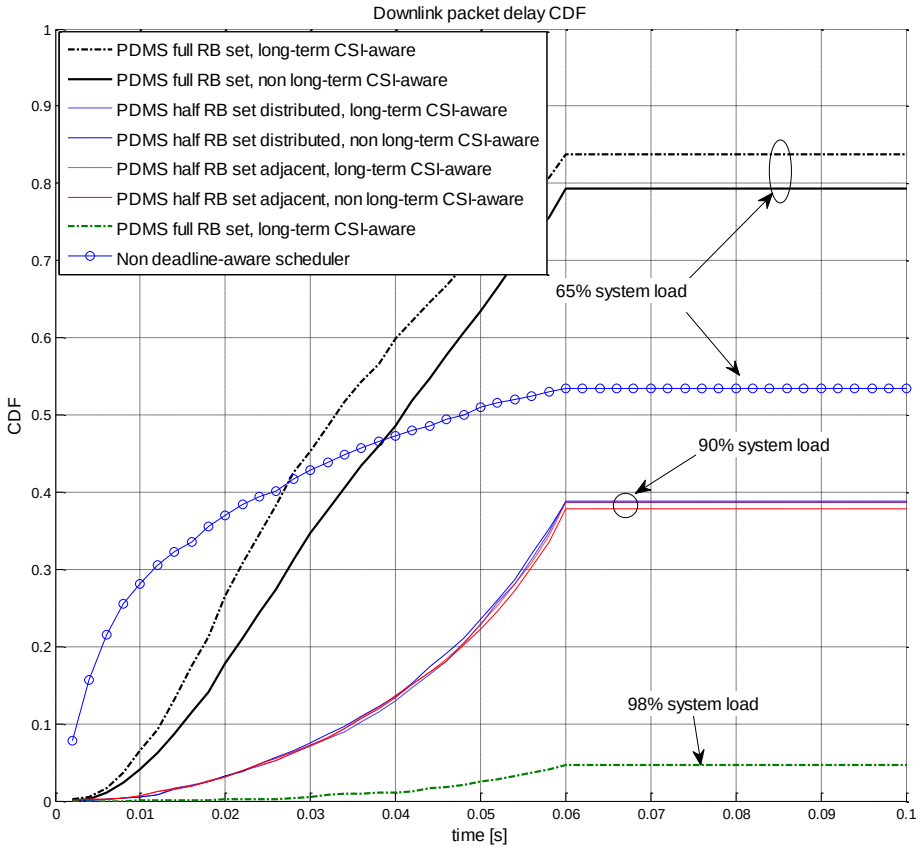


Figure 4-8. Performance of PDMS scheduler. Downlink packet delay for different load configurations within studied load regimes.

4.11 Context-Based Device Grouping and Signalling

Context-based device grouping and signaling aim at reducing the signalling overhead for MMC traffic and mitigating the potential congestion in the signalling channels. In general it is observed that the traffic coming from machines is topologically (i.e., machines that are located in the same, or close, or similar locations) and time related. Thus, we exploit the bursty nature of the MTC traffic and effectively remove the redundancy in the transmitted messages by either suppressing or compressing the messages with redundant content. Furthermore, event-dependent messages that are not redundant are being scheduled and transmitted in a coordinated manner. This enables the reduction of the delay since the devices are accessing the medium using less RACH attempts and at the same time it increases the success probability for the first attempt, even with large number of devices.

Figure 4-9 shows the uplink signalling overhead produced by the MTC devices that attempt to access the network. The considered scenario assumes up to 30000 MTC devices (within the coverage of 1 macro BS) accessing the network. We observe that when compared to SOTA 3GPP solutions, the highlight of the proposed approach is to exploit the redundancy in the signalling messages, which is typical during the congestion period for MMC traffic. The redundancy reduction substantially decreases the signalling overhead and mitigates the traffic congestion. Additionally, with the compression of content-related messages, the signaling overhead in the uplink is further reduced and thus the protocol efficiency is improved [MET13-D11].

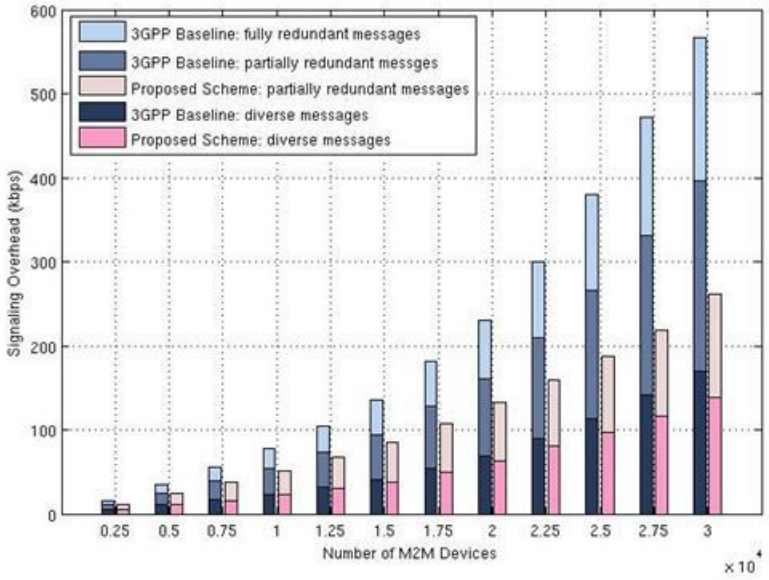


Figure 4-9. Signalling overhead for different number of devices

4.12 Context-Aware Smart Devices and RATs/Layers Mapping

A growing demand for high data rates, seamless connectivity, and a heterogeneous landscape of RATs requires efficient means for call admission control and RRM. Since mobile network operators are interested in an efficient and optimized utilization of their infrastructure, a context-aware radio resource mapping approach is proposed that combines the method of Joint Call Admission Control (JCAC) and the Dynamic Bandwidth Adaption (DBA) [KLM+11a] for maximizing the overall system utilization and revenue.

The JCAC method is handled as a General Assignment Problem (GAP), which belongs to the class of bin packing problems. In the GAP N items need to be assigned to M bins. Each bin has a certain capacity c and each item has a weight w and a utility u depending on the corresponding bin. Here, the items correspond to user service requests and the bins to RAT access nodes, which differ with respect to the bandwidth that can be provided. The main objective is to maximize the overall resource utilization by finding an optimum mapping of services onto RAT access nodes. The overall utility is calculated based on utility functions that can be specified by the network operator.

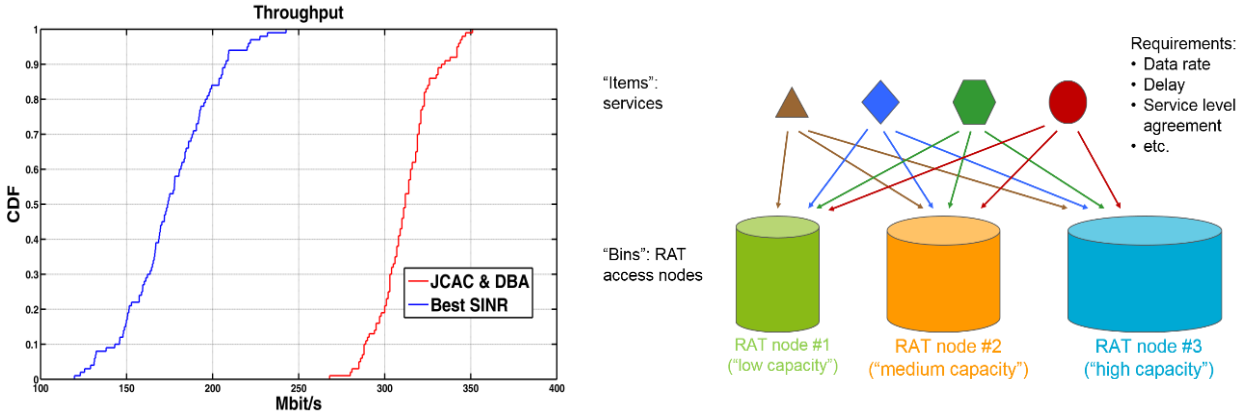


Figure 5-7. (a) CDF of throughput gain, (b) principle of GAP

Figure 5-7 shows the throughput CDF in a specified scenario for combined JCAC & DBA compared to a reference case. The overall throughput using JCAC & DBA is significantly increased. Overall resource utilization is highly dependent on the chosen utility functions.

4.13 Impact on the METIS 5G System Concept

Context awareness proved to be a useful technology to boost the performance of xMBB and M-MTC. Two large areas where context awareness could be applied is decision for appropriate mapping of users to the right cells/layers/RATs (in order to efficiently exploit available radio resources we may allow users to attach not to the best cell available (in terms of signal quality), but to the one which is optimal from the holistic radio resource point of view) and the self organizing network (SON)-like area, for smart tuning of RRM/mobility related settings. Especially the latter option allow for a reliable data connection. Some of the schemes use context information directly for RRM. Context information can also be used for efficient handling of MMC traffic, and specifically for suppression of redundant messages.

4.14 Addressed METIS Goals

The mechanisms proposed in this chapter address several METIS goals. Optimal allocation of users to different cells/layers/RAT help to achieve 1000x higher traffic volume in realistic deployments, suppression of redundant MMC enable support of 10-100x higher number of machine type devices, but most of all, context awareness helps to achieve a personalized user experience (cf. Section 4.12).

4.15 Conclusions

Context awareness proved to be a useful method for enhancing overall user experience, but also for optimizing system metrics such as capacity (Section 4.11) or mobility robustness (cf. Section 4.7). Gains coming from utilization of context information can manifest in multiple ways. Several of the technology components developed in METIS and described in this document, exploit prediction of user's position. This result in a reduced number of call drops caused by handover failures, but the same mechanisms of user's position prediction can help with providing a seamless data connection for users on the move. This capability was shown for both indoor and outdoor users. Indoor users can benefit from having their data exchange finished, before entering a coverage hall (e.g. an elevator) or heavily congested cell. Outdoor users were proved to gain from anticipatory resource allocation by e.g. avoiding video buffer underrun when going through a tunnel. In some of the considered cases performance of context aware solutions depends heavily on precise estimation of current/future device location. Especially indoors, this is still a challenge that needs to be tackled in order to make these solutions work. However, there are also mechanisms exploiting context information, where prognosis on next cell that the user will attach to is sufficient.

Context awareness allows also bringing down the cost of network maintenance. A good example here is a context-aware optimization of handover parameters. In the past, mobile network operators have spent a lot of efforts on manually tuning access node-specific settings. In order to reduce OPEX and improve QoS for mobile users, a context-aware self-tuning mechanism is proposed that yields significant performance improvements [Kle15], e.g., with respect to connection drops, handover failures, and ping-pong handovers (cf. Section 4.5).

Exploitation of context awareness brings biggest benefits with repetitive user behaviour in heterogeneous environments in terms of user requirements. Therefore it is also suitable as an enhancement to machine type communication. Requirements for MTC can be directly translated to PHY/MAC layer parameters. QoE for humans is far more unpredictable, subjective and needs an individual assessment in order for context awareness to work.

For an efficient exploitation of the context information, schemes reducing the signalling overhead should be considered. Thus the valuable content shall be extracted from the available information and only this should be transferred closed to the decision points. Information Gain statistical tool (cf. Section 4.2) seem to be a useful approach for identifying which parameters may be used for differentiating the users.

Since the signalling channels will be a bottleneck for the MTC traffic we have to exploit the bursty nature of the traffic and the redundancies in the messages so as to reduce the signalling overhead and mitigate signalling congestion. Thus, avoidance of transmitting redundant messages over air interface, as well as the scheduling of the event-dependent messages that are not redundant minimizes the congestions. Additionally, the correlation of the content-related messages further reduces the signaling overhead. The previously described coordination improves significantly the protocol efficiency.

5 D2D and V2X Mechanisms

5.1 Introduction

The benefit of D2D operations, such as proximity, hop or reuse gains have been described in numerous publications e.g., [JKR+09], [FDP+12], [DRW+09], [MLP+11], [RF11] or [BFA11]. One of the major identified showstoppers for direct D2D that prevent smooth introduction of this mode of operation in e.g., 4G LTE-A, are compatibility issues [FPB+13], [PFM+14]. Therefore 5G have a unique opportunity to introduce this feature with a native support from the very first revision of 5G standard. Direct D2D mode of operation is not only foreseen for human centric data exchange, but it also piggybacks several other use cases, e.g., direct communication between sensor/machine type devices, Vehicle-to-Anything (V2X) communications or inbuilt support in national security, public safety or emergency situations [FPS+15]. V2X is considered to be one of the most interesting applications for the 5G that can directly influence the life standard and safety level of a vast group of people in the future mobile society. Autonomous vehicles and mission-critical automotive services allowing in-time reaction for a potentially dangerous situations can bring numerous benefits, starting from driver's or vulnerable road user's safety and ending on optimization of the fuel consumption. Still, from the technical point of view, the nature of direct D2D connection needs redefinition of RRM aspects comparing to classical cellular communications via RAN infrastructure. Solutions presented below tackle these key aspects: discovery of devices, power control and radio resource management.

5.2 Unified Resource Allocation Framework for D2D Discovery

In contrast to existing device discovery schemes that designed for a completely flat architecture, the proposed unified framework for D2D discovery is its ability to operate in both in-coverage and out-of-network coverage scenarios [QQF+14], [FPS+15]. Our design takes advantage and extends the concept of D2D communications underlaying a cellular network - in which the eNB owns the discovery resources – such that the allocation of discovery resources is efficiently managed in both in-coverage and out-of-coverage situations. The key idea is that in the absence of network coverage, so called cluster heads (CH) with special capabilities, take over some of the functionalities of the network and control the resource utilization for a group of attached devices. When network coverage is partially available, the CH devices can be smoothly integrated in the infrastructure. In the absence of infrastructure nodes, the CHs provide network assistance required by the underlaying devices.

An efficient system is characterized by a minority of devices acting as CHs, while most devices are camping on a cluster associated with a CH in order to reduce energy consumption and interference associated with synchronization signal transmission and detection. Besides, when the clustering concept is not feasible, the framework switch to a fully distributed device-centric approach where all the devices autonomously select discovery resource from a prior known resource pool and directly discover each other in a fully or partially synchronous manner.

Compared with non-network assisted discovery processes this unified approach provides benefits in terms of high power efficiency, short discovery time and efficient resource usage [QQF+14]. It also enables a seamless transition between both scenarios and smooth integration of clusters and cells (where available) and as a consequence the reuse of all the network-assisted D2D mechanisms such as RRM and power control under partial or out of network coverage.

Figure 5-1 and shows that the clustering concept is able to achieve 99.9% discovery ratio in METIS Emergency communication test case [MET13-D11] and below 5 seconds discovery time, respectively. We evaluated three different ways of selecting the CHs: cluster-head based hybrid and threshold-based. More details about different clustering procedures can be found

in Section A.4.1. The results of the threshold-based approach show that taking into account measurement about the signal strengths and the quality of links between the devices achieves a trade-off between the coverage ratio, the discovery time and the energy consumption.

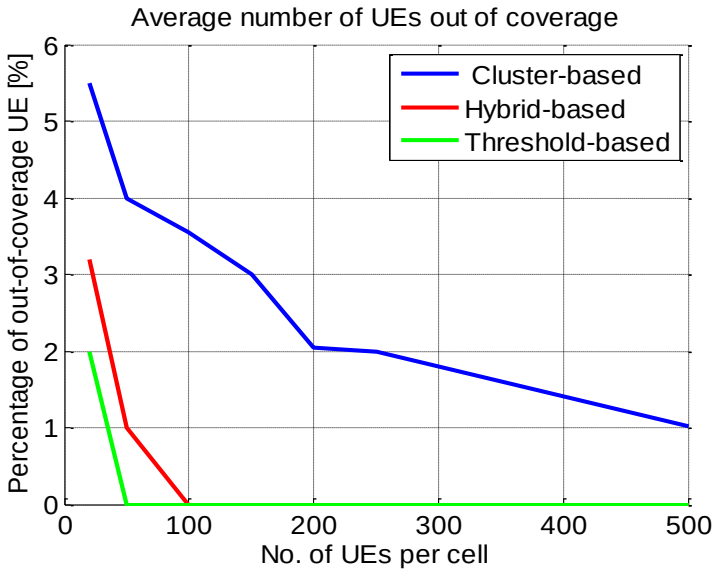


Figure 5-1. Percentage of users under coverage i.e. UEs that connect to one CH

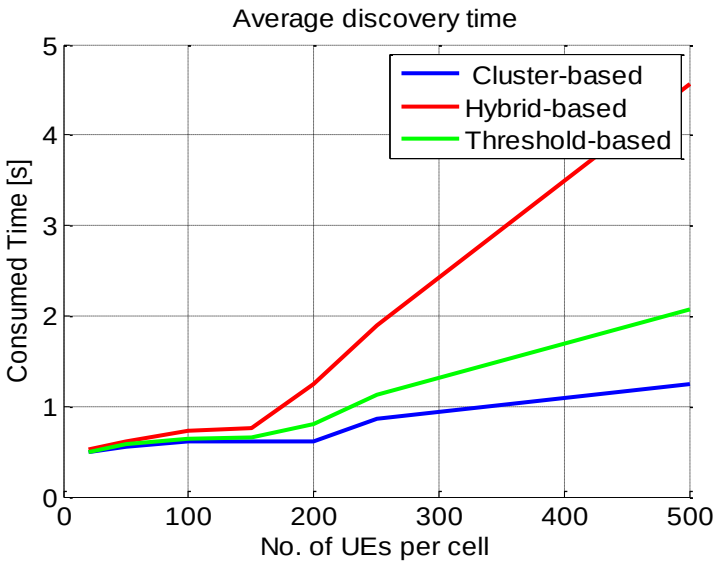


Figure 5-2. Discovery time for METIS emergency communication test case (TC10)

5.3 Distributed Channel State Information Based Mode Selection for D2D Communications

Existing mode selection algorithms are typically confined to single hop D2D connections; assume a central entity, and/or the availability of perfect CSI set and disregard advanced receiver capabilities and physical layer algorithms at the devices and/or network nodes. In contrast, the proposed distributed CSI based mode selection scheme can scale to multihop D2D paths [SFM14], lends itself to distributed implementation [RF12], does not assume full CSI information [PFM+14] and can take into account advanced receiver capabilities such as variants of physical layer network coding [FPG14-1], [FPG14-2].

Our design addresses the above requirements without requiring location information. Location-based mode selections a complementary technology component that can be used in

conjunction with or separately from the proposed distributed CSI based mode selection scheme.

The key idea is based on the observation that mode selection and the management of resources between the cellular and D2D layers are inherently intertwined. Therefore, in contrast to SOTA, the proposed mode selection scheme distinguishes (1) cellular mode, (2) direct mode with dedicated D2D resource and (3) direct mode with cellular resource reuse. Two versions of this basic idea can be implemented: the balanced random allocation (BRA) utilizes the available resources adaptively to the load in the cellular and D2D layers, while the cellular protection algorithm (CPA) protects the cellular layer from the interference caused by the D2D layer [PFM+14], [FBP+15].

Compared with 3GPP Rel-11 systems that do not support D2D communications and lack mechanisms for D2D mode selection, the distributed CSI based mode selection increases the total system rate depending on the geometry, and the availability of proximity users in the coverage area of a cellular system. Also, when devices are close to one another (less than 50 m), CSI based mode selection leads to significant total power gain as compared with 3GPP Rel-11 while the total average system rate can be increased. In realistic examples we find up to 150% rate gain while reducing the total power consumption by 50% [FBP+13]. Detailed results are available in Section A.4.2.

5.4 Location-Based Mode Selection for D2D

Introduction of D2D communication in cellular networks imposes the need for transmission mode selection algorithms, and in particular it is significant for network assisted D2D communication. Existing mode selection solutions typically rely on channel knowledge, this implies the need for channel measurements which in the case of existence of D2D links might result in increase of signaling overhead. The proposed location-based mode selection method for D2D enables a seamless transition between indirect and direct transmission modes by utilizing context information in the form of user's location. Location can be mapped to distances, and distance can be used to estimate path-loss and power required for transmission. The mode selection procedure is triggered by a network and all processing and decision making is transferred to the BS. The criterion for mode selection decisions is the transmit power of the UE, whereas the objective of the location based mode selection is to minimize the transmit power of the UEs. The direct transmission mode is selected only when the estimated transmit power required for the communication is lower than the estimated transmit power in the indirect mode.

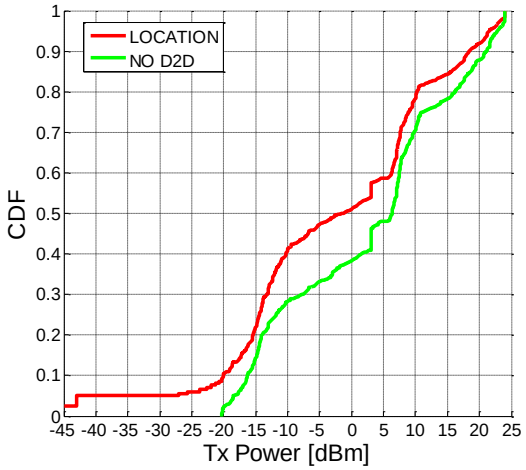


Figure 5-3. Transmit power of UEs

Figure 5-3 presents the transmit power CDF for UEs when location based mode selection is utilized along with the open loop power control (OLPC) method. A significant reduction of transmit power can be achieved by using direct transmission mode while also reducing packet delay, due to the single hop nature of the D2D communication.

5.5 Multi-cell Coordinated and Flexible Mode Selection and Resource Allocation for D2D

This technology component focuses on the optimization of the overall system performance and efficiency through RRM mechanisms that address both cellular and direct/indirect D2D traffic. In contrary to SOTA 3GPP solutions, direct D2D transmission is treated as an inherent part of the 5G system. Additionally, D2D operations are evaluated using dynamic UL/DL TDD based air interface that was developed in METIS for UDN deployments [MET15_D24]. Using such air interface allows for a dynamic allocation of time domain slots to UL, DL or direct D2D transmission on a slot by slot basis, depending on the instantaneous traffic needs and scheduling metrics.

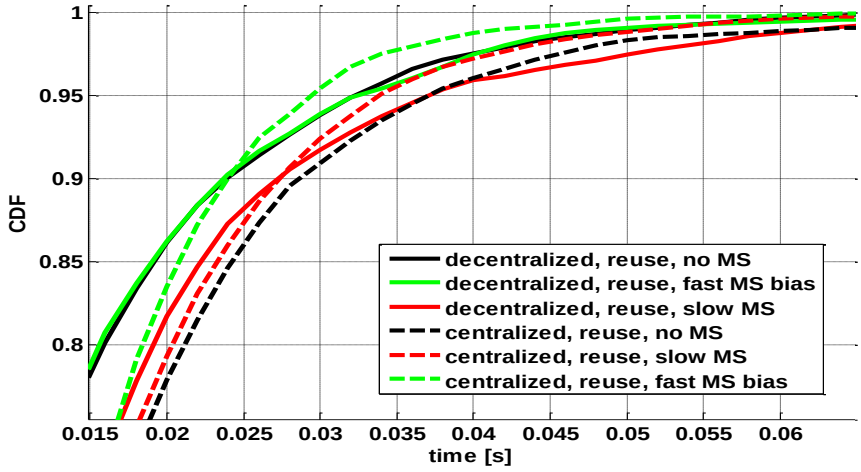


Figure 5-4. D2D packet delay CDF for different modes of operation

The performance assessment depicted in Figure 5-4 is done in two dimensions. The first dimension focuses on the comparison of completely centralized RRM mechanisms (central coordinator, e.g. a macro BS manages the resource allocation in the investigated cells, such as small cells within macro coverage) compared with a fully decentralized scheme, in which the cell performs resource allocation autonomously. In the second domain we show the influence of the mode selection time dynamics on the system level performance.

The achieved results show that fast mode selection (based on previous TTI's SINR) is outperforming slow mode selection (based on pathloss measurements) by 14% in terms of packet transmission delay (cf. decentralized fast and slow mode selection represented by solid red and green curves in Figure 5-4). An additional delay reduction, close to 10%, can be achieved by introduction of centralized scheduling (cf. fixed and solid green curves in Figure 5-4). However, both fast and centralized modes of operations come at the expense of increased signalling exchange. Further findings can be found in Section A.4.4.

5.6 Location-Based D2D Resource Allocation

This technology component investigates location-based resource allocation scheme for D2D communication. In the considered approach a centralized resource allocation algorithm for D2D that exploits users' location information to estimate the distances between nodes in the network is proposed. This centralization of resource allocation gives more control over D2D links and as a result allows for a better control of interference related to D2D communication. The selection of the best candidate for resource reuse is performed based on distance

maximization. Assuming that the large scale fading including path loss and shadowing has the largest impact on interference, increasing the distance between interferers also minimizes the interference.

The advantage of this simple resource allocation method is that it requires no CSI from the D2D users (DUEs) due to the fact that context information is used. This and the fact that the scheme allows for multiple DUEs to share the same cellular resources differentiates this approach from some of the existing solutions [BY12] [JKR+09] [WC12][ZHS10].

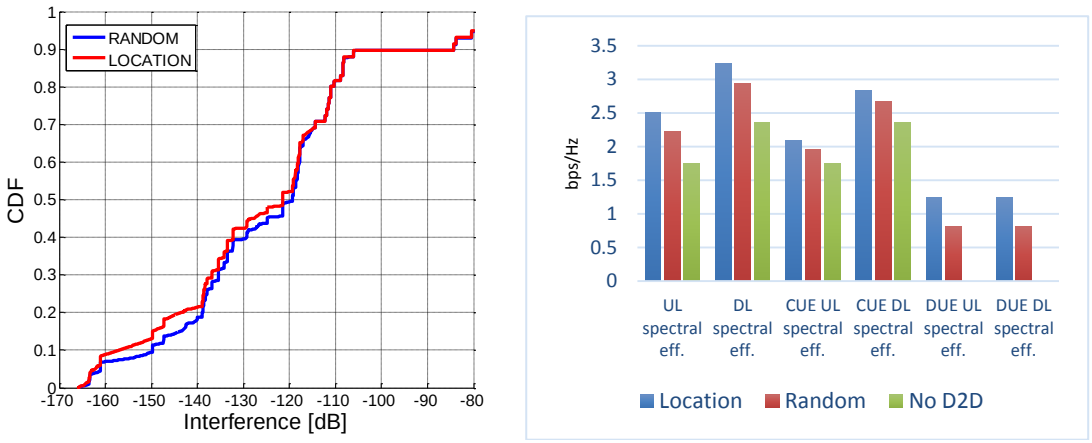


Figure 5-5. (a) D2D introduced interference CDF (b) DUE SINR CDF

Figure 5-5 (a) contains the CDF curves of the interference levels introduced by D2D pairs that reuse resources used in the uplink by cellular users. In this figure the most relevant conclusion comes from the high interference region. It can be noticed that for high level of interference all the investigated cases give similar results. Thus, introduction of a more complex method of resource allocation does not necessarily lead to the reduced interference from DUEs. However, looking at the bars in Figure 5-5 (b) that show spectral efficiency achieved in the network, improvement of spectral efficiency for location-based resource allocation approach can be seen. This implies that the proposed location-based resource allocation enables better protection against interference for the DUEs compared to the random resource allocation approach.

5.7 Context-Aware Resource Allocation Scheme for Enabling D2D in Moving Networks

The proposed Location Dependent Resource Allocation Scheme (LDRAS) aims to enable the integration of automotive services in 5G networks. To this end, it exploits the properties of V2X communications (short range broadcast transmissions, periodicity, packet size) to accommodate such services in the envisioned D2D underlay. The challenge of interference management in D2D communications is even greater in the context of a highly mobile environment. The acquisition of the CSI would be extremely costly or even infeasible in most practically relevant automotive use cases. Hence, the proposed scheme relies on cell partitioning and the introduction of zones instead. The exact shape, size, and number of these zones is determined by the deployment scenario applying the guidelines established in [BKK+14]. The spatial resource reuse is enabled through reservation of the resources for V2X (i.e., D2D) communication in each zone and the restriction of their reuse for cellular communication in certain neighboring zones. In this manner the proposed scheme enables simple RRM and efficient signaling in the D2D underlay during the operations of the network, while the complex interference alignment is carried out in a planning phase. Hereby, scheduling decisions are based on crude location information instead of infeasible channel measurements.

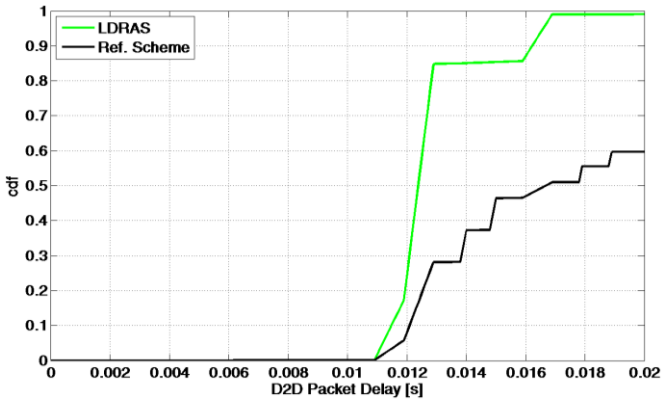


Figure 5-6. Cumulative distribution function (CDF) of the D2D packet delay under different RRM schemes

In contrast to the currently available solutions for D2D communication by 3GPP and other parties, the proposed scheme allows for the adoption of additional (i.e., automotive, V2X) services in the 5G network. The performance comparison to a selected SOTA approach [ZHS10] shows significant gains in terms of reliability. Figure 5-6 compares the achieved D2D packet delays considering a packet size of 1600 bytes and an inter-arrival time of 1.5 s. It can be seen that LDRAS guarantees a delay of less than 17 ms for 99% of the generated packets, while the reference scheme leads to infinite delay (i.e., the packets are dropped) for approximately 40% of the generated packets. For a comprehensive interpretation of the results see Section A.4.6.

5.8 Network Assisted Resource Allocation for Direct V2V Communication

This technology component considers bringing the advantages of network assisted D2D to the field of vehicular traffic safety. The key idea is based on a cluster concept where the cellular network assists few vehicles or CHs to take over the control of the direct Vehicle-to-Vehicle (V2V) communications for a group of attached vehicles. The role of the CH is to manage the radio resources on the basis of resource assignment received from the cellular node. The network control enables efficient spatial reuse of resources and allows the prediction of the traffic load using the context information collected by the network nodes related to the mobility and the density of vehicles.

Compared to the fully distributed methods used in existing 802.11p standard, this solution provides a fair and predictable channel access since the real time context information is used to adjust the resource assignment. In addition, the network assistance helps decreasing the message collision probability hence reducing the E2E delays and increasing the reliability of critical vehicular communications. Finally, compared to an approach that only use 3GPP systems and does not support D2D communications, this approach provides a higher scalability, better resource efficiency and lower E2E delay by splitting the traffic generated by safety applications between the cellular spectrum for the major part of the control traffic and the dedicated V2V spectrum for safety data traffic.

More details on performances and the gains are shown in Section A.4.7.

5.9 Resource Allocation and Power Control Scheme for D2D-Based V2V Communications'

This technology component studies RRM schemes, i.e., resource block (RB) allocation and power control algorithms, when applying the (direct) D2D underlay network to safety critical V2X communications. Under the condition of satisfying vehicular users' strict QoS requirements on latency and reliability, it aims at maximizing cellular users' sum rate with fairness consideration. To do so, the eNB is required to conduct RB allocation and power control in either long-term or short-term and either centralized or (semi-)distributed manners. In this way, the performances of both vehicular and cellular users can be improved with moderate algorithm complexity and signalling overhead.

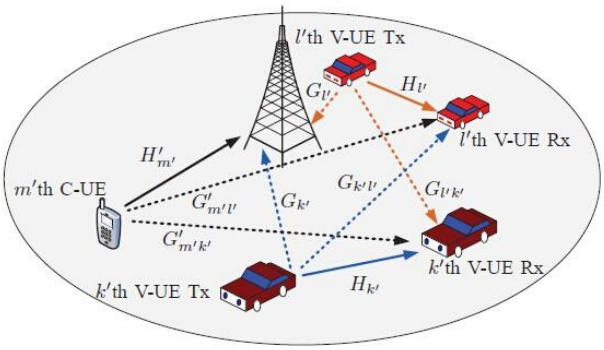


Figure 5-7. Interference illustration between D2D-based V2X and cellular communications

Compared to current 3GPP system which only supports backend-based V2X communications where the transmissions being relayed by the eNB, this technology component can largely reduce the E2E delay. Moreover, another possible legacy solution for V2X is ad-hoc communication over the 802.11p standard, which is mainly optimized for an environment with no or very low mobility and also has a scalability problem.

By implementing the proposed algorithms, several gains can be achieved. Firstly, with the assistance of the eNB, the stringent requirements on latency and reliability of safety critical V2X communications can be satisfied by utilizing the hop gain and proximity gain of direct D2D links. Secondly, the spectrum which has already been allocated to vehicular communications may be further utilized by cellular users to improve their performances. Further information on the proposed algorithm of this TeC can be found in Section A.4.8.

5.10 Joint Methods for SINR Target Setting and Power Control for D2D Communications

When D2D communication is integrated in cellular networks, transmit power control (PC) is crucial to enable the coexistence of cellular and D2D traffic sharing a common cellular-D2D resource pool licensed to the cellular infrastructure [RF11]. In our design, SINR target setting is used in conjunction with PC, and helps the design of distributed PC algorithms that can address the inherent trade-off between spectral and energy efficiency [FPB+13]. The key idea of the joint design of SINR target setting and D2D PC is to adaptively adjust the SINR targets that are associated with both cellular and D2D transmissions and set the transmit power levels that meet such SINR targets [FBP+13], [FBP+15]. This way, SINR target setting and PC work jointly and iteratively in a distributed fashion such that a system wise (overall) utility function is maximized. The overall utility function can take into account both spectral and energy efficiency objectives [FSS14].

Although the iterative approach of this technology component enables it to reach optimality in terms of spectral and energy efficiency, the number of iterations may not be practical in a fast fading environment. Therefore, the proposed scheme can also operate in a near optimal fashion, in which case the transmit power levels are set close to the optimal value, but with a much less number of iterations. Thereby, the scheme can be deployed in fast fading scenarios. In addition, the joint SINR target setting and PC scheme can smoothly be integrated into existing cellular networks, because it can coexist with the widely used LTE PC algorithms [FPB+13], [PFM+14].

Compared with 3GPP Rel-11 systems that do not support D2D communications and lack mechanisms for D2D power control, the proposed D2D power scheme can significantly increase both the spectral and energy efficiency, if there are proximal communication devices in the coverage area of the cellular system [FBP+13], [FBP+15]. In realistic scenarios, the SINR levels can increase by 5-15 dB while reducing the consumed energy by up to 50% of both the D2D and cellular layers [PFM+14]. Detailed considerations can be found in Section A.4.9.

5.11 Location-Based Power Control Algorithm for D2D

The aim of this technology component is to determine the boundaries for the maximum and the minimum transmit powers of DUEs that are sharing resources with cellular users. This is achieved by using location information for distance estimation. By using a distance dependent path-loss model, the path-loss between nodes can be estimated. This estimation can be used to determine the transmit power boundaries for a given DUE taking into account the SINR target of both receiving DUE and the sharing cellular UE.

This PC mechanism should be used on top of instantaneous PC methods e.g., open loop power control (OLPC). The upper transmit power boundary is determined to minimize the probability of severe interference to primary (cellular) resource owner caused by D2D transmission. The main purpose of the lower bound for transmit power is to set up initial transmit power level for D2D pair which should be later adjusted using either legacy PC algorithms or other PC algorithms proposed by METIS.

Since the location-based PC is a part of location-based methods framework for D2D communication the performance evaluation is performed jointly with location-based mode selection (Section 5.4) and resource allocation (Section 5.6).

5.12 Further Enhanced ICIC in D2D Enabled HetNets

Using coordinated muting (such as blank or almost blank subframes in LTE-A) for interference coordination between co-channel deployed macro and small cells can result in inefficient use of resources. For example when a macro BS mutes some of its DL resources to enable range extension of small cells within its coverage area, those DL resources are actually used only

within the coverage of the small cells. This, however, creates new transmission opportunities for D2D communication.

We propose to use the resources reserved for small cells for D2D communication when both UEs of the D2D pair are far enough from the small cells (because the D2D communication uses significantly lower transmit power than a macro cell). To support this, the UEs perform measurements during the muted macro subframes to detect if they are near an active small cell. If there is no small cell nearby, the small cell's resources can be used for D2D communication. Otherwise the BS allocates other resources for the D2D pair. This concept is illustrated in Figure 5-8 (a).

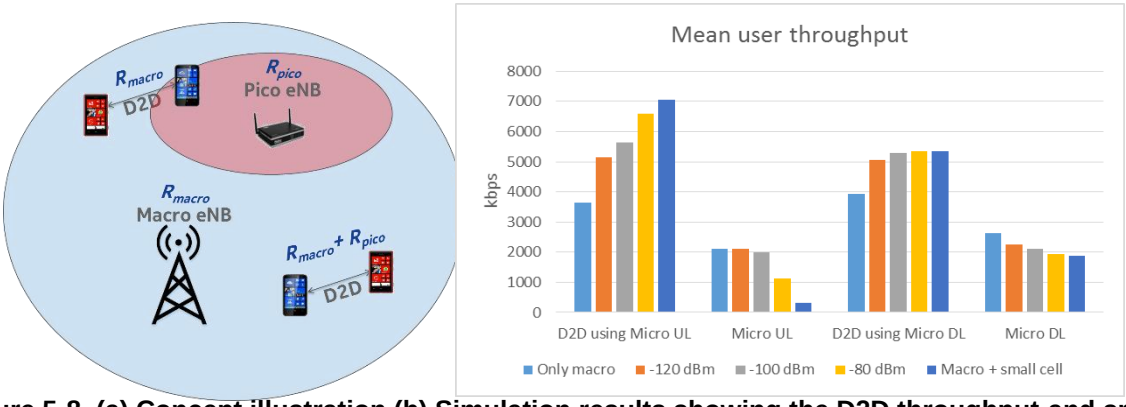


Figure 5-8. (a) Concept illustration (b) Simulation results showing the D2D throughput and small (micro) cell throughput for different D2D safety distance thresholds (small cell RSRP threshold ranging from -120 dBm to -80 dBm).

The baseline assumption of this technology component is that, as it has wide area coverage, the macro cell network is in control of D2D resources and allocation of those. Compared to the baseline, this solution enables more efficient use of radio resources by opportunistic allocation of additional resources for D2D communication. These additional resources are taken from the resources reserved for small cells. Outside the small cell coverage, these can be used for D2D without creating too much interference.

As seen from system simulation results shown in Figure 5-8 (b), this solution can provide substantial D2D throughput gain with only a minor loss to small cell throughput. The loss is due to the interference from D2D to small cells. The trade-off between D2D gain and cellular loss can be controlled by the network by setting the safety-distance threshold appropriately (e.g., using RSRP measurements), as seen in Figure 5-8 (b). Further details of the simulation setup as well as further analysis can be found in Section A.4.11.

5.13 Impact on the METIS 5G System Concept

D2D proved to be a useful form of transmission used in all METIS services. In case of xMBB it helps to increase the spectral efficiency and offload regular cellular traffic, if smartly applied. It can be also used as a basis for U-MTC, especially in the V2X communication where delays are a key factor. Finally M-MTC solutions can benefit in numerous ways from direct D2D, e.g. in machine-to-machine relaying (for coverage extension) or in CH operations. All these conditions suggest that D2D should be an inherent part of the future 5G system, with native support from the first release to avoid backward compatibility problems.

5.14 Addressed METIS Goals

Direct D2D is one of the most promising technologies that address almost every METIS objective. By smart reuse of radio resources used for the ongoing cellular transmission it can help in achieving 1000x higher traffic volume, proximity gain allows for faster data exchange in more favourable channel quality conditions which translates to 10-100x higher data rate and energy efficient operations, hop gain can be extremely useful in use cases that require really

low latencies (5x lower E2E latency). Finally, D2D with CH operations were proved to be an efficient way forward for handling of massive amount of devices (10-100x higher number of devices). It is worth to underline that D2D operations were also proved valid with flexible air interface based on dynamic UL/DL in TDD mode.

5.15 Conclusions

In this chapter we proposed a holistic network level concept for direct D2D communication, starting from device discovery, through mode selection and ending on efficient RRM and power control mechanisms. In contrast to existing D2D mode selection algorithms proposed mode selection schemes do not only select between cellular or D2D mode, but are intertwined with reusing cellular resources or allocating orthogonal resources to cellular and D2D users. One of the proposed extended mode selection algorithms (cf. Section 5.3) can be executed on a slow time scale (100 ms) and can be advantageously combined with a power control algorithm operating on a faster time scale to maintain feasibility and near optimality in terms of spectral and energy efficiency (cf. Section 5.10). An alternative mode selection mechanism (fast mode selection, cf. Section 5.5) allows for a better adaptation to instantaneous channel conditions, but requires increased signalling overhead. Centralized or decentralized D2D operations can be also dependant on the spectrum sharing architecture (if spectrum sharing is employed by a given operator). Studies provided in [MET14-D53] investigate such possibilities for cross-operator D2D communication.

Previously proposed distributed power control schemes suffered from inefficiency due to low SINR targets or infeasibility regions due to high SINR targets. In contrast, our proposed distributed power control scheme does not only achieve global optimum in terms of energy or spectral efficiency, it is always feasible due to adaptively adjusting the SINR targets. This global optimality can be traded off for a low number of iterations which makes this approach a viable approach for D2D communications even in fast fading environments.

Mode selection, resource allocation and power control for D2D communication was also evaluated in location based framework. In this approach information on the location of the devices is the criterion used for the selection of the appropriate transmission settings. Results provided in Sections 5.4 and 5.6 show overall system spectral efficiency improvements. This framework can be also combined with studies in [MET14-D53] that propose utilization of geo-location database for the short term ultra-reliable communication for V2V.

Our investigations focus also on proving that D2D communication can be used with no (or limited) deterioration for ongoing cellular transmission. This is also visible in V2X studies exploiting direct transmission between devices. The context-aware resource allocation scheme for enabling D2D in moving networks fulfils its design goals and offers significant performance improvement as compared to the SOTA. However, meeting the METIS requirements for V2D communication necessitates further enhancements in the 5G system. V2X schemes utilize location information, CH operations or smart RRM algorithms. The latter two were also compared against 802.11p showing higher performance. We've shown that we can not only satisfy the stringent latency and reliability requirements of safety critical V2X communication, but also bring benefits to cellular users by reusing the spectrum which has already been allocated to vehicular communication.

6 Dynamic Reconfiguration Enablers

6.1 Introduction

Network and deployment flexibility and the ability of reconfiguring network nodes are considered as one of the key enablers for the 5G networks. It is of the uttermost importance for the future cellular networks, primary due to the diverse requirements as well as the foreseen access node densification, which is expected in order to accommodate the massive traffic growth in 2020 and beyond. Undeniably, increasing the number of access nodes brings several inevitable side effects. From the economic and environmental point of view one of the most noticeable repercussions is the increased overall static (related to 'stand by' operations) and dynamic (related to data transmission/reception and processing) energy consumption [MCC+09], [Ols12]. Already nowadays the telecommunication industry is responsible for consuming around 1% of the overall global energy intake [Roy08], and this share in the future will become larger, considering the growing usage of cellular mobile technologies. This affects not only the ecological aspects of cellular networks, such as CO_2 footprints, but also leads to increased OPEX experienced by operators.

Additional economical and technical challenges related to the network densification are efficient provision of backhaul for access nodes [GCG+14], [BJM+14] and dynamic configuration of network elements. Solutions aiming at solving these problems are necessary for optimal handling of the time varying traffic demand.

A particularly new aspect of the future 5G systems are nomadic nodes (NNs) – cells with wireless backhaul that can change their position, as their operating antennas are mounted on vehicles such as cars, buses, trams or trains [3GPP14-36836], [SPS12]. This new infrastructure mobility aspect triggers several concerns for network planning and RRM schemes, but it is also a huge opportunity for future systems to provide additional capacity exactly where it is needed - for the people travelling with or in proximity of these vehicles.

Support of dynamic network topology may be also achieved using clustering. In clustering operation, we group elements of the radio network with similar properties. Both UE and radio nodes can be clustered. By grouping the entities based on their similarities, we can better manage input and output data related to the operation of such a bundle. Additional challenges related to clustering operations are the dynamic behaviour of users (in the sense of mobility and traffic patterns) as well as heterogeneity of the network and user demands [Bas99], [Don00], [BDG+08]. Some examples of clustering mechanisms related to cluster head operations can be found Sections 2.14, 4.11 and 5.7.

This chapter covers several examples of solutions proposed by the METIS project, which allow dynamic reconfiguration of the network. First three technology components evaluate schemes for activation/deactivation of access nodes. Next ones focus on nomadic nodes and different aspects of clustering. Finally, considerations on selected network interfaces are given.

6.2 Activation and Deactivation of Small Cells in UDN

The proposed dynamic activation and de-activation of small cells in UDNs facilitates the dynamic re-organization of available resources, according to users' varying traffic volume and necessary service requirements. It is proposed to provide network connectivity at all desired locations and sufficient bandwidth to satisfy clients' communication needs, reducing the underutilisation of network resources and saving energy. Small cells' dynamic activation and de-activation is decided based on the monitoring reports from both small cells and associated user equipment. This information includes UL/DL resources that are used (e.g., reserved resource blocks and available capacity) as well as wireless network graph. The latter describes graph that is formed among small cells and user equipments according to the overlap of their transmission range. The capacity usage ratio of each cell (i.e. capacity

indictor) and the overlapping factor of cells (i.e., coverage indicator) are calculated based on the above-mentioned collected information and are used for the decision making. The overlapping factor provides an indication of the overlap of the transmission range of neighbouring cells and it is calculated using graph theory. The de-activation of small cells is triggered when there is too high coverage overlap, while the capacity usage ratio is too low. In the case of low coverage, low channel quality indicator (CQI) or even high blocking probability, the activation of one or more small cells is checked, with the goal to provide more resources wherever it is necessary. Both actions are followed by the handover of a set of UEs.

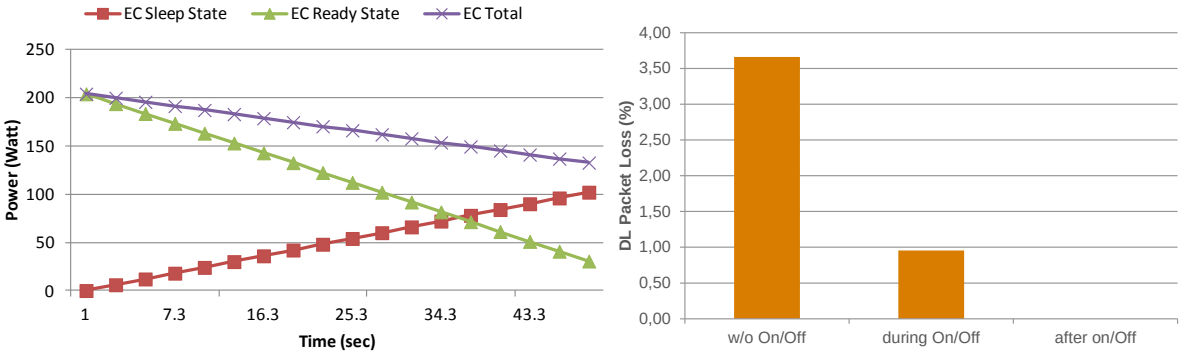


Figure 6-1. (a) Small cells energy consumption (EC) for the simulation time (sec), (b) DL packet loss

In contrast to the other available solutions the proposed scheme takes into account additional context information as well as the formed network graph for the selection of the appropriate small cell to activate/de-activate. A topology of 20 small cells and 10 UEs has been deployed in order to evaluate the network performance with and without the usage of the proposed dynamic activation/de-activation mechanism. The simulation results for the de-activation case show significant gains in terms of energy consumption. Figure 6-1 depicts the total energy consumption as well as the consumed energy for the sleep state (deactivated small cells) and the ready state (active small cells) during the simulation time (i.e., axis x). The total consumed energy for the sleep state of all small cells of the topology increases as more small cells are de-activated, while the opposite happens for the energy consumed in the context of the ready state. In addition the de-activation of small cells leads to the decrease of the detected packet loss for the DL, due to the reduced inter-cell interference and reduced number of handover events [YSC+13].

6.3 Energy Savings Schemes for Phantom Cell Concept Systems

Phantom Cell Concept (PCC) systems follow a network architecture paradigm in which UE can connect to both a macro and a small cell in a dual connectivity fashion, as illustrated in Figure 6-2(a). This dual connectivity feature offers the possibility to flexibly put unused small cells to a sleep mode in which they consume a reduced amount of energy but cannot serve any user [TAHD14]. However, connectivity to the cellular network is not disrupted in the process, since users keep a connection to the macro cell. Furthermore, the connection to a small cell can be performed in a macro-assisted fashion, where the macro cell manages UE-small cell connections in a centralized way, thereby potentially optimizing resource utilization in the small cell network. This technological component evaluates three macro-assisted small cell energy savings schemes: a downlink signalling-based scheme, an uplink signalling-based scheme, and a database-aided scheme.

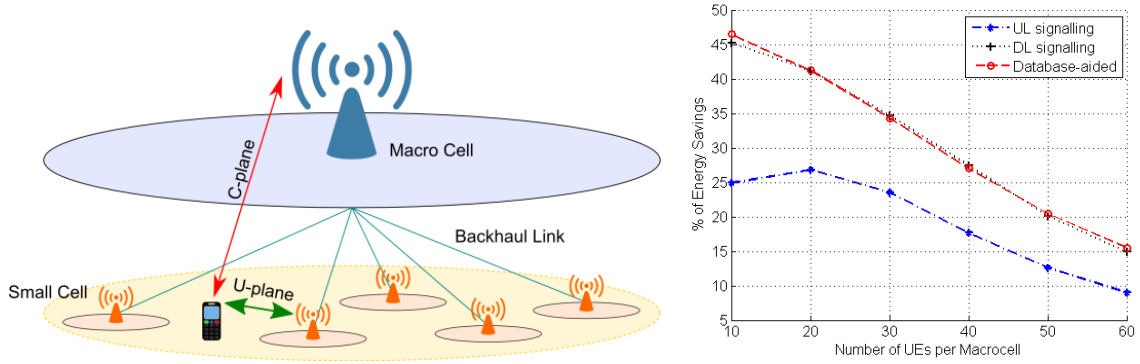


Figure 6-2. (a) Representation of the PCC network architecture, (b) Percentage of energy savings obtained with the various schemes

Small cell energy savings schemes currently discussed in 3rd Generation Partnership Project (3GPP) are focused on the case of networks in which macro cells and small cells operate independently. Very little attention has been paid to macro-assisted schemes where a master-slave relationship is present between macro cells and small cells. Due to their centralized nature, macro-assisted connection schemes offer a large number of advantages with respect to conventional connection schemes. One example is the possibility to select the best small cell for a UE to connect to, not only based on the small cell channel estimations performed by UEs, but also on additional information sent by small cells via the backhaul link (e.g. the load of each small cell).

Detailed signalling of the three energy savings schemes can be found in section A.5.1 of this document. Performance results of the considered schemes in terms of energy savings with respect to a baseline scheme, where small cells are always turned on, can be observed in Figure 6-2(b). Up to 45% of energy consumption reduction can be achieved when a low number of users are present in the network when using the downlink signalling-based scheme or the database-aided scheme.

6.4 Activation and Deactivation of Nomadic Cells

This technology component describes the mechanism and the algorithms for activating and deactivating nomadic relays. A nomadic relaying network is a network with vehicle mounted relays acting for potential relaying functionalities. It extends the legacy network by randomly located non-operator deployed relay nodes. Compared with the traditional relay nodes, the nomadic nodes do not require site-leasing, and therefore no site planning is in question. On the contrary, the large amount of unplanned nodes needs a high effort in coordination and management. This technology component deals with the architectural designs that enable the activation and deactivation of nomadic cells. As part of the design, optimization algorithms are proposed to satisfy the minimum quality of service requirements of all the users in the network. Furthermore, both centralized [RSF14] [RSF+14] and distributed [RSS14] algorithms are proposed for energy saving purposes.

Figure 6-3 depicts the architecture and information flow for the management of nomadic network. A central control unit decides the connection assignments between the base stations, relays and users based on the feedback information on channel states and user or nomadic nodes profiles. In the case when a central control unit is not available, users and NRs are autonomously connected and disconnected based on their own measurements. The assignment list will be delivered to the users and the nomadic nodes through base stations, upon which the nomadic cells and users send the corresponding connection or handover request. Then, admission control will be performed at the base station and nomadic cells to allow or reject the connection requests.

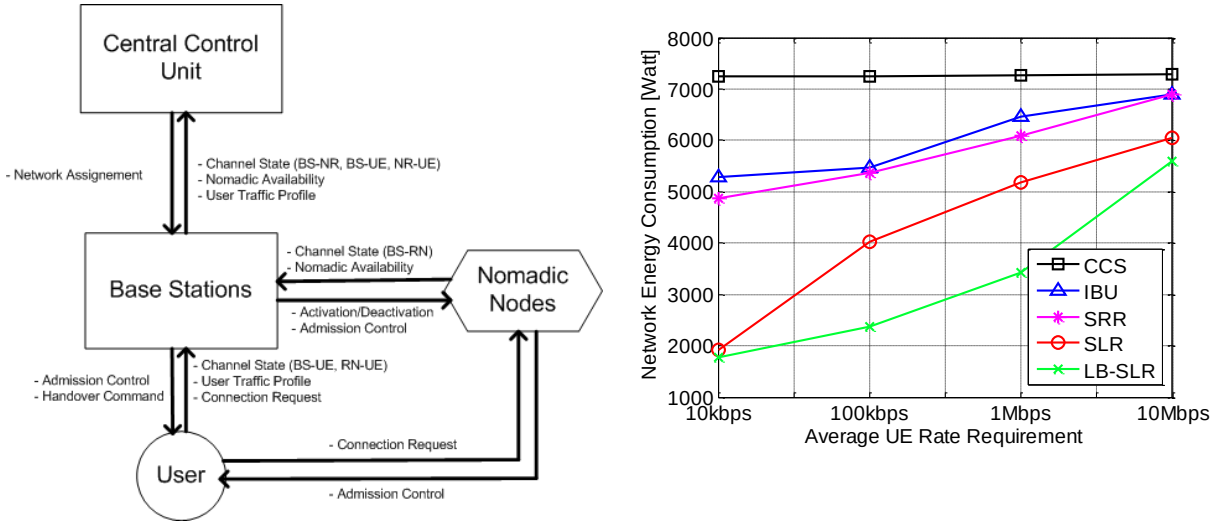


Figure 6-3. a) Architecture for activation/deactivation, b) Energy saving performance

Based on the collected channel information and the user's QoS, an optimization framework can be formulated with the network assignments as input argument. The decision of activation or deactivation can be made based on the optimal assignment, such that if no users or relays are assigned to the cell, the cell can be deactivated. During the whole procedure, there are two general requirements for the network optimization: 1) All nodes must be connected to the network. 2) The amount of radio resources is limited. Based on these two constraints, a general optimization framework can be formulated for energy savings or other utility functions of throughput or SINR. Example results for energy savings are shown in the right of Figure 6-3, where three assignment schemes are displayed, including a conventional algorithm (CCS), an algorithm proposed in [RSF14] that consider worst-case interference (IBU) and an algorithm proposed in [RSF+14] considering dynamic interference (SLR). A great potential of energy saving in the nomadic network can be observed in Figure 6-3, particularly when the average rate requirement is low (less human activity, e.g., in the night or a in sparsely populated area).

6.5 Dynamic Nomadic Node Selection for Backhaul Optimization

One of the targets of the 5G system is a flexible deployment, where inhomogeneous distribution of traffic demand over time and space shall be cost efficiently covered in an agile manner. In this regard, demand-driven temporal network densification via NNs within the dynamic RAN framework [MET15-D64], [MET15-D66] is a promising enhancement. NNs require flexible backhaul to provide broadband access on demand, for which one cost-efficient realization is in-band relaying. The capacity of the wireless backhaul link between an NN and its serving BS has a crucial role in the achievable E2E performance, especially when limited by severe fading characteristics. Flexibility of dynamic NN selection can be exploited to overcome the limitations of the backhaul link. Proposed method introduces a mechanism for identification of the optimum serving NN based on the backhaul link quality, namely SINR. The method exploits the diversity via the availability of multiple NNs in a confined region, e.g., a parking lot; the coarse NN selection takes into account long-term channel quality measurements based on shadowing, whereas, the optimal selection relies on the short-term channel quality measurements based on both shadowing and multi-path fading.

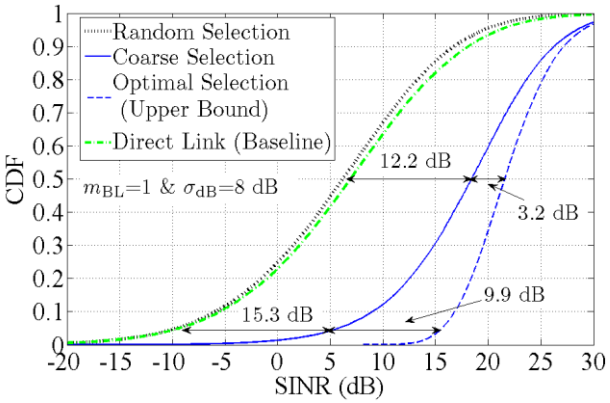


Figure 6-4. SINR CDFs; SINR improvement on the backhaul link via NN Selection Schemes

Figure 6-4 shows the CDFs of the SINR on the backhaul link and direct link (user directly connects to the BS) in a scenario with severe fading characteristics (Rayleigh-lognormal with 8 dB shadowing standard deviation). Relative to both random NN selection and direct link, coarse NN selection provides significant SINR gains of about 15 dB and 12 dB at lower and median CDF percentiles, respectively. Optimal NN selection can further improve the performance at the cost of increased signalling overhead. The SINR gains can translate into E2E throughput gains depending on the access link quality [BRZ+14].

6.6 Nomadic Node Attachment Procedure for Integrated Fixed Backhaul Management

The goal of the proposed approach is to allow an existing mobile network infrastructure to dynamically adapt to a sudden increase in service demands without overprovisioning the network. By analysing, e.g., the test case Traffic Jam [MET13-D11] in dense urban environments, the traffic load created by NNs, such as cars or buses, during the traffic jam can result in an increase of up to 30% in the backhaul link demands of a macro site only [SS14]. This fact implies the need for over-dimensioned fixed backhaul networks, e.g., due to the traffic jams in busy hours, which can be avoided by employing the proposed solution. In this way, an NN requesting access to the mobile network could be either attached or detached depending on the existing transport network resources.

In the algorithm description below, by definition, NNs are the ones which are nomadic and request access to the mobile network, anchor nodes are the ones which are already attached to the existing mobile network infrastructure which include either NNs (in case of mesh networks) or the stationary BSs already deployed. The network control entity in this case can be located in, e.g., a data center, being part of a Software Defined Network (SDN) transport controller or an end-to-end (E2E) network orchestrator, as discussed in [MET15-D64].

The presented solution couples dynamic link requirements in the RAN (e.g. moving networks) with the transport link capabilities of the network infrastructure. Furthermore, the solution makes use of the evolution in transport network architecture solutions, which is going towards flexible and programmable networks. In this work, a SDN based Dense Wavelength Division Multiplexing (DWDM)-centric optical transport network was considered [PO+13][BB+15]. Such architecture allows optimizing the problem of NN integration to the fixed network infrastructure with the possibility of on demand reallocation of transport network resources, i.e. optical wavelengths, to adapt tithe corresponding backhaul link capacities [SS14].

Algorithm description:

Step 1: Attachment REQUEST (Nomadic Node 'N')
SEND information about actual data traffic/service demands to the nearest anchor node(s).
Step 2: Acknowledgment (Anchor node redirects request to Network Control Entity)
CHECK new node N connection allowance based on the existing transport network resources
IF network (local and/or global) is able to fulfil 'N's requirements THEN
ACCEPT attachment. END
NOT able to fulfil 'N's requirements THEN
ALLOCATE additional network resources
IF resource allocation is NOT available
DENY attachment. END
IF resources allocation is possible
ACCEPT attachment. END
END
Step 3: SEND final decision (Network Control Entity)
Attachment ACCEPTED or DENIED.

6.7 Dynamic clustering

The objective of this approach is to create cell clusters for each UE and to select a power budget to be used by the cells of each cluster. The signalling of a particular UE is obtained after a joint processing in all the cells of the cluster. The power budget is optimized to ensure certain QoS in the presence of interference. As a consequence, the joint processing is done independently of the rest of UEs, that is to say, no interference cancellation technique is used, since interference is mitigated by means of the power optimization.

The inputs of this approach are the SNR values from all cells to all UEs. No channel coefficients are required by the clustering algorithm. However, for an efficient joint transmission to the UEs, the cluster members are assumed to know the channel coefficients. This implies that only a limited amount of control data is interchanged since the channel coefficients are not spread to many more cells in a larger area.

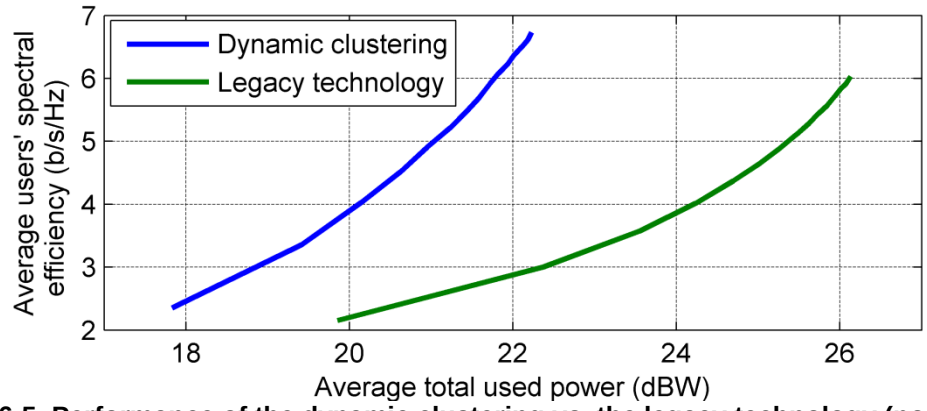


Figure 6-5. Performance of the dynamic clustering vs. the legacy technology (no clustering)

In a scenario with a fixed number of cells and compared against current standards in which no clustering is performed, the proposed approach has a 12% of spectral efficiency gain together with a reduction of the 60% of the total used power. The power reduction, which is due to the power budget optimization, allows the deployment of more cells with no extra power consumption. In a scenario in which the number of cells for the dynamic clustering is increased until the power consumed equals that consumed by the legacy technology, the

spectral efficiency is multiplied by a factor of 2. Figure 6-5 shows the spectral efficiency in terms of the average total used power for the second scenario.

6.8 Overlapping super-cells for dynamic user scheduling across bands

The objective of this approach is to create overlapped cell clusters (super-cells) to ensure that all UEs are in the centre of a super-cell. This way, the low SNIR values of cluster edge users can be avoided. Each individual cell will belong to several super-cells, in which it will operate in a different band (or set of resources). In particular, cells will belong to a number of super-cells that equals the number of bands, so that the reuse factor is 1 for all cells. However, super-cells that overlap in a particular cell will operate in different bands to prevent interference.

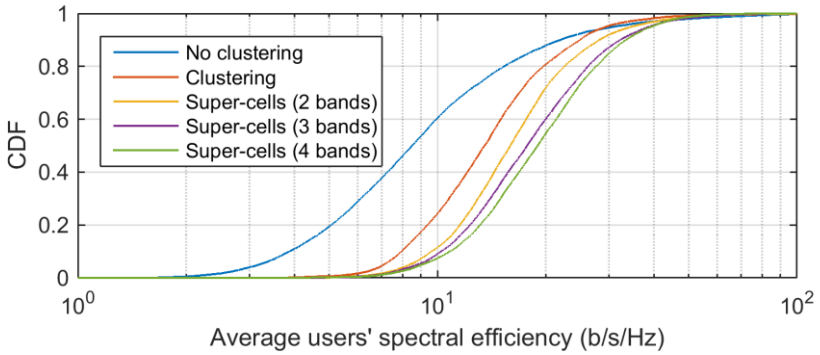


Figure 6-6. Performance of the overlapping super-cells vs. no clustering and a non-overlapping clustering

Super-cells are obtained using hypergraph partitioning algorithms [SK06]. This approach is independent of the underlying CoMP or eICIC technique used inside each super-cell.

Figure 6-6 shows the CDF of the spectral efficiency of the proposed approach for 2, 3 and 4 bands, and two additional approaches with no clustering and a non-overlapping clustering (obtained with hypergraph partitioning but for only one band). Each super-cell transmits to their UEs using dirty paper coding and with a fairness maximization objective [CGY14]. The signalling processing is done without taking into account the interference of the other co-band super-cells. In the case of 4 bands, the average users' spectral efficiency is multiplied by a factor of 1.4 with respect to a non-overlapping clustering, and by a factor of 2 with respect to no clustering. The 5th percentile is multiplied by a factor of 3 with respect to no clustering.

6.9 Clustering Toolbox

The need for a proper coordination of nodes in mobile networks, where the number of devices is constantly changing, is important for network management of future communication systems (e.g., MTC, D2D). In this context, clustering mechanisms are investigated so as to examine their potential application and impact on nodes' coordination schemes. Such mechanisms focus on the grouping of nodes into smaller clusters and grouping nodes with high similarity in the same cluster. The high signal quality from neighbouring nodes and the close distance among nodes are two of the criteria for similarity identification in a mobile networking environment. At the same time, moving networks are being considered as potential enhancement for mobile communications. The characteristics in such deployments differ from the ones in typical cases; nonetheless, the requirements remain the same. Therefore clustering mechanisms for both static and moving networks need to be examined. In order to select the most promising algorithm for each case a variety of clustering algorithms are tested and evaluated based on the clusters they result in. The application of clustering mechanisms is expected to facilitate the network be means of i) efficient spectrum allocation inside smaller groups of nodes, ii) energy saving from access point activation/deactivation and iii) efficient support of moving nodes. The presented results are based on the new Aggregate Local

Mobility (ALM) Clustering Algorithm for VANETs [SNG10]. Other algorithms that have been taken into account are [HSZ+14] and [RM12].

Table 6-1. Number of formed clusters

Number of moving nodes	Low Transmission range (~50m)	Medium Transmission range (~70m)	High Transmission range (~160m)
10	3	2	2
20	7	5	3

Table 6-2. Delay (sec) for the formation of clusters

Number of moving nodes	Low Transmission range (~50m)	Medium Transmission range (~70m)	High Transmission range (~160m)
10	1,5 sec	0,32 sec	0,41 sec
20	0,61 sec	0,58 sec	0,6 sec

Table 6-1 and Table 6-2 present the experiments that were conducted for clusters formation using a moving nodes scenario. For each scenario the number of formed clusters and the delay to form the clusters has been measured. The initial inter-nodes distance is 20 m and each node has a randomly distributed velocity in range [20, 30] m/s. Clusters formation is based on the distance the moving nodes and the modularity of the formed graph. As expected the number of formed clusters is inversely proportional to the transmission range of each node, while the delay to form the identified clusters increases with the increase of the number of moving nodes. Existing 3GPP networks have not incorporated mechanisms for the formation of clusters. The latter could contribute to the improved coordination of nodes as well as to facilitate the exchange of time critical information without network assistance (e.g., vehicular safety).

6.10 New Management Interfaces Between the Operator and the Service Provider and Interfaces for Information Exchange and Action Enforcement

We expect that 5G scenarios will integrate a variety of technologies and solutions. Many of these solutions converge on the need for exploitation of context information from different networks and service infrastructures thus posing requirements for network flexibility as well as for integrating information from the mobile network operator's (MNO) elements and external service providers (SPs). Information coming from the SPs will enhance the network management and traffic engineering according to the actual service requirements.

In the following we elaborate on architectural considerations for new management interfaces improving network performance based on MNO and SP collaboration and we highlight enhanced interface requirements and information flows for such collaboration through analysis of technology components evaluated in this report in MTC and UDN scenarios.

Figure 6-7 depicts involved entities and required interfaces for enabling MNO-SP interactions and flexible network configuration. Notably, this integrates building blocks (BBs) and concepts defined in [MET15-D64]. Network entities such as the User Device and the Infrastructure Radio Node incorporate the Radio Node Management BB (METIS RNM) whilst the Network Manager incorporates the Novel Network Interface Termination BB as capturing the MNO-SP management interface control points. Information on the mentioned building blocks can be found in [MET15-D64].

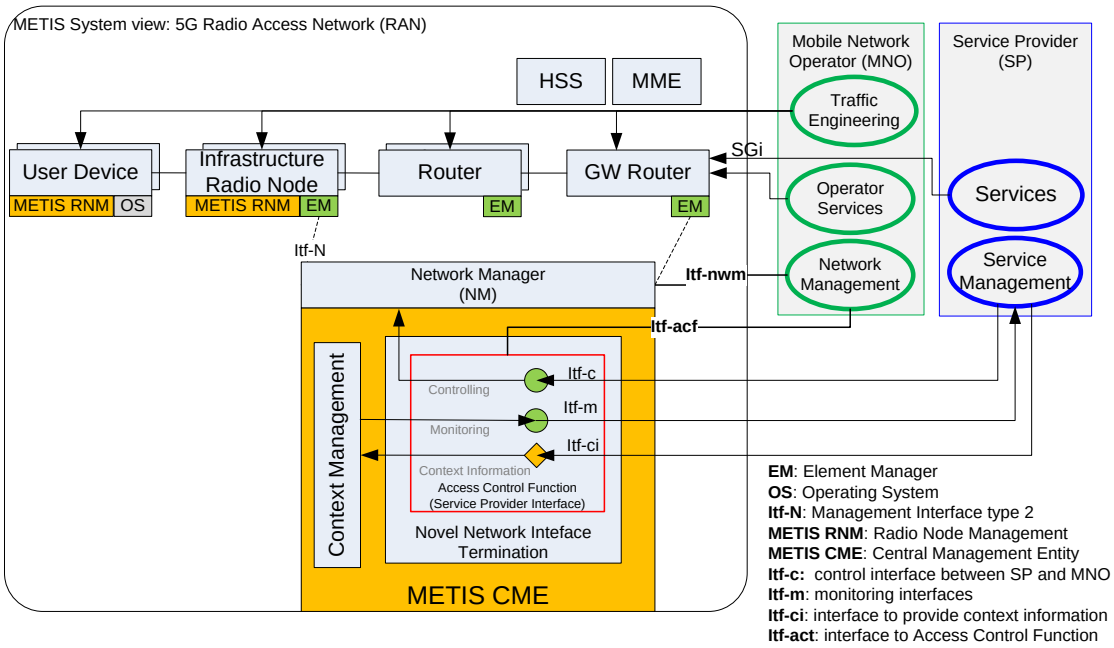


Figure 6-7. Interfaces between MNOs and SPs

Figure 6-7 depicts context information exchange and actions enforcement for METIS solutions utilising service information which may end up to the decision points from the UEs, or directly from Application Service Delivery (ASD) servers. Analysis of the service information exchange is focusing at UDN deployments, and massive machine type communication. The wireless communication may take place either through the network infrastructure or via direct links (D2D communication).

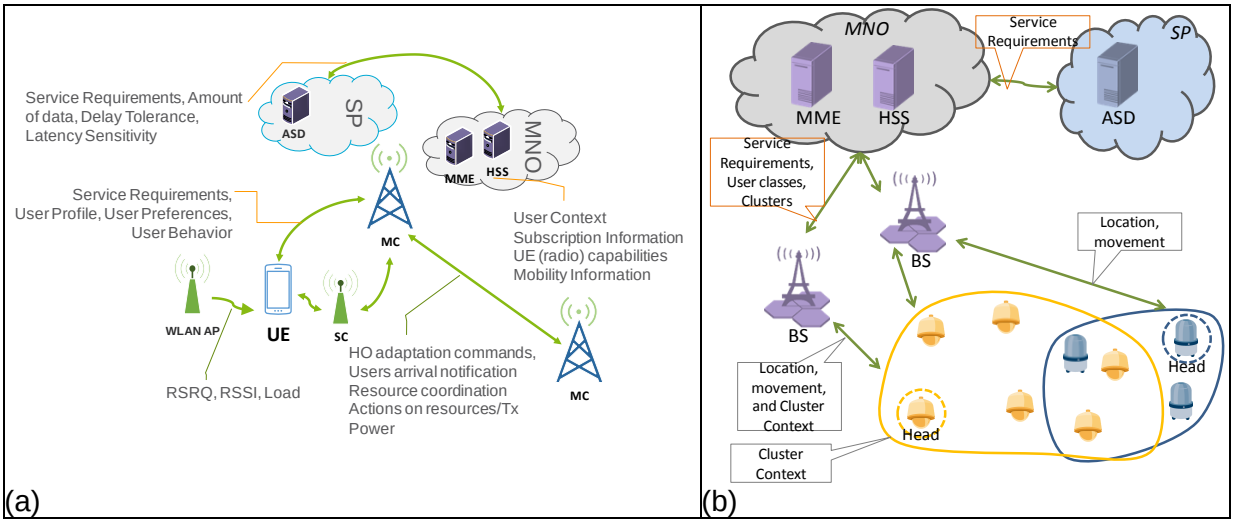


Figure 6-8. Information flows including interactions with the service providers for (a) ultra dense deployments (b) machine type communications

From the METIS considered technology solutions service information is used mainly for resource allocation (e.g., RACH preambles, service mapping to connectivity type, etc.) and RAT/Layer mapping. Certain aspects of service information have been exploited (as in technology components provided by this report) such as service and QoS requirements (e.g. amount of data, delay tolerance) and service priorities. Further, specific consideration has been carried out on those METIS solutions incorporating service provision aspects; such consideration indicated that the delay tolerance is a key service requirement to be taken into account by the MNO for resource allocation aiming at the appropriate service provision for

both UDN and MTC scenarios [MET15-D64]. Additional information required is posing interface requirements including user context information, UE capabilities, cluster context information as well as commands, notifications and resource coordination actions. Such information flows and interface requirements are indicative within UDN and MTC scenarios and they have been extracted by corresponding research technology solutions (reported in [MET14-D42]).

6.11 Interfaces for Context Exchange in RANs and Among Operators

In future mobile networks, a much broader scope of services and applications will have to be supported, which may result in more challenging requirements with respect to reliability, data rates, energy efficiency, spectral efficiency, latency, coverage, mobility, etc. In order to enable mobile network operators to perform smart resource mapping and facilitate context-aware mobility support and spectrum sharing across RANs and operator domains, several new kinds of interfaces are proposed:

- Interface between terminal (UE, MTC device) and access node (e.g., eNode, NodeB, Access Point (AP), 5G access node),
- Interface between access nodes,
- Interface between terminals (e.g., for non-network assisted cases).

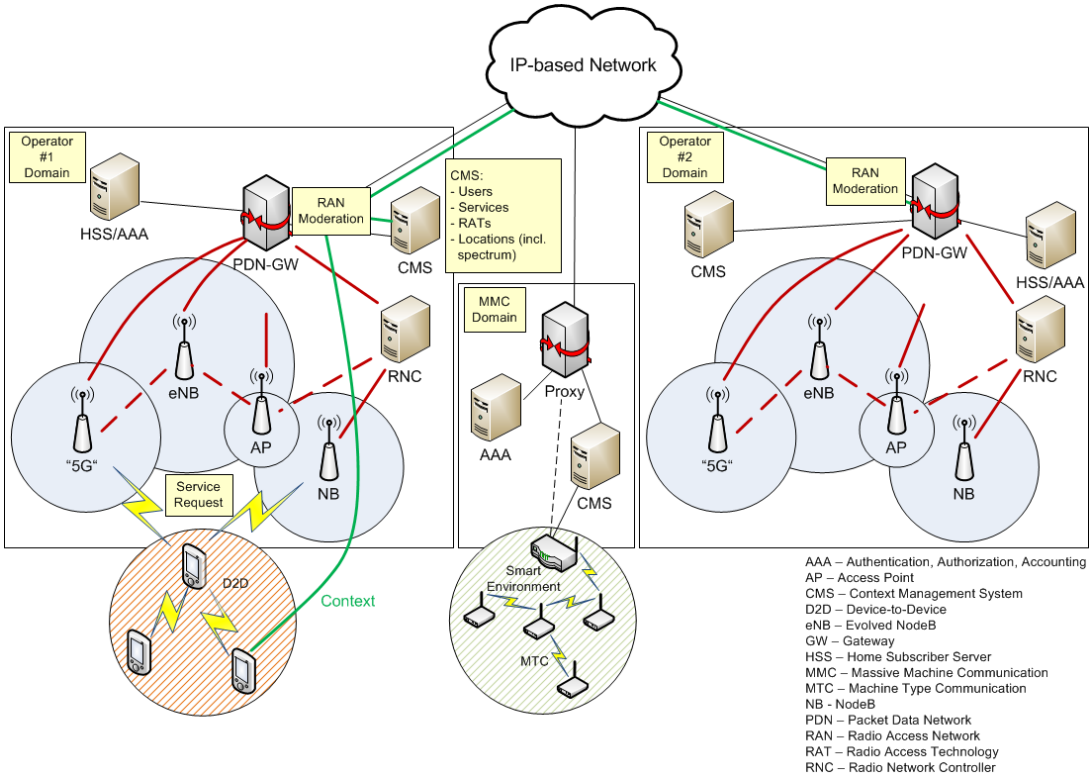


Figure 6-9. Concept of intra- and inter-operator RAN moderation (Network entities provide interfaces for context information exchange. "Local" context is exploited for RAN moderation within and across operator domains)

The context information to be exchanged via these interfaces covers various static and dynamic attributes of the terminals and access nodes relevant for improved service provisioning in potential multi-operator scenarios. These context information may be hosted by domain-specific Context Management Systems and represented using Extensible Markup Language. Examples of solutions exploiting context awareness can be found in Section 4.

6.12 Impact on the METIS 5G System Concept

Proposed solutions were analysed from the point of view of xMBB services and are targeting at flexible and dynamic provision of high data rate connections, exactly when and where it is needed. They allow for cost efficient operations of both UDN and NN. Especially in the area of network energy efficiency, proposed solutions can provide a great improvement when comparing to legacy systems.

Clustering solutions can be employed in all main services proposed by METIS. M-MTC and U-MTC could benefit from e.g., CH operations, whose benefits were discussed in Sections 2.14, 4.11 and 5.7. xMBB benefits mostly from the coordination of RRM and interference alignment. Clustering is also required for many of the schemes proposed in METIS, e.g. CoMP [MET15-D33], coexistence-aware resource allocation for extremely close nodes, or base station clustering for inter-operator spectrum sharing under realistic network deployment. Description of the two last solutions can be found in [MET14-D53]

6.13 Addressed METIS Goals

The prime area where the mechanisms presented in this chapter can help the future solutions to meet METIS objectives is the provision of 1000x higher capacity at the similar cost and energy efficiency. Provision of high capacity via network densification is feasible (RRM and mobility challenges in UDN are introduced in next chapters), but comes at a high cost. It can be lowered primarily by dynamic activation and deactivation of network elements (exact gains depend on the traffic dynamics in individual cells) as well as through shifting the capacity/coverage challenges from fixed network infrastructure onto NNs (efficient wireless backhaul, also tackled in this chapter, can be treated as one of the major enablers for moving networks).

6.14 Conclusions

Economical challenges related to network densification (provision of low cost and efficient backhaul and most of all energy footprint) are emphasising the network solutions that help us to dynamically activate and deactivate specific access nodes, depending on the instantaneous traffic demand. In order to tackle the side effects (handovers, cell reselection, cell discovery), a control/user plane decoupling could be exploited (e.g., macro operating the control plane and small cells used user plane - high data rate transfers). One proposal is macro-assisted small cell energy savings schemes, enabled by the PCC architecture, which allows to put a large number of small cells to sleep without disrupting the user's connectivity, thereby significantly decreasing the energy consumption of the network, and potentially increasing the QoS offered to users, due to the reduction of the experienced interference level. The viability of energy efficiency schemes based on dynamic activation and deactivation of access nodes has been proven in a number of independent METIS studies, including the ones involving nomadic cells. NN based flexible deployment is a promising 5G technology that enables an agile and demand-driven network extension [BRZ+14]. To this end, dynamic NN selection schemes exploit flexible backhaul link and the availability of multiple NNs to provide significant gains on the backhaul link and, thus, overcome backhaul limitations on the end-to-end communication link [BRZ+14] [RSF14].

Clustering is one of the main enablers to provide to 5G mobile operators a flexibility to adapt dynamic network technology. Clustering mechanisms can be used for dynamic creation of cooperating nodes. Described clustering mechanisms reduce interference, and allow the network to serve UEs with the same rates but using much less power. Thanks to this, denser networks can be deployed with no extra power usage. The advantage of denser networks is that the base stations are closer to the UEs, and this drastically increases the UEs' rates. Additionally user fairness is maximized with overlapping clusters that minimize the UEs in a cluster-edge situation. Clustering toolbox proposed in this chapter can be used to select the most appropriate clustering algorithm for required application.

7 Summary

One of the biggest challenges, which is often overlooked when discussing requirements and scenarios for 5G, is the operations in a highly heterogeneous environment. This heterogeneity will manifest in various forms. Devices of IoT are going to have completely different requirements than users watching ultra-high definition videos, not to mention V2X services. Future deployments will comprise of cells of different coverage, capacity, transmit power, operating frequency, etc. Finally, previous generations of telecommunication standards, developed by 3GPP or IEEE, will still play a vital role in 2020 and beyond. This complicated technological ecosystem has to be exploited in order to provide highest E2E performance for the future society.

Predicted network densification, direct D2D and dynamic UL/DL air interface in TDD mode (necessary in order to provide maximal utilization of available spectrum) will result in challenging interference constellations. Therefore proposed mechanisms which help us to identify sources of interferences, out of which Multi-Kernel based method shows best performance, are of a high importance. When it comes to RRM mechanisms for the future networks, we observe reasonable gains of more centralized schemes that show their full potential in an unplanned deployment of UDN. In centralized schemes macro could play a coordinator role, distributing scheduling decisions and collecting necessary information from small cells within its coverage. In order to provide reliable and timely signalling over the air signalling should be used when backhaul can't provide sufficient performance. Over the air signalling from macro can be also used to provide a synchronization to the small cells that is a necessary enabler for schemes targeting at the reduction of intercell interference. Reduction of the intercell interferences by proactive interference management is one of the most widely discussed topics in this report and a number of mechanisms were proposed that target different scenarios: utilization of TDD spectrum for coordination of FDD operations, using fountain codes to provide a high data rate connection for the vehicles on the move, exploitation of JT-CoMP with no CSI at the receiver, SON-like operations based on resource auctioning, or replacing access based on OFDMA with SCMA. Initial studies of the latter option show visible performance gains over OFDMA, especially when comparing MIMO at higher speeds, when closed loop information fails due to the channel fading. However, final comparison of these two approaches needs all encompassing studies including similar feature set (MU-MIMO, CoMP), comparable receiver types, complexity etc. Additional RRM aspects that should be further investigated are the impact of mmW technology (especially in combination with large antenna systems) and implications of spectrum sharing methods on suitable resource allocation schemes.

One of the key challenges of the highly heterogeneous deployment is handling of mobility. We've proved that it is beneficial to utilize information about load and capacity of cells as well as delay tolerance of the transmission in order to schedule data-heavy traffic (e.g., software updates or release of popular content) in underutilized cells offering favourable transfer conditions. Information about the load can also be exchanged using user devices. We've also demonstrated significant energy saving gains when exploiting radio fingerprints broadcasted by the network that can help the devices to detect small cells in the proximity. Additionally, we've shown that utilization of radio measurements coming from the sub 6 GHz bands for detection of the mmW cells significantly alleviates a cumbersome detection of access nodes operating on higher frequency bands. Last, but not least, it was demonstrated that the performance of users moving inside the vehicles is enhanced by the utilization of moving relay nodes.

A vast amount of technology components focused on the development of mechanisms that exploit context information. Performed research shows that it is a very promising approach in order to provide customised user experience as well as improving system metrics. Solutions showing good performance gains are based on prediction of user position. Its usefulness was

proven in reality for the scenarios with users watching a video transmission moving with a car (GPS related measurements were used). In order to harvest these gains by indoor users, we also recommend to research positioning algorithms and methods that work inside buildings. Context awareness when used to group devices of IoT allows avoiding superfluous data transmission. Using this approach and cluster head operations we've managed to accommodate far greater number of devices than existing 3GPP solutions. Finally, context awareness benefits SON-like operations that adapt different mobility related parameters (handover settings) that are adapted based on different capabilities of user devices. This hypothesis was validated by simulating a number of different algorithms aiming to reduce the number of handover failures, call drops, etc. Still, context awareness operations can't be blindly carried out across the network as it may lead to a significant signalling exchange. Further trade-off investigations for increase in signalling overheads and achieved gains are hence necessary

One of the most widely exploited research areas where we've managed to provide a holistic set of solutions is D2D domain. Gains of this mode of operations (proximity, hop or reuse gains) are unquestionable and in our studies we've focused on how to introduce this technology without harming transmission of cellular users. The set of aspects that we've addressed ranges from efficient device to device discovery (we've proposed mechanisms that work with and without network coverage), through appropriate resource allocation including mode selection (choice between direct or indirect form of transmission, using orthogonal radio resources or reuse of the ones used already for ongoing cellular transmission) and ending on power control mechanisms that allow transmission with far less power than SOTA solutions. A holistic analysis was provided for fast vs slow resource allocation and centralized vs decentralized D2D operations that could be used for mapping of selected methods for appropriate applications (although fast and centralized mode selection and resource allocation performs better, it also requires higher signalling overhead which is a limiting factor for several services). Increased interest of automotive industry in 5G is also reflected in our studies. We've provided a set of solutions and algorithms that address this area, by using a smart resource allocation that enable high performing V2X data exchange without significant degradation of cellular users. Also operations exploiting cluster head entity showed a promising performance in efficient (compliant with requirements for V2X services) distributions of available radio resources for devices in proximity. The area which was identified as missing in order to provide a holistic vision for D2D and V2X are the activities in inter-operator setup and in spectrum sharing scenarios.

One key area where studies captured in this deliverable contribute to the development of 5G is the reduction of environmental footprint (and costs of the energy consumption) of the network infrastructure. A straightforward approach to achieve this goal is a native support of dynamic activation and deactivation of unused access nodes. Two ways forward proposed in this report are the exploitation of control/user plane split with dual connectivity and nomadic nodes.. Control/user split allows performing activation/deactivation of access nodes without mobility procedures needed for cell reselection or handover. On the other hand, nomadic nodes will help the network operators to provide a high performing connection to the users inside or in proximity of the vehicle, while deactivating fixed cells that would be used instead. Wireless backhaul connection can be easily provided using large antenna systems that could be mounted on the vehicles without constraints on antenna size and their separation known e.g., from handheld devices. Exploitation of nomadic nodes may become a true revolution brought by 5G that breaks the established paradigm of users being bounded to the coverage of fixed access infrastructure. Using nomadic nodes, it is the network that follows users, not the other way around.

8 References

[3GPP10-36304] 3GPP TS 36.304, “User Equipment (UE) procedures in idle mode (Release 9)”, September 2010.

[3GPP11-37868] 3GPP TS 37.868, “Study on RAN Improvements for Machine-type Communications (Release 11)”, September 2011.

[3GPP12-36423] 3GPP TS 36.423, “X2 application protocol (Release 12)”, December 2014.

[3GPP13-36843] 3GPP TR 36.843, “Feasibility Study on LTE Device to Device Proximity Services - Radio Aspects,” June 2013.

[3GPP13-36888] 3GPP TR 36.888, “Study on provision of low-cost Machine-Type Communications (MTC) User Equipments (UEs) based on LTE (Release 12)”, June 2013.

[3GPP14-23203] 3GPP TS 23.203, “Policy and charging control architecture (Release 13)”, December 2014.

[3GPP14-36133] 3GPP TS 36.133, “Requirements for support of radio resource management (Release 12)”, December 2014.

[3GPP14-36331] 3GPP TS 36.331, “Radio Resource Control (RRC); Protocol specification (Release 12)”, December 2014.

[3GPP14-36355] 3GPP TS 36.355, v12.2.0, “Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Positioning Protocol (LPP)”, July 2014.

[3GPP14-36836] 3GPP TR 36.836, “Study on mobile relay”, June 2014

[And05] J. Andrews, “Interference cancellation for cellular systems: a contemporary overview,” IEEE Wireless Communications, vol. 12, no. 2, pp. 19–29, 2005.

[Bas99] S. Basagni, "Distributed Clustering for Ad Hoc Networks", Fourth International Symposium on Parallel Architectures, Algorithms, and Networks, 1999. (I-SPAN '99), Proceedings, p.p.: 310-315, 1999

[BB+15] B. Skubic, G. Bottari, A. Rostami, F. Cavaliere, P. Öhlén, “Rethinking Optical Transport to Pave the Way for 5G and the Networked Society. Lightwave Technology, Journal of (Volume:PP , Issue: 99), Jan 2015.

[BCG09] P. Bellavista, A. Corradi and C. Giannelli, "Mobility-aware Management of Internet Connectivity in Always Best Served Wireless Scenarios," Mobile Networks and Applications, vol. 14, pp. 18-34, 2009.

[BDG+08] U. Brandes, D. Delling, M. Gaertler, et al "On Modularity Clustering", IEEE Transactions on Knowledge and Data Engineering, p.p.: 172-188, 2008

[BFA11]M. Belleschi, G. Fodor, A. Abrardo, “Performance analysis of a distributed resource allocation scheme for D2D communications”, in Proc. IEEE Workshop on Machine-to-Machine Communications, 2011

[BHF+09] S. Buljore, H. Harada, S. Filin, P. Houze, K. Tsagkaris, O. Holland, K. Nolte, T. Farnham and V. Ivanovz, "Architecture and Enablers for Optimized Radio Resource Usage in Heterogeneous Wireless Access Networks: The IEEE 1900.4 Working Group," IEEE Communication Magazine, vol. 47, no. 1, pp. 122-129, 2009.

[BJM+14] N. Bhushan, D. Malladi, R. Gilmore, et al. “Network densification: the dominant theme for wireless evolution into 5G”, IEEE Communications Magazine, Volume: 52, Issue: 2, February 2014

[BKK+14] M. Botsov, M. Klügel, W. Kellerer, and P. Fertl, “Location dependent resource allocation for mobile Device-to-Device communications,” in 2014 IEEE Wireless Communications and Networking Conference (WCNC), April 2014, pp. 1679–1684.

[BRZ+14] Ö.Bulakci, Z. Ren, C. Zhou, J. Eichinger, P. Fertl and S. Stanczak, “Dynamic Nomadic Node Selection for Performance Enhancement in Composite Fading/Shadowing Environments,” IEEE VTC 2014-Spring, Seoul, South Korea.

[BY12] P. Bao, G. Yu, "An interference management strategy for device-to-device underlaying cellular networks with partial location information," in Proc. Personal Indoor and Mobile Radio Communications (PIMRC) 2012, pp.465-470, 9-12 Sept. 2012

[CGY14] D. Calabuig, R. H. Gohary and H. Yanikomeroglu, "Optimization of a class of non-convex objectives on the Gaussian MIMO multiple access channel: Algorithm development and convergence analysis," in IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Toronto, Canada, 22-25 June 2014.

[CHP+09] M. Charafeddine, Z. Han, A. Paulraj, and J. Cioffi, “Crystallized rates region of the interference channel via correlated equilibrium with interference as noise,” in Proceedings of the IEEE International Conference on Communications (ICC ’09), Dresden, Germany, June 2009.

[CMK+08] I. Cananea, D. Mariz, J. Kelner, D. Sadok, G. Fodor, "An On-Line Access Selection Algorithm for ABC Networks Supporting Elastic Services," in IEEE Wireless Communications and Networking Conference (WCNC), Las Vegas, NV, USA, 2008.

[CT06] T. Cover and J. Thomas, “Elements of information theory,” 2nd edition, John Wiley & sons, 2006.

[Dey00] A.K.Dey, ‘Providing Architectural Support for Building Context-Aware Applications’, PhD Thesis, College of Computing, Georgia Institute of Technology, December 2000

[Don00], S. van Dongen, “Graph Clustering by Flow Simulation”. PhD thesis, University of Utrecht, May 2000.

[DRW+09] K. Doppler, M. Rinne, C. Wijting,C.B. Riberio, K. Hugl, “D2D communications underlaying an LTE cellular network”, IEEE Communications Magazine 7(12), 42–49 (2009),

[EHB08] J. Ellenbeck, C. Hartmann, L. Berlemann, Decentralized Inter-cell interference Coordination by Autonomous Spectral Reuse Decisions, Proc. of European Wireless 2008, 22-25 June 2008, Prague.

[FBP+15] G. Fodor, M. Belleschi, D, D. Penda, A. Pradini, M. Johansson, A. Abrardo, “Benchmarking Practical RRM Algorithms for D2D Communications in LTE Advanced”, Wireless Personal Communications (Springer), January 2015. DOI 10.1007/s11277-014-2258-1; arXiv:1306.5305 [cs.IT], [http://arxiv.org/abs/1306.5305].

[FDP+12] G. Fodor, E. Dahlman, S. Parkvall, G. Mildh, N. Reider, G. Miklos, Z. Turanyi, ”Design Aspects of Network Assisted Device-to-Device Communications”, IEEE Communications Magazine, March 2012.

[FET03] G. Fodor, A. Eriksson and A. Tuoriniemi, "Providing QoS in Always Best Connected Networks," IEEE Communications Magazine, 2003

[Fet14] G. P. Fettweiss, “The Tactile Internet: Applications and Challenges”, IEEE Vehicular Technology Magazine, Volume: 9, Issue: 1, 2014

[FK12] S. Fernandes and A. Karmouch, "Vertical Mobility Management Architectures in Wireless Networks: A Comprehensive Survey and Future Directions," IEEE Communication Surveys Tutorials, vol. 14, no. 1, pp. 45-63, 2012.

[FPB+13] G. Fodor, D. D. Penda, M. Belleschi, M. Johansson and A. Abrardo, "A Comparative Study of Power Control Approaches for Device-to-Device Communications", IEEE International Conference on Communications (ICC '13), Budapest, Hungary, June 2013.

[FPG14-1] G. Fodor, A. Pradini, A. Gattami, "On Applying Network Coding in Network Assisted Device-to-Device Communications", European Wireless '14, Barcelona, Spain, May 14-16, 2014.

[FPG14-2] G. Fodor, A. Pradini and A. Gattami, "Device-to-Device Communications and Network Coding: Friends or Foes?", IEEE Comsoc MMTC E-Letter, Volume 9, No. 1, pp. 33-35, January 2014.

[FPS+15] G. Fodor, S. Parkvall, S. Sorrentino, P. Wallentin, Q. Lu, N. Brahmi, "Device-to-Device Communications for National Security and Public Safety", IEEE Access, Vol. 2, pp. 1510-1520, DOI: 10.1109/ACCESS.2014.2379938, January 2015.

[FSS14] G. Fodor, S. Sorrentino and S. Sultana, "Network Assisted Device-to-Device Communications: Use Cases, Design Approaches, and Performance Aspects", Book Chapter in: S. Mumtaz, J. Rodriguez (Eds), "Smart Device to Smart Device Communications", Springer, ISBN: 978-3-319-04962-5 (Print) 978-3-319-04963-2 (Online), March 2014.

[FU98] G. D. Forney and G. Ungerboeck, "Modulation and coding for linear Gaussian channels", IEEE Transactions on Information Theory, vol. 44, no. 6, pp. 2384-2415, October 1998.

[GCC+10] L.G.U Garica, G.W.O Costa, A. F Cattoni, K.I Pedersen, P. E Mogensen, Self-Organising Coalitions for Conflict Evaluation and Resolution in Femtocells, GLOBECOM 2010, pp 1-6, December 2010.

[GCG+14] Xiaohu Ge, Hui Cheng , M. Guizani, Tao Han, "5G wireless backhaul networks: challenges and research advances", IEEE Network, Volume: 28, Issue: 6, December 2015.

[GJ03] E. Gustafson and A. Jonsson, "Always Best Connected," IEEE Wireless Communications Magazine, vol. 10, no. 1, pp. 49-55, 2003.

[HSZ+14] B. Hassanabadi, C. Shea, L. Zhang, and S. Valaee, "Clustering in vehicular ad hoc networks using affinity propagation," Ad Hoc Networks, vol. 13, pp. 535–548, 2014.

[IKT12] H. Ishii, Y. Kishiyama, H. Takahashi, „A novel architecture for LTE-B: C-plane/U-plane split and Phantom Cell concept", Globecom 2012.

[ITUR14-BT2020] Recommendation ITU-T BT.2020-1, "Parameter values for ultra-high definition television systems for production and international programme exchange", June 2014.

[ITUR14-M2320] Report M.2320, "Future technology trends of terrestrial IMT systems", November 2014.

[JKR+09] P. Janis, V. Koivunen, C. Ribeiro, J. Korhonen, K. Doppler, and K. Hugl, "Interference-aware resource allocation for device-to-device radio underlaying cellular networks," in IEEE 69th Vehicular Technology Conference Spring 2009.

[JN13] M. R. Jeong and N. Miki, "A simple scheduling restriction scheme for interference coordinated networks," IEICE Transactions on Communications, vol. E96-B, no. 6, pp. 1306–1317, 2013.

[Kle15] A. Klein, "Context Awareness for Enhancing Heterogeneous Access Management and Self-Optimizing Networks", PhD thesis at University of Kaiserslautern (publication in progress), 2015.

[KLM+11a] Klein, A.; Lottermann, C.; Mannweiler, C.; Schneider, J.; Schotten, H.D., "A Novel Approach for Combined Joint Call Admission Control and Dynamic Bandwidth Adaptation in

Heterogeneous Wireless Networks", 7th EURO-NF Conference on Next Generation Internet (NGI), Kaiserslautern, Germany, pp. 1–8, 2011.

[Laa05] K. Laasonen, "Clustering and Prediction of Mobile User Routes from Cellular Data", 9th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD), Porto, Portugal, 2005.

[LH05] S.Liou and Y.Huang, "Trajectory Predictions in Mobile Networks", International Journal of Information Technology Vol.11, No.11, 2005.

[LHW+10] Wenxin Liu, Chunjing Hu, Dongyan Wei, Peng, M. and Wenbo Wang, "An overload indicator & high interference indicator hybrid scheme for inter-cell interference coordination in LTE system," Broadband Network and Multimedia Technology (IC-BNMT), pp. 514-518, October, 2010.

[LLK+11] L. Lindbom, R. Love, S. Krishnamurthy, C. Chunhai; N. Miki, V. Chandrasekhar, "Enhanced Inter-cell Interference Coordination for Heterogeneous Networks in LTE-Advanced: A Survey", [online] Available:http://arxiv.org/abs/1112.1344, 2011.

[LMS14] T.R. Lakshmana, B. Makki, and T. Svensson "Frequency Allocation in Non-Coherent Joint Transmission CoMP Networks," in Proc. IEEE International Conference on Communications (ICC 2014), Jun. 2014.

[Lor15] J. Lorca, "Increasing coverage and maximum CFO in DFT-s-OFDM for Machine-Type Communications", to appear in Proceedings of the IEEE International Conference on Communications (ICC 2015), London, June 2015.

[LZW+11] S. Lv, W. Zhuang, X. Wang, and X. Zhou, "Context-aware scheduling in wireless networks with successive interference cancellation," in Proc. IEEE International Conference on Communications (ICC 2011), Jul. 2011.

[Maz75] J. E. Mazo, "Faster-than-Nyquist signaling", Bell System Technical Journal, vol. 54, no. 8, pp. 1451-1462, October 1975.

[MCC+09] M. A. Marsan, L. Chiaraviglio, D. Ciullo, M. Meo, "Optimal Energy Savings in Cellular Access Networks", IEEE ICC, GreenComWksp., Dresden, Germany, June 2009

[MDS+12] Prodromos Makris, Dimitrios N. Skoutas, Charalabos Skianis, "A Survey on Context-Aware Mobile and Wireless Networking: On Networking and Computing Environments' Integration," IEEE Communications Surveys & Tutorials, vol. PP, no. 99, pp. 1 - 25, 2012.

[MET13-D11] ICT-317669 METIS, Deliverable 1.1 Version 1, "Scenarios, requirements and KPIs for 5G mobile and wireless system", April 2013.

[MET13-D41] ICT-317669 METIS, Deliverable 4.1 Version 1,"Summary on preliminary trade-off investigations and first set of potential network-level solutions", October 2013.

[MET13-D61] ICT-317669 METIS, Deliverable 6.1 Version 1, "Simulation Guidelines", Oct. 2013.

[MET14-D42] ICT-317669 METIS, Deliverable 4.2 Version 1,"Final report on trade-off investigations", September 2014.

[MET14-D53] ICT-317669 METIS, Deliverable 5.3 Version 1, "Description of the spectrum needs and usage principles", August 2015.

[MET15-D24] ICT-317669 METIS, Deliverable 2.4 Version 1, "Proposed solutions for new radio access", March 2015.

[MET15-D33] ICT-317669 METIS, Deliverable 3.3 Version 1, "Final performance results and consolidated view on the most promising multi-node/multi-antenna transmission technologies", March 2015.

[MET15-D54] ICT-317669 METIS, Deliverable 5.4 Version 1, “Future spectrum system concept”, May 2015.

[MET15-D64] ICT-317669 METIS, Deliverable 6.4 Version 1, “Final report on Architecture,” January 2015.

[MET15-D66] ICT-317669 METIS, Deliverable D6.6 Version 1, “Final report on the METIS 5G system concept and technology roadmap,” April 2015.

[MK05] D.J.C. MacKay, "Fountain codes," Communications, IEE Proceedings- , vol.152, no.6, pp.1062,1068, 9 Dec. 2005.

[MLP+11] H. Min, J. Lee, S. Park, D. Hong, “Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks”, IEEE Transactions on Wireless Communications 10(12), 3995–4000 (2011).

[MPM12] S.Michaelis, N. Piatkowski and K. Morik, “Predicting next network cell IDs for moving users with Discriminative and Generative Models”, Mobile Data Challenge Workshop in conjunction with International Conference on Pervasive Computing, Newcastle, UK, 2012.

[Ols12] M. Olsen, "Final Integrated Concept", ICT-247733 EARTH, July 2012

[PBS+12] F. Pantisano, M. Bennis, W. Saad, S. Valentin, and M. Debbah, "Matching with externalities for context-aware cell association in wireless small cell networks," in IEEE International Conference on Communications (ICC), 2012.

[PFM+14] A. Pradini, G. Fodor, G. Miao, M. Belleschi, ”Near Optimal Practical Power Control Schemes for D2D Communications in Cellular Networks", European Conference on Networks and Communications (EuCNC ‘14), Bologna, Italy, June 23-26, 2014.

[PKW+12] M. Proebster, M. Kaschub, T. Werthmann, S. Valentin, "Context-Aware Resource Allocation to Improve the Quality of Service of Heterogeneous Traffic," in IEEE International Conference on Communications (ICC), 2011.

[PO+13] P. Öhlén et al., ”Software-Defined Networking in a Multi-Purpose DWDM-Centric Metro/Aggregation Network”, IEEE Globecom WS 2013.

[PWH13] K. Pentikousis, Y. Wang and W. Hu, “Mobileflow: Toward software-defined mobile networks”, IEEE Communications Magazine, vol. 51, no. 7, pp. 44-53, July 2013.

[QQF+14] L. Qianxi, M. Qingyu, G. Fodor, N. Brahmi, "Clustering Schemes for D2D Communications Under Partial/No Network Coverage", VTC Spring-2014, May 18-21, Seoul, Korea.

[RF11] N. Reider and G. Fodor, "A Distributed Power Control Scheme for Cellular Network Assisted D2D Communications", IEEE Globecom, Houston, TX, USA, December 5-9, 2011.

[RF12] N. Reider and G. Fodor, "A Distributed Power Control and Mode Selection Algorithm for D2D Communications", EURASIP Journal on Wireless Communications and Networking, 2012:266, 21 August 2012.

[RFL+13] Z. Ren, P. Fertl, Q. Liao, F. Penna, and S. S. and, “Street Specific Handover Optimization in Future Cellular Networks,” in Processing of IEEE Vehicular Technology Conference Spring (VTC-Spring), Dresden, Germany, June, 2013, pp. 1–5.

[RM12] Z. Y. Rawashdeh, and S. M. Mahmud, “A novel algorithm to form stable clusters in vehicular ad hoc networks on highways”, EURASIP Journal on Wireless Communications and Networking, 2012.

[Roy08] S. Roy, ”Energy logic for telecommunications”. White Paper, Global Marketing, 2008

[RSF14] Z. Ren, S. Stanczak, P. Fertl, and F. Penna, “Energy-Aware Activation of Nomadic Relays for Performance Enhancement in Cellular Networks,” Proceedings of IEEE International Conference on Communications (ICC), Sydney, Australia, June, 2014, pp. 1–6.

[RSF+14] Z. Ren, S. Stanczak, and P. Fertl,, “Activation of Nomadic Relays in Dynamic Interference Environment for Energy Savings,” in Proceedings of IEEE Global Conference on Communications (GLOBECOM), Austin, Texas, December, 2014, pp. 1–6.

[RSS14] Z. Ren, S. Stanczak, Mahdy Shabeeb, P. Fertl, and Lars Thiele, “A Distributed Algorithm for Energy Saving in Nomadic Relaying Networks,” in Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, November., 2014, pp. 1–5.

[RTR07] A. Racz, A. Temesvary and R. Reider, “Handover Performance in 3GPP Long Term Evolution (LTE) Systems”, Proc. Of Mobile and Wireless Communications Summit, 2007, 16th IST, ppp. 1 – 5, July 2007.

[SB12] M. Schubert and H. Boche, “Interference Calculus - A General Framework for Interference Management and Network Utility Optimization”, Springer, Feb. 2012.

[SFM14] J. M. B. da Silva, G. Fodor, T. F. Maciel, ”Performance Analysis of Network Assisted Two-Hop D2D Communications” IEEE Broadband Wireless Access Workshop, Austin, TX, December 12, 2014.

[SGS15] Y. Sui, I. Guvenc and T. Svensson “Interference Management for Moving Networks in Ultra-Dense Urban Scenarios”, EURASIP Journal on Wireless Communications and Networking. Special Issue on 5G Wireless Mobile Technologies, Mar. 2015.

[SK06] N. Selvakkumaran and G. Karypis, “Multi-Objective Hypergraph Partitioning Algorithms for Cut and Maximum Subdomain Degree Minimization”, IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems, vol. 25, no. 3, pp. 504–517, March 2006.

[SNG10] E. Souza, I. Nikolaidis, P. Gburzynski, "A New Aggregate Local Mobility (ALM) Clustering Algorithm for VANETs,“ IEEE International Conference on Communications (ICC), pp.1-5, 2010.

[SPS12] Y. Sui, A. Papadogiannis, T. Svensson, “The Potential of Moving Relays - A Performance Analysis”, VTC Spring 2012.

[SRS+13] Y. Sui, Z. Ren, W. Sun, T. Svensson, and P. Fertl, “Performance study of fixed and moving relays for vehicular users with multi-cell handover under co-channel interference,” in Proc. IEEE Int. Conf. on Connected Veh. and Expo (ICCVE), Las Vegas, Dec. 2013.

[SS14] Sharma S., “Integrated Backhaul Management for Ultra-Dense Network Deployment” Master Thesis at KTH, TRITA-ICT-EX-2014:189.

[SSB+14] W. Sun, E. G. Ström, F. Brännström, Y. Sui, and K. C. Sou, “D2D-based V2V communications with latency and reliability constraints”, in IEEE GLOBECOM Workshops, Dec. 2014.

[STB11] S. Sesia, I. Toufik, and M. Baker, LTE - The UMTS Long Term Evolution: From Theory to Practice, 2nd ed. West Sussex, UK: John Wiley & Sons Ltd., 2011.

[SYS+15] W. Sun, D. Yuan, E. G. Ström, and F. Brännström, “Resource sharing and power allocation for D2D-based safety critical V2X communications”.

[TAHD14] E. Ternon, P. Agyapong, Liang Hu; A. Dekorsy, "Database-aided energy savings in next generation dual connectivity heterogeneous networks," Wireless Communications and Networking Conference (WCNC), 2014 IEEE , vol., no., pp.2811,2816, 6-9 April 2014.

[TAZ+13] A. Tzanakaki, M. P. Anastasopoulos, G. S. Zervas, B. R. Rofoee, R. Nejabati and D. Simeonidou, “Virtualization of heterogeneous wireless-optical network and IT infrastructures in support of cloud and mobile cloud services”. IEEE Communications Magazine, vol. 51, no. 8, pp. 155-161, August 2013.

[TBM+13] F. M. L. Tavares, G. Berardinelli, N. H. Mahmood, et al., “On the Potential of Interference Rejection Combining in B4G Networks”, VTC Fall, 2013.

[WC12] H. Wang, X. Chu, “Distance-contrained resource-sharing criteria for device-to-device communications underlaying cellular networks”, Electronics Letters, Vol.48, No.9, April 2012.

[WSP+09] H. Wei-jen, T. Spyropoulos, K. Psounis, A. Helmy, "Modeling Spatial and Temporal Dependencies of User Mobility in Wireless Mobile Networks", IEEE/ACM Transactions on Networking, vol.17, no.5, pp.1564,1577, Oct. 2009

[YLV+14] O.N.C. Yilmaz, Z. Li, K. Valkealahti, M.A. Uusitalo, M. Moisio, P. Lundén, C. Wijting, "Smart Mobility Management for D2D Communications in 5G Networks", IEEE WCNC 2014 WS on Device-to-Device and Public Safety Communications (WDPC), Istanbul, Turkey, April 2014.

[YSC+13] Q. Ye, M. Shalash, C. Caramanis, and J. G. Andrews, “On/off macrocells and load balancing in heterogeneous cellular networks” in Proc. IEEE GLOBECOM, Atlanta, GA, USA, pp. 3814–3819, 2013.

[YWL+14] O.N.C. Yilmaz, C. Wijting, P. Lundén, J. Hämäläinen, "Optimized Mobile Connectivity for Bandwidth-Hungry, Delay-Tolerant Cloud Services toward 5G", 11th International Symposium on Wireless Communication Systems (ISWCS’2014), Barcelona, Spain, August 2014.

[YYK+14] Yanyan Ma ; Yu Jiang ; Kakishima, Y. ; Nagata, S. ; Kishiyama, Y. ; Nakamura, T. “System-level throughput evaluation of multiuser MIMO using enhanced codebook considering user mobility in LTE-Advanced downlink,” 11th International Symposium on Wireless Communications Systems (ISWCS), 2014 , pp. 432-437.

[ZHS10] M. Zulhasnine, C. Huang, and A. Srinivasan, “Efficient resource allocation for device-to-device communication underlaying LTE network”, in IEEE WiMob, Oct. 2010.

Annex A

A.1 Interference Management and Resource Allocation Schemes

A.1.1 Coordinated Fast Uplink and Downlink Resource Allocation in UDN

Studies shown in Section 2.6 capture two corner cases of RRM studies with flexible UL/DL TDD based air interfaces. These corner cases (fully centralized and distributed/standalone mechanisms) can be complemented with distributed RRM mechanism that realizes network-level KPI optimization based on exchange of information between small cells BS. LTE-A eICIC is an example of distributed RRM. For 5G networks distributed RRM needs to be more proactive and frequency domain muting, fast uplink downlink adaptations can be better exploited for UDN. Mechanism that was evaluated (Figure A.1-1) is based on following rules:

- Each BS is randomly allocated a set of primary and secondary resources.
- BS never mutes on primary resources. Sets a primary SNR target for all other BSs on its primary resources.
- BS mutes on a secondary resource if it does not meet the SNR target of its neighbor and if its metric is less than its neighbor
- Metric exchange is thus made between the BS to proactively decide frequency domain muting
- Muting decision is made after UL/DL scheduling decision

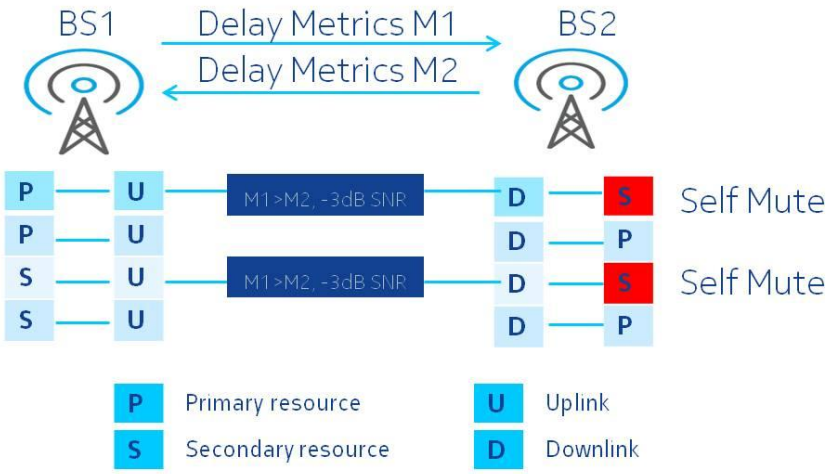


Figure A.1-1. Distributed RRM scheme for centralization/decentralization studies

All centralized, distributed and decentralized schemes were evaluated using assumptions as in Table A.1-1.

Performed simulations for a distributed scheme are captured in Figure A.1-2. Achieved 99th percentile packet delay results position distributed solution between centralized and fully decentralized/standalone schemes (centralized RRM show best performance in terms of packet delay) for the packets transmitted in UL. Experience of the worst case user can be hence improved without going for the fully centralized scheme.

Table A.1-1. Parameters used for RRM centralization level studies

Parameter	Value
Packet size	UL packet size = 640 kByte/sec/user on 20 MHz DL packet size = 4 x 640 kByte/sec/user on 20 MHz
Packet inter-arrival time	1 sec for all users
TX powers	DL: -3 dBm UL average: -5.4 dBm, UL max: -1.58 dBm
Wall loss	3 dB
Cell deployment	25 cells, 5 m x 5 m each
Distributed scheme parameters	Primary resource ratio: 0.4 Primary SNR target = 3 dB Secondary SNR target = 0 dB Only path-loss information is used in SNR computation.
Scheduling slot	2 ms comprising of few 0.25 ms TTIs
Antenna configuration	4X4 MIMO in all simulations. MCS and rank determined for MIMO.

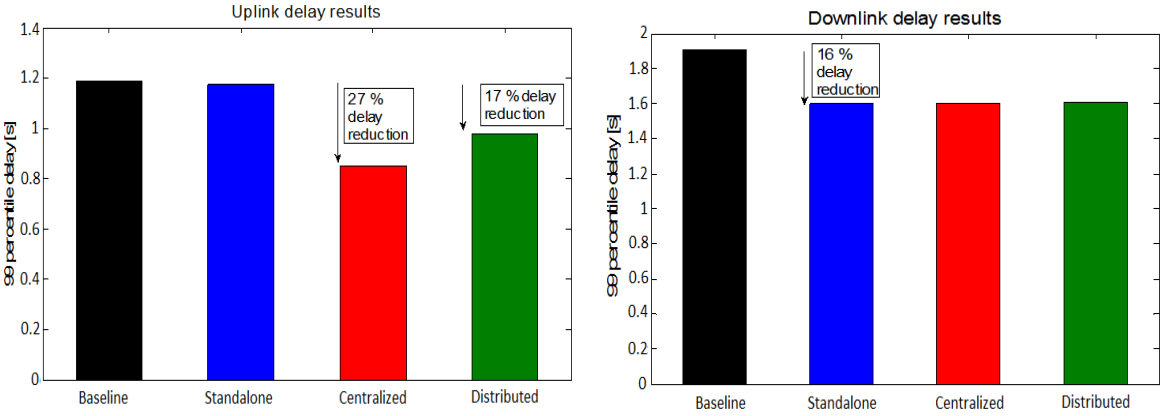


Figure A.1-2. UL and DL packet delay of the 99th percentile CDF, different RRM schemes

Presented material captures changing performance of different RRM schemes depending on the level of centralization. For 5G systems location of RRM functionality may change in RAN depending, e.g., on the availability of centralized entity (such as macro cell), signalling overhead constraints, backhaul quality, etc. Further considerations (e.g., inclusion of D2D mode of operations, providing reliable signalling for centralized operations) can be found in Sections 2.8 and 5.5

A.1.2 CSI-based Coordination Scheme for Macro or Small Cells with Non-coherent JT CoMP

Given a bandwidth W, we investigate the frequency allocation using a slider s as shown in Figure A.1-3. The frequency resources can either be dedicated or shared between UEs. The various degrees in which the resources are allocated are controlled using the slider s. With s=0, the complete bandwidth is shared between users, and with s=3 the frequency resources are disjointly divided amongst the users. In Figure A.1-3 s=2 each UE gets $W/3$ dedicated resource (DR) and $W/3$ of the shared resource (SR). With s=1, more frequency resources are shared than being dedicated. Note that the slider position maintains fairness among users. Group of subcarriers experiencing a flat fading have a coherence bandwidth,w_c.

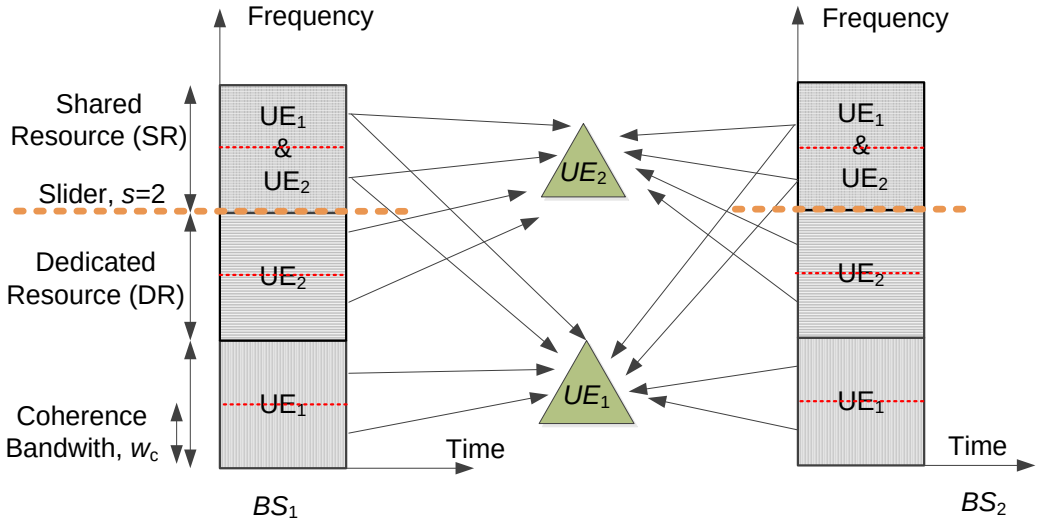


Figure A.1-3. Frequency allocation between two UEs, where each UE has the dedicated resource and shared resource equally distributed

We study the trade-off between DR and SR for a given UE, and provide a generic model in evaluating the rates. The rate achieved [CT06, eq. 9.62] due to the dedicated resource for the mth user and equal power allocation is

$$R_m^{\mathrm{DR}} = \sum_{\substack{n=(m-1)sN+1 \\ s\neq 0}}^{msN} \frac{w_c}{N}\log_2(1+\gamma_{m;n}^{\mathrm{DR}})$$

where N is the number of subcarriers, M is the number of users, and the SNR on nth subcarrier with power $P_{m;n}^{\mathrm{DR}}$ is $\gamma_{m;n}^{\mathrm{DR}} = \frac{\frac{N}{w_c}\left|\sum_{k=1}^{K} h_{m,k;n}\sqrt{P_{m;n}^{\mathrm{DR}}}\right|^2}{\sigma^2}$ where $h_{m,k;n}$ is the channel from the kth BS to the mth user in the nth subcarrier, and σ^2 is the receiver noise. For the SR, the rate achieved is

$$R_m^{\mathrm{SR}} = \sum_{\substack{n'=MsN+1 \\ s\neq M+1}}^{M(M+1)N} \frac{w_c}{N}\log_2(1+\gamma_{m;n}^{\mathrm{SR}})$$

where the SINR on the n'th subcarrier with power $P_{m;n}^{\mathrm{SR}}$ is $\gamma_{m;n'}^{\mathrm{SR}} = \frac{\frac{N}{w_c}\left|\sum_{k=1}^{K} h_{m,k;n'}\sqrt{P_{m;n'}^{\mathrm{SR}}}\right|^2}{\frac{N}{w_c}\left|\sum_{l\neq m}\sum_{\substack{k=1 \\ l\in\mathcal{L}_i}}^{K} h_{m,k;n'}\sqrt{P_{m;n'}^{\mathrm{SR}}}\right|^2 + \sigma^2}$,

where $\mathcal{L}_i$ is a subset of users sharing the same frequency resource. Note that the broadcast

approach is followed here, hence the interference is on the same channel, and that the interference is treated as noise. The long term throughput is governed by

$$\eta_{T+1} = \sum_{m=1}^{M} \sum_{t=1}^{T+1} \frac{r_m}{t} \Pr\left(\frac{r_m}{t} < R_m \leq \frac{r_m}{t-1}\right)$$

where $R_m = R_m^{\mathrm{DR}} + R_m^{\mathrm{SR}}$, r_m is the initial transmission rate, and T is the total number of retransmissions. The outage probability after $(T+1)$th retransmission for the mth user is

$$P_{m,T+1}^{\mathrm{outage}} = \Pr\left(R_m \leq \frac{r_m}{T+1}\right).$$

This is captured in Figure A.1-4, where the sharing the frequency resource performs the best, for an initial transmission rate of $r_m = 4$ bps. The table below summarizes the possible frequency allocations for the mth user when number of BSs are the same as the number of users, $K = M = 2$, and $W = 1$ Hz. More details of this work can be found in [LMS14].

Table A.1-2. Possible frequency allocations for the mth user

Slider position	Dedicated Resource (DR)	Shared Resource (SR)	Interpretation	Legend
$s = 0$	0	1	All shared	Blue asterisk
$s = 1$	1/6	4/6	Some dedicated	Magenta crosses
$s = 2$	2/6	2/6	Uniformly allocated	Red dots
$s = 3$	3/6	0	Dedicated	Black triangles

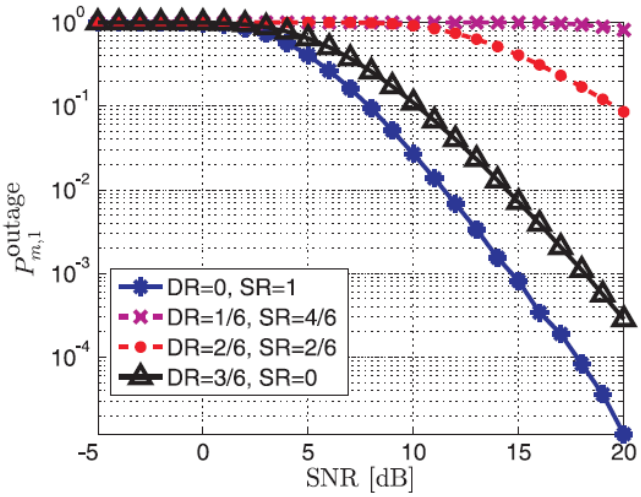


Figure A.1-4. Outage probability versus SNR for a given $r_m = 4$ bps in an open loop system (without feedback, $T = 0$)

A.1.3 Time-Sharing Approach to Interference Mitigation Using Resource Auctioning and Regret-Matching Learning

The proposed approach based on regret-matching procedure has been compared in Monte-Carlo simulations with the results achieved for baseline LTE-A system without interference mitigation and also with the scheme based on LTE-A eICIC using ABSs. The investigated performance indicators were average cell spectral efficiency and UE throughput distribution. The following table summarizes the simulation assumptions.

Table A.1-3. Simulation parameters

Parameter	Value
Number of macro BSs	3 (1 site)
Number of micro BSs	18 (9 sites)
Maximum macro BS transmit power	43 dBm
Maximum micro BS transmit power	30 dBm
Carrier frequency	2.6 GHz
System bandwidth	20 MHz
Range expansion threshold	{5 dB, 10 dB}
MIMO	4×2 transmit diversity
UE distribution	uniform (only outdoor) - 300 UEs
UE mobility	none

The LTE-A eICIC scheme considered in simulations was a static one with two different ABS configurations: with 2 and 4 ABSs in a radio frame. Similarly, for the proposed approach based on regret-matching two power reduction levels P_d have been considered: -10 dB and -30 dB.

Table A.1-4. Average cell spectral efficiency

	SE (bps/Hz)	Gain (%)	SE (bps/Hz)	Gain (%)
Range expansion threshold	5 dB		10 dB	
no ICIC	2,105	-	2,135	-
eICIC (4 ABSF)	1,926	-9,28	2,082	-2,53
eICIC (2 ABSF)	1,821	-15,58	1,965	-8,63
Regret-matching (P_d = -10 dB)	2,272	7,37	2,267	5,83
Regret-matching (P_d = -30 dB)	2,157	2,44	2,167	1,48

Table A.1-4 presents the obtained values of average cell spectral efficiency and the corresponding gains using proposed schemes compared against the LTE-A with no ICIC. One can notice that the approach based on regret-matching provides gains even over 7%.

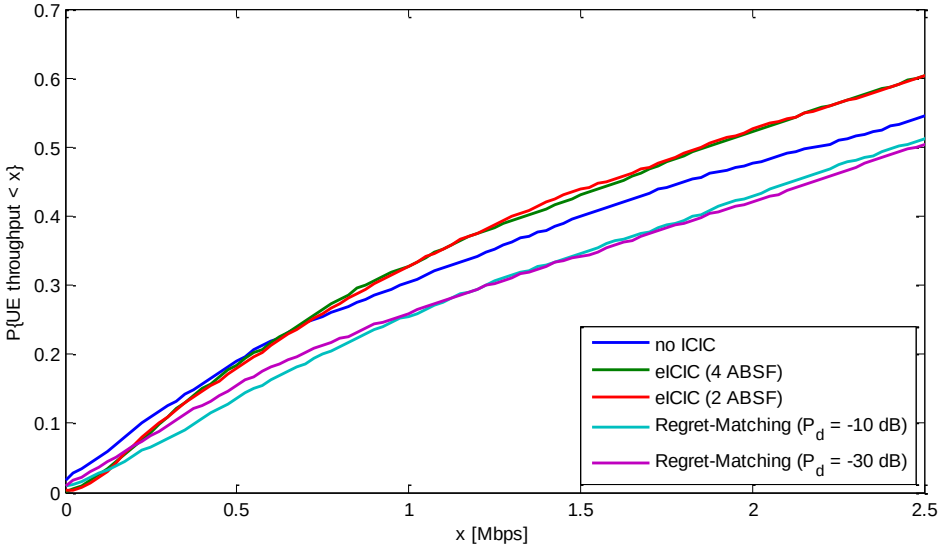


Figure A.1-5. UE throughput CDF for 5dB range expansion threshold

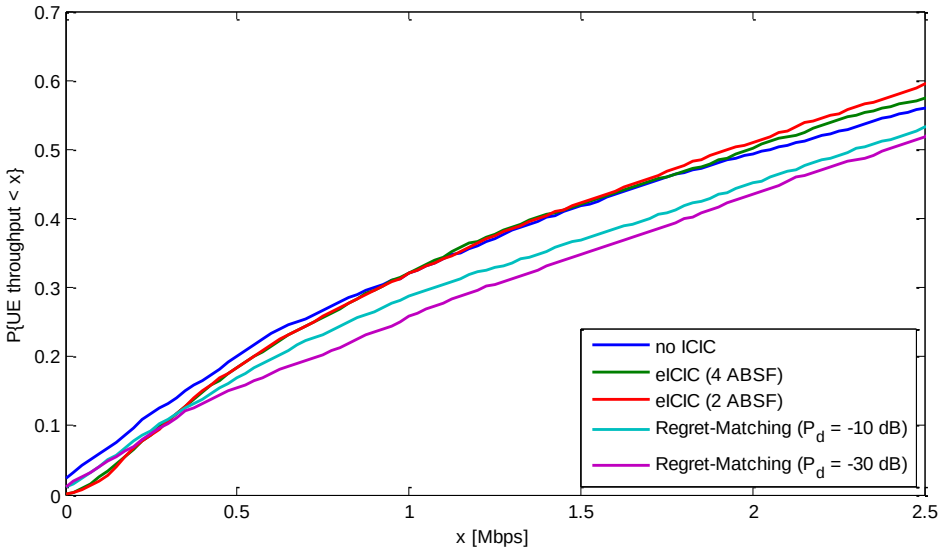


Figure A.1-6. UE throughput CDF for 10dB range expansion threshold

Figure A.1-5. and Figure A.1-6. present the UE throughput distribution for the range expansion threshold equal to 5 dB and 10 dB respectively. These are also summarized in more details in Table A.1-5 and Table A.1-6, where the achieved 5th and 50th percentile UE throughput is outlined, with the corresponding gains vs. baseline LTE-A without ICIC. One can notice that the proposed interference mitigation scheme based on regret-matching procedure provides higher UE throughputs with up to 40% gain in 5th percentile throughput comparing to LTE-A without ICIC. On the other hand the eICIC scheme provides higher gains in terms of 5th percentile UE throughput (up to 50%), however these come at a cost of reduced cell spectral efficiency. This is also shown in Table A.1-6, where the use of eICIC results in lower 50th percentile UE throughput than in case of no ICIC. On the other hand, with the use of regret-matching gains over 10% in 50th percentile UE throughput can be achieved.

Table A.1-5. 5th percentile UE throughput

	SE (bps/Hz)	Gain (%)	SE (bps/Hz)	Gain (%)
Range expansion threshold	5 dB		10 dB	
no ICIC	0,122	-	0,095	-
eICIC (4 ABSF)	0,187	35,15	0,189	49,36
eICIC (2 ABSF)	0,188	35,35	0,193	50,42
Regret-matching (P_d = -10 dB)	0,215	43,42	0,148	35,65
Regret-matching (P_d = -30 dB)	0,169	27,96	0,157	39,17

Table A.1-6. 50th percentile UE throughput

	SE (bps/Hz)	Gain （%）	SE (bps/Hz)	Gain （%）
Range expansion threshold	5 dB		10 dB	
no ICIC	2,200	-	2,071	-
eICIC (4 ABSF)	1,891	-16,35	2,011	-3,01
eICIC (2 ABSF)	1,875	-17,33	1,948	-6,35
Regret-matching (P_d = -10 dB)	2,442	9,92	2,322	10,79
Regret-matching (P_d = -30 dB)	2,503	12,11	2,411	14,09

A.1.4 Self-Organization of Neighbouring Femtocell Clusters: Mathematical Description of the Beacon Signal

The proposed beacon signal is intended to be broadcast by already synchronized BSs (generally, but not limited to, macro cells) and is designed with enhanced detection properties. The beacon should have no impact on legacy terminals and BSs; therefore it is included in a number of resource elements (REs) not previously used for other purposes. Out of the 72 REs comprising the six central RBs of each subframe, only 62 are occupied by the Primary and Secondary Synchronization Signals (PSS/SSS) at specific subframes, thus leaving 5 unused REs above and below. Four of these REs can be exploited for the Synchronization Beacon, leaving one free RE above and below the PSS/SSS. This free RE will facilitate the use of a 64-point Fast Fourier Transform (FFT) by the device when trying to decode the synchronization signal with a low-complexity receiver.

The Synchronization Beacon would comprise two bipolar Walsh-Hadamard length-8 sequences. These sequences are very simple to decode and remain orthogonal when perfectly aligned. Leaving only one orthogonal sequence in subframe 0 aids in the detection process as inter-sequence interference is avoided. A total of eight orthogonal Walsh-Hadamard sequences of length-8 can be constructed, from which only two are required.

Each sequence is split into two half-sequences that are mapped on the symbols corresponding to the PSS and SSS, and on the REs above and below those occupied by PSS/SSS (see Figure A.1-7). The mapping must be done in increasing order of the resource element number, reading the elements from left to right and mapping them to REs with increasing subcarrier frequency indices. The same Synchronization Beacon will be transmitted by all cells.

Figure A.1-7 illustrates the resource elements involved for frame structure type 1 and normal cyclic prefix, choosing two possible Walsh-Hadamard length-8 sequences. The two REs located above and below the PSS/SSS must not be used, in order to help the receiver in acquiring synchronization by means of a 64-point FFT. The 4 remaining REs would have essentially no inter-cell interference (if the network is supposed to be fully time-synchronized), suffering only from thermal noise.

A non-coherent detection technique can be employed for detection by unsynchronized small cells given that the channel transfer function between macro and small cell is unknown prior to synchronization. Given a received signal $r[n]$, maximum likelihood non-coherent detection (MLD) involves finding the time offset k for which the following expression is maximized:

$$\hat{k} = \arg\max |\rho_k|,$$

where the correlation coefficient in the frequency domain ρ_k can be defined as the correlation divided by the square root of the product of the energies of both the received and expected signals, after shifting in time the input signal by *k* samples. Correlations can be separately calculated above and below the PSS/SSS. These correlations will also be split into two terms of four REs, covering both OFDM symbols. The correlation coefficient would thus be:

$$|\rho_k| = \frac{1}{\sqrt{8}} \cdot \frac{|Corr^+| + |Corr^-|}{\sqrt{En^+} + \sqrt{En^-}},$$

where

$$Corr^+ = \sum_{i=0}^{3} S_k^{+(1st\_symb)}[i] q_1[i] + \sum_{i=4}^{7} S_k^{+(2nd\_symb)}[i] q_1[i]$$

$$Corr^- = \sum_{i=0}^{3} S_k^{-(1st\_symb)}[i] q_2[i] + \sum_{i=4}^{7} S_k^{-(2nd\_symb)}[i] q_2[i]$$

$$En^+ = \sum_{i=0}^{3} \left| S_k^{+(1st\_symb)}[i] \right|^2 + \sum_{i=4}^{7} \left| S_k^{+(2nd\_symb)}[i] \right|^2$$

$$En^- = \sum_{i=0}^{3} \left| S_k^{-(1st\_symb)}[i] \right|^2 + \sum_{i=4}^{7} \left| S_k^{-(2nd\_symb)}[i] \right|^2$$

$S_k[i]$ represents the frequency contents of $r[n-k]$ at the REs containing the proposed beacon; $q[i]$ is the Synchronization Beacon sequence; and *i* runs over the sixteen resource elements comprising the upper and lower parts of the spectrum. Figure A.1-8 shows a possible detector structure based on two stages.

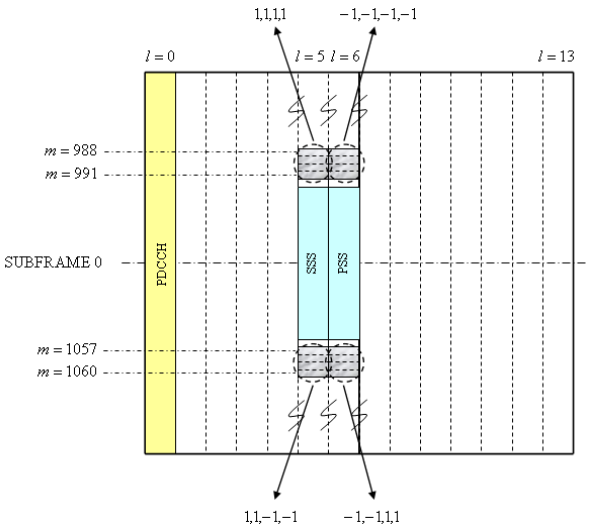


Figure A.1-7. Resource mapping of the beacon signal for FDD and normal cyclic prefix

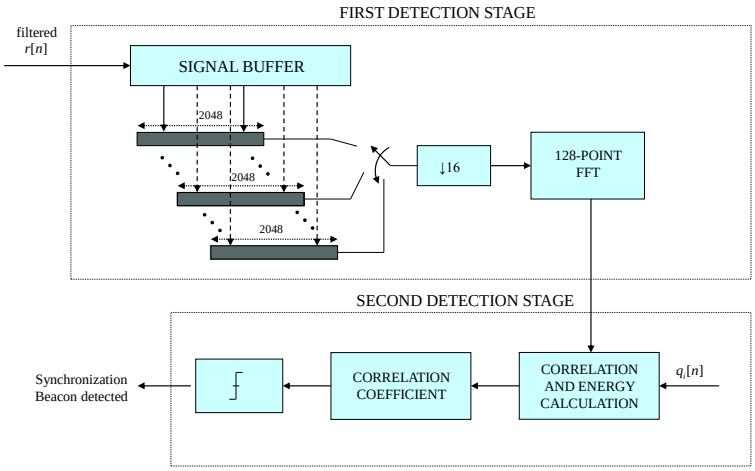


Figure A.1-8. Non-coherent detection scheme for the Synchronization Beacon

A.1.5 Support of Low-Cost MMC Devices Under Low Coverage and High Frequency Offset Conditions

Two approaches are considered in this technology component for the support of low-cost MMC devices. The first one reuses the concept of LTE broadcast/multicast subframes (MBSFN) for dedicated MMC traffic, where cells are grouped into MMC Coordination Areas under control of an MMC Controller node (MMCC). All cells would operate in a synchronized way within an MMC Coordination Area, so that MMC traffic is broadcast in the downlink and jointly received in the uplink, thus increasing coverage. Management of MMC traffic can be made independent from the other types of traffic, therefore allowing tailoring resources to the actual MMC traffic needs while ensuring interoperability with non-MMC devices. Figure A.1-9 shows a possible integration of the MMCC node in an LTE network, and a possible arrangement of MMC Coordination Areas.

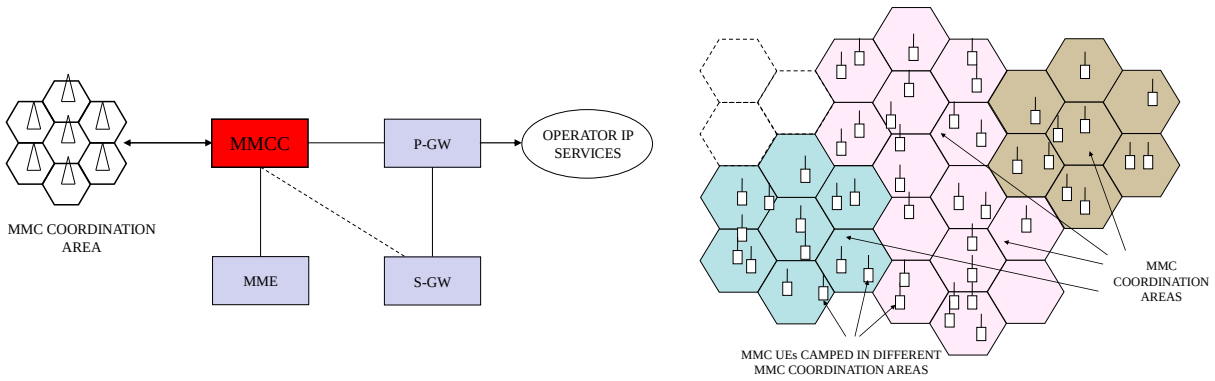


Figure A.1-9. (a) Integration of MMC Controller in an LTE network; (b) example of MMC Coordination Areas

The MMC Controller would be in charge of (de)registering MMC devices, broadcasting and combining UL/DL MMC traffic data, performing RRM and scheduling for MMC users, and providing anchor point for mobility.

A subset of MMC MBSFN subframes is proposed in the downlink to carry dedicated downlink MMC traffic and MMC scheduling. Devices could apply DRX in between MMC MBSFN subframes for battery savings, as shown in Figure A.1-10. Special reserved uplink subframes would also concentrate UL MMC traffic independently from other types of traffic. MMC MBSFN subframes would in turn be sub-divided into data and scheduling subframes, the latter concentrating all MMC scheduling assignments in UL and DL.

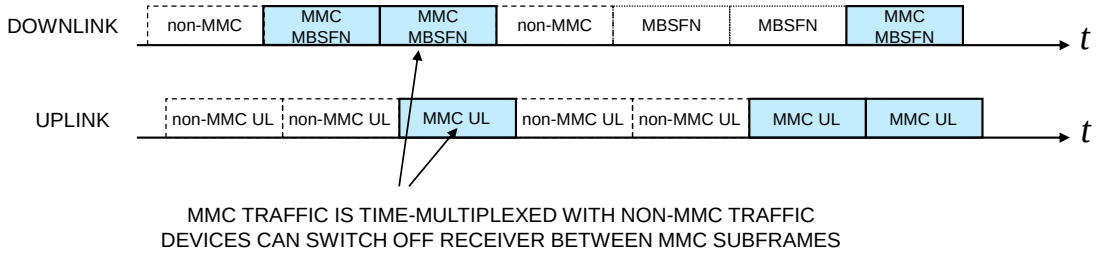


Figure A.1-10. Coexistence of MMC MBSFN subframes in an LTE-like frame structure

Dimensioning of MMC Coordination Areas would be a trade-off between MMC coverage and overall capacity, with the possibility to dynamically adapt it to traffic.

Another approach for the support of low-cost MMC devices proposes improvements in DFT-s-OFDM access technique for the uplink aimed at increasing coverage and maximum supported CFO. Figure A.1-11 shows the modified DFT-s-OFDM transmitter chain where the repetition block just transforms the sequence of complex baseband information symbols into *L* repetitions of *M* symbols prior to the DFT operation. Repetitions in the time domain result in an increase in the power spectral density by a factor $10\times\log(L)$ in dB, equivalent to reducing the bandwidth by *L* but with the added advantage of not losing frequency diversity.

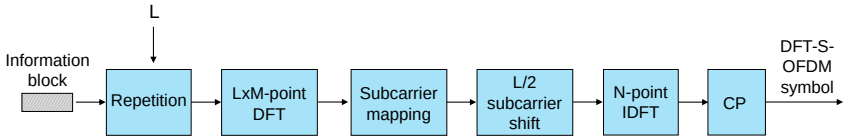


Figure A.1-11. Modified DFT-s-OFDM transmitter chain

The*L*-1 nulls between each pair of consecutive non-zero subcarriers must be left blank in order to allow the support of large CFO values. The time-domain complex waveform comprises *L'* repetitions of *N*/*L'* samples each, where *N* is the symbol length and $L'=\gcd(L,N)\le L$, and this increases the maximum supported CFO by a factor *L'* compared to 3GPP SOTA. Table A.1-7 collects the resulting SNR values for EPA and EVA channel models and several values of *L* vs. legacy *L*=1, showing both significant coverage improvements (from the enhanced frequency diversity gain) and the ability to support large CFO values. The best results were obtained with *L*=6 when CFO is not higher than 0.5Δ*f*; *L*=4 when CFO=Δ*f*; and *L*=8 when CFO=2Δ*f*. It is to note that the BLER curve for EVA and CFO=0.25Δ*f* saturates above 1% (thereby yielding infinite SNR, ∞), because of the very basic CFO estimation method employed in the simulations, which did not average values over multiple OFDM symbols. Better CFO estimations yield lower SNR values in all cases, but the conclusions are similar.

Table A.1-7. Required SNR values for 1% BLER and several values of *L*

CFO	Case	*L*=1	*L*=2	*L*=4	*L*=6	*L*=8
0	EPA, 552 kbps	4	1.5	-0.8	-2	-3.3
	EVA, 91 kbps	-4.7	-7	-9.2	-10	-10.5
0.25Δ*f*	EPA, 552 kbps	20	14	12.8	10.8	13.5
	EVA, 91 kbps	∞	2.5	2.5	1	4.5
0.5Δ*f*	EPA, 552 kbps	∞	14	13	10.7	13.5
	EVA,	∞	2.1	2.5	1	4.5

	91 kbps					
Δf	EPA, 552 kbps	∞	∞	13	∞	13
	EVA, 91 kbps	∞	∞	2.5	∞	4.5
$2\Delta f$	EPA, 552 kbps	∞	∞	∞	∞	13.5
	EVA, 91 kbps	∞	∞	∞	∞	4.5

A.1.6 User Anchored Multi-RAT Self-Managed Load

UEs in idle mode could exploit the cell load indications of multiple neighbour cells as additional inputs to cell selection and reselection algorithms. A suitable cell selection policy could be:

$$S_{rxlev} > 0 \text{ AND } S_{qual} > 0 \text{ AND } L_{serving} < L_{serving,max} \Rightarrow \text{camp in the cell}$$

where

$$S_{rxlev} = Q_{rxlevmeas} - \left(Q_{rxlevmin} + Q_{rxlevminoffset}\right) - P_{compensation}$$
$$S_{qual} = Q_{qualmeas} - \left(Q_{qualmin} + Q_{qualminoffset}\right)$$

Here $L_{serving}$ denotes the cell load, $L_{serving,max}$ is the maximum allowed cell load, and the rest of parameters are as defined in [3GPP10-36304]. If the load is above a certain threshold the terminal understands that the cell is highly loaded and evaluates other candidate cells. This threshold may also be broadcast by the cell by making use of a suitable System Information Block or any other similar procedure.

After camping in the cell, and if more than one second has elapsed, the terminal may perform neighbour cells' measurements aimed at possible cell reselections in order to find more suitable cells. Considering intra-frequency reselections, if any of the following conditions are not fulfilled then neighbour cell measurements shall be performed by the terminal device:

$$\left.\begin{aligned} S_{rxlev} &> S_{intrasearch,P} \\ S_{qual} &> S_{intrasearch,Q} \\ L_{serving} &< L_{serving,max} \end{aligned}\right\} \Rightarrow \text{not perform neighbour measurements}$$

where $S_{intrasearch,P}$ and $S_{intrasearch,Q}$ are power and quality thresholds for intra-frequency cell reselection, respectively. Further measurements provide the neighbour cells' received power and quality levels, and cells are ranked according to the quantities R_s, R_n defined for the serving and neighbour cells according to:

$$\left.\begin{aligned} R_s &= 10\log\left[f\left(Q_{meas,s}\right)\times\left(1-L_s/100\right)\right] + Q'_{hyst,s} \\ R_n &= 10\log\left[f\left(Q_{meas,n}\right)\times\left(1-L_n/100\right)\right] + Q'_{off,n} \end{aligned}\right\} \Rightarrow \text{choose the cell with highest R}$$

where f represents any downlink "utility" curve taken as a reference for the technology being considered (e.g. a truncated Shannon curve), and L_s and L_n are the cell loads in serving and neighbour cells (given in %). $Q_{meas,s}$ and $Q_{meas,n}$ are the received signal quality for serving and neighbour cells, and the offsets $Q'_{hyst,s}$, $Q'_{off,n}$ are similar (although perhaps with different values) to the ones defined in [3GPP10-36304]. Given a R_s value for the serving cell and one or several R_n values for the neighbour cells, the terminal reselects to the neighbour cell having the highest R_n value during a time $T_{reselection}$ (also broadcast by the cell) given that it is above the serving value R_s.

Inter-frequency/inter-RAT reselections would also be similar to the intra-frequency case, but with some additional rules for cell measurements that depend on the priorities of the different frequencies/RATs as broadcast by the network (defined in [3GPP10-36304]).

UEs in connected mode would also exploit the cell load indications but in a different way: the BSs should know the cell load indications of neighbour cells for handover decisions, and these can be provided by the UEs as part of normal measurement reports. Neighbour cell loads should take part in handover decisions in such a way that candidate cells with excessive cell loads would not be eligible for handover whenever possible. When the UE switches to connected mode, the serving cell should be aware of the neighbour cell load indications because otherwise it could steer the UE towards a different cell than the one selected by the UE in idle mode, thereby creating ping-pong effects. For this reason the UE should send neighbour cell load indications to the serving cell after entering connected mode. Figure illustrates the process of sending cell load indications by the UE as part of the measurement reports, after entering connected mode, upon request from the serving BS, and after a successful handover to a new serving cell.

In case that cell load indications are not available (or are incorrectly received by the UE) cell loads could be estimated by analysing some of the characteristics of the air interface signals. This alternative however lacks from additional information on backhaul and baseband processing loads, which may also impact load balancing.

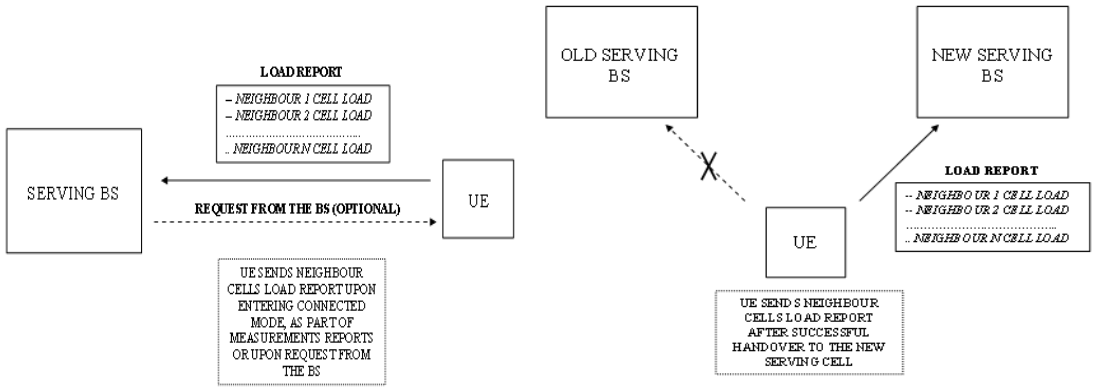


Figure A.1-12. Neighbour cell load report sent by the UE: (a) upon entering connected mode, as part of the measurement reports, or upon request from the base station; (b) after a successful handover

Simulation results for Madrid dense urban scenarios both with macros and with macros + micros are summarized in Table A.1-8. Of special importance is the reduction in the ratio of 50^{th} vs. 10^{th} percentile throughput values, as a measure of the consistency in quality of experience throughout the cell. Average and cell edge (10^{th} perc.) throughput values are also clearly enhanced with minimal network upgrades.

Table A.1-8. Simulation results for Madrid dense urban scenarios, showing relevant KPIs for those cell locations where the best server changed as a result of the proposed load balancing procedure

	macros, w/o load balancing	macros, w/ load balancing	**macros, improvement**	macros + micros, w/o load balancing	macros + micros, w/ load balancing	**macros + micros, improvement**
Average throughput (Mbps)	0.743	1.707	**229.7%**	0.6	1.263	**210.5%**
50^{th} percentile throughput (Mbps)	0.705	1.526	**216.4%**	0.325	1.049	**322.7%**

10th percentile throughput (Mbps)	0.16	0.896	**560%**	0.058	0.194	**334.4%**
50th perc. / 10th perc. throughput	4.412	1.702	**-**	5.591	5.401	**-**

A.1.7 Interference Management for MNs in Ultra-Dense Urban Scenarios

This work is an extension of the work in [MET15-D34] and Section 4.8, the study of moving relay nodes (MRNs)., We showed that the use of half-duplex MRNs can improve the QoS at the VUEs in a noise limited system or in a system with limited amount of interference. This study extends our investigations to the deployment of full-duplex MNs in a densely deployed outdoor urban scenario. The setup is based on the simplified Madrid grid model for METIS test case 2 [MET13-D11], and the simulation parameters are summarized in Table A.1-9.

In this study, each sector of the macro BS is configured with a single antenna, while the receivers at the MN backhaul links are equipped with multiple antennas. Outdoor UEs and VUEs are equipped with single antennas.

Both time domain method and the multi-antenna receivers are used at macro sectors to deal with the inter-cell interference. For the time domain method, the use of ABSs is considered to be used both at macro and micro cells. At the macro cells, ABSs are configured to mitigate the inter-cell interference, while ABSs are configured at micro cells to protect the access links of MNs. For the multi-antenna receivers, in this study we have considered the maximum ratio combining (MRC) and the interference rejection combining (IRC) receivers. In order to ensure fairness between different UEs, a modified proportional fairness scheduling methods based on [JN13] is used, and the regular proportional fairness scheduler is used at the micro cells.

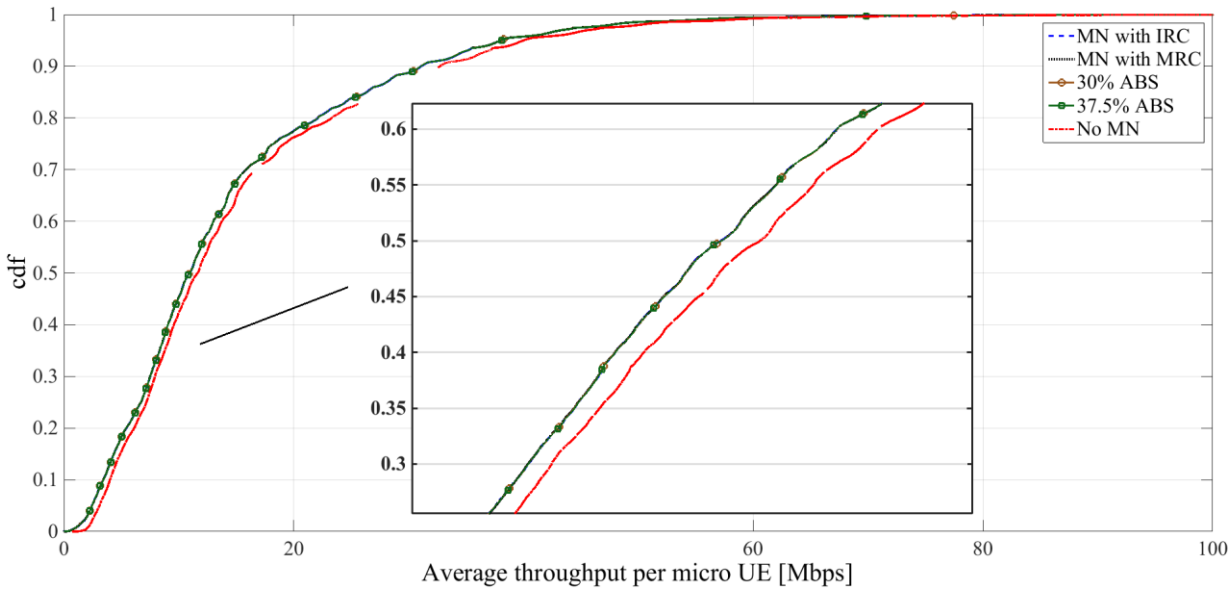


Figure A.1-13. The CDF of micro UE throughput when VPL is at 30 dB

Figure A.1-13 shows the CDF of the micro UEs in the presence of MNs in the system. In this study, we configure ABSs at the micro cells to protect the access links of the MNs. All the micro cells use the same ABS patterns. As we can see, the performance of the micro UEs is

not significantly impacted by MNs if the VPL is high. This is because, firstly, the bandwidth of the micro cell is relatively large, and therefore not so many ABSs are needed at the micro cells. Secondly, high VPL can also attenuate the interference from the access links of the MNs to the nearby micro UEs.

The study in [SGS14] shows that together with the VPL, complicated inter-cell interference enhanced by the street canyon effect is the most important factor that limits the performance of the VUEs in a densely deployed urban scenario. By employing effective interference management schemes, the VUEs can significantly benefit from the use of MNs, as MNs can accommodate more antenna elements, and implement more advanced signal processing methods. Moreover, the studies in [SGS14] also indicate that little impacts are observed on the regular outdoor UEs, after introducing MNs to the system. Detailed evaluation results of using different interference management schemes can be found in [SGS14].

Table A.1-9. Simulation parameters

Component	Configuration Parameters
Buildings and streets	9 buildings 120 meters by 120 meters with 6 floors (3.5 meter height of each floor), based on simplified Madrid grid model. Road width 21 meters (including sidewalks and parking lanes)
Macro BS	Height: 5 meters above the top of the middle building, based on simplified Madrid grid model. Maximum transmit power (per 10 MHz): 43 dBm Carrier: 800 MHz Bandwidth: 20 MHz Antenna configuration: 17 dBi gain, 3 sectors (one antenna per sector), 0, 120 and 240 degrees with respect to the north
Micro BS	Height: 10 meters above the ground close to middle point of south and east walls Positions: see the based on simplified Madrid grid model. Maximum Transmit power (per 10 MHz): 30 dBm Carrier: 2.6 GHz Bandwidth: 80 MHz Cell range expansion bias: 5 dB Antenna configuration: 17 dBi gain, 2 sectors (one antenna per sector), pointing to the main street with an angle of 20 degrees with respect to the closest wall
Moving Network	Full-duplex Speed: 50 km/h Height: 3.5 meters above the ground Position: randomly generated according to a Poisson distribution with λ = 0.5 in each direction of the road Maximum transmit power indoor (per 10 MHz): 10 dBm Carrier: 800 MHz for backhaul links, and 2.6 GHz for access links Bandwidth: 20 MHz for backhaul links, and 80 MHz for access links Antenna configuration both indoor and outdoor: single antenna, 0 dBi gain omnidirectional antenna Receiver noise figure: 5 dB
Outdoor UEs (Macro UE, Micro UE)	Speed: 0 to 3 km/h Height: 1.5 meters above the ground Positions: uniformly randomly dropped, 50 UEs per road Cell selection: based on received power with 5 dB CRE bias

	for micro cells Receiver noise figure: 9 dB
VUEs	Height: 1.5 m above the ground Position: uniformly randomly dropped inside a vehicle. Number of VUEs in each vehicle: uniformly from the interval [1, 50]; Cell selection: 1) same as macro UEs (baseline case); 2) always connect to the MN of their own vehicles (other cases) Receiver noise figure: 9 dB

A.1.8 Scalable Solution for MMC with SCMA

We considered LTE Release 11 as our performance baseline. For a large number of devices in the MMC cases, it was assumed that a user device is in idle state when it has no data to transmit in order to save energy and reduce an interaction with the network. Once a data packet arrives, the device needs to transit to active state from idle state before it can transmit any packets, regardless of packet size being small or large.

To transmit a data packet, a device will start with a random access using physical RACH (PRACH) to establish connection and service setups with eNB and network, as well as sending scheduling request to get granted resources for uplink transmissions. As a result, it may incur significant signalling overhead and a large latency to transmit data packets, especially when irregular, single small packets (e.g., as short as 20 bytes) are transmitted. It is possible that more than one user wakes up and access the network at the same time for data transmissions. As they randomly select preambles from a predefined set for contention access, RACH collisions due to a limited number of preamble resources may happen, where a RACH collision means two or more users (randomly) select the same preamble in the same access region. We assume that packets will be dropped if a RACH collision happens in any access group; while if no RACH collision occurs, there will be no access/signal detection errors at the receiver end. As a result, we evaluate the LTE-A Release 11 performance by considering the packet drop rates resulting from RACH collisions. The performance criterion used here is the number of devices that can be supported for a given average packet drop rate in the system (e.g., 1%). Detailed parameters are given in the following table.

Table A.1-10. Evaluation assumptions for LTE-A Release 11 baseline [3GPP11-37868]

Parameters	Setting
PRACH configuration index	14
Preamble format	0
Total # of preambles used for random access	64
Traffic model [TC11 NRT]	Poisson distributed arrivals with IAT of T (300 seconds)
Number of MTC devices (N_d)	System/cell loading: λ = N_d/(T*1000) (arrivals/ms)
Packet access (packet TX) failures	Due to RACH collisions only; No signal detection errors were assumed for non-collision access and signal transmissions @ the receiver

Capacity threshold metric:	1%packet TX failure rate

SCMA performance was fully simulated by uplink system-level simulator. The evaluation criterion used here is the number of devices that can be supported for a given average packet drop rate (e.g., 1%) in the system. System-level simulation for the performance evaluation is performed on a 19-cell 3-sector network, assuming the 2 GHz carrier frequency. Different numbers of users are dropped randomly in the network and statistics are collected for a sector in the centre cell. Uplink traffic for each user follows a Poisson distribution with a mean packet inter-arrival time of 300 s. A specific traffic load in each sector is obtained by configuring a different number of active users in each sector. A contention region is defined by a size of 4 LTE RBs over one OFDMA sub-frame (TTI), and the flat Rayleigh fading channels are modelled.

For a fair comparison of the SCMA and LTE Release 11 schemes, same data size that fits into the equivalent of one RB OFDMA resource is considered with a fixed spectral efficiency of 1 bit/subcarrier. The data transmitted in SCMA is spread over the 4 RBs (i.e., each encoded bit is spread into 4 subcarriers); for LTE baseline, one packet is transmitted using one of the 4 RBs. OLPC from LTE is applied to determine the transmission power at the user terminal with SCMA 4-point 6 codebooks. Each of the neighbour cells experiences the same traffic load as the cell of interest (i.e., the center cell). As a result, the time-varying interference from the neighbouring sectors to the center sector can be captured. The MPA receiver is employed for SCMA to handle non-orthogonal joint blind signal detection. Detailed simulation parameters are given in the following table.

Table A.1-11. Simulation setup and assumptions for SCMA

Parameters	Settings
Network layout	19 cells with 3 sectors per cell
Inter-site distance	500 m
Channel	Flat Rayleigh fading
Antenna configuration	SIMO 1x2, uncorrelated antennas
Modulation & coding	SCMA: Codebook size of 6, Spreading factor 4 with 2 non-zero elements, code rate ½
Channel estimation	Perfect
Traffic pattern	Packet size of 20 bytes. Poisson packet arrival with IAT of 300 s; configurable number of active users in each sector
Interference Model	Same distributed traffic in neighbor cells
OLPC	$\alpha = 0.95$, $P0 = -93$ dBm
Receiver model	SCMA multiuser detection with MPA

A.1.9 Downlink Multi-User SCMA for Mobility-Robust and High-Data Rate Moving Network Mobility (MN-M)

To further improve the spectral efficiency, SCMA can be combined with MIMO techniques. In this section, we propose some open-loop MIMO transmission schemes for SCMA.

MIMO for multiplexing: Open-loop single-user and multi-user multiplexing exploiting both space and code domains to multiplex layers of data stream. This combined technique is called

code-space multiplexing (CSM) which is the extension of the spatial multiplexing (SM). Ignoring the complexity issue, one can apply ML detector over all combined SCMA layers of all MIMO streams. However to enjoy the sparsity feature of SCMA and hence limit the complexity of detection, the MIMO streams are first separated by a MIMO detector and then SCMA layers of each MIMO stream are decoded separately following the MPA reception technique.

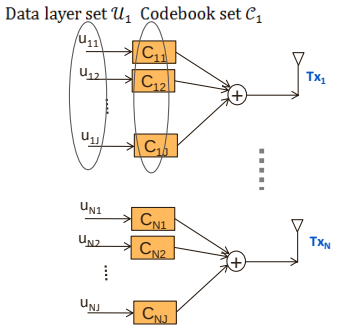


Figure A.1-14. MIMO CSM for SCMA

MIMO for transmit diversity: Block-wise space-time/frequency coding is introduced to apply space-time coding techniques such as Alamouti over SCMA blocks to achieve further diversity. Figure A.1-15 illustrates the block-wise Alamouti code for SCMA. A block-wise Alamouti detector is used at the receiver to recover the SCMA block before feeding into the MPA detector. Notably, the columns of the block can be re-arranged to control the inter-block interference caused by the channel variation.

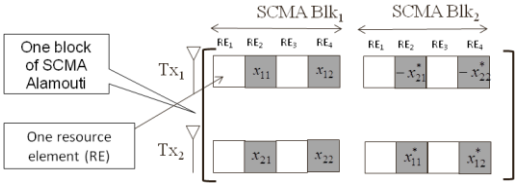


Figure A.1-15. Block-wise Alamouti encoder for SCMA

DL MU-SCMA enables open-loop user multiplexing with robustness to mobility and low rate of feedback requirement. With a very limited need for channel knowledge in terms of CQI, a transmit point (TP) simply pairs users together while the transmit downlink power is properly shared among multiplexed layers in code domain. Code domain pairing along with near optimal MPA detector provides a flexible multiplexing and UE pairing solution. Multiplexed layers are separated and decoded at the terminals using MPA receiver. Compared to MU-MIMO, DL MU-SCMA is more robust against dynamic channel variations in high speed scenarios and less sensitive to the error in channel information. DL MU-SCMA is evaluated to show the advantage of SCMA user pairing to increase throughput of downlink for a highly loaded network for both low and high speed users.

SCMA can provide an open-loop CoMP solution without knowledge of short term multi-TP CSI. It can bring two main advantages to the system: i) dramatic reduction of the overhead caused by dynamic multi-TP CSI feedback, and ii) significant increase of throughput and coverage as well as the robustness to channel aging.

To see the benefit of SCMA in CoMP scenario, DL MU-SCMA/CoMP is implemented and evaluated in various scenarios. These scenarios are illustrated in Figure A.1-16.

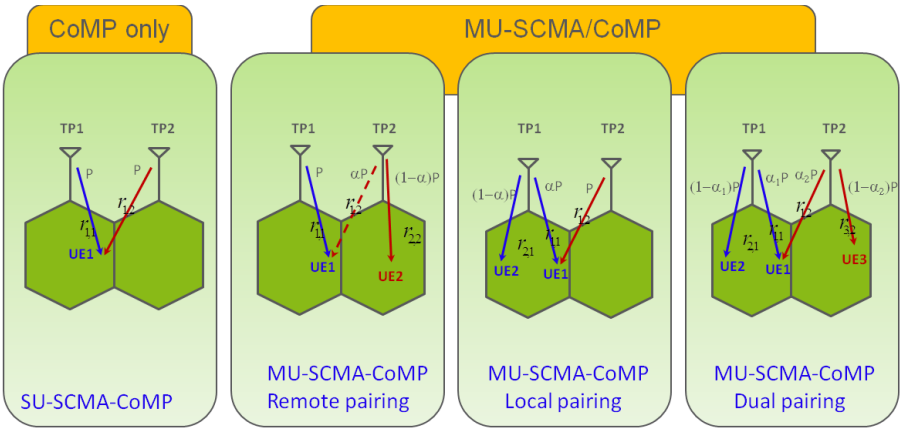


Figure A.1-16. Typical DL MU-SCMA/CoMP scenarios

A system-level simulation tool that is based on simplified Madrid grid has been implemented. This tool has been used to evaluate DL OFDMA and DL MU-SCMA/CoMP for both low (3 km/h) and high speed (120 km/h) scenarios. For full buffer traffic, the aggregate cell throughput and 5%tile coverage rate (cell edge throughput) are evaluated. For non-full buffer traffic, the delay distribution and perceived throughput are evaluated. The performance of OFDMA is compared with MU-SCMA in CoMP and non-CoMP transmission modes. Antenna pattern and path-loss of Madrid grid follows the METIS model and evaluation methodology [MET13-D61]. Users are stationary for the duration of the simulation. The list of simulation assumptions and parameters for full buffer and non-full buffer are summarized in Table A.1-12.

Table A.1-12. Simulation assumptions for full buffer and non-full buffer simulation of DL MU-SCMA and DL OFDMA

Parameter	Value
Deployment	Simplified Madrid grid with wrap around, 18 micro cells with outdoor users only
Number of users	560 users uniform drop
TP transmission power	28dBm for micro cell
Number of transmit antennas	2 uncorrelated antennas
Number of receive antennas	2 uncorrelated antennas
System bandwidth	10 MHz at 2.6 GHz carrier frequency
Channel type	ITU-TU fading channel
Transmission Mode	2×2 SM or SFBC, open-loop
User speed	3 km/h or 120 km/h
HARQ	Incremental redundancy (IR) HARQ with up to 3 retransmissions
Scheduler	Proportional fair scheduler
RBG size for scheduling	50 RBs for wideband scheduling
Waveform	Multiuser SCMA, OFDMA
SCMA codeword dimension	4 OFDMA tones
SCMA receiver assumption	MPA joint detector PHY abstraction
OFDM receiver assumption	MLD PHY abstraction
CQI feedback	Perfect CQI. Feedback report every 10 TTIs
OLLA	Enabled with 10% BLER for first transmission
MCS table	Follows LTE standard for OFDMA and specific MCS table for SCMA
MU-SCMA	Follows designed pairing algorithm, power allocation, and multi-user detection scheme
Traffic model	Full Buffer, or FTP2 non-full buffer with 0.1sec inter arrival time and exponential distribution and different packet sizes to set network load

The relative throughput and coverage gains of MU-SCMA over OFDMA are shown in Figure A.1-17 with full buffer traffic for both SM and Alamouti transmission modes and low and high speed scenarios. As illustrated in this figure, the relative gain is stable regardless of the user speed. The relative gain is between 23-48% for SM mode and 50-95% for Alamouti mode.

These results confirm the capability of MU-SCMA to provide high throughput and user experience independent of the user mobility status and their speeds. Moreover, the relative gain of MU-SCMA over OFDMA maintains for high speed scenario where the closed-loop multiple access schemes such as MU-MIMO fails due to the channel aging effect and high rate of CSI feedback signaling.

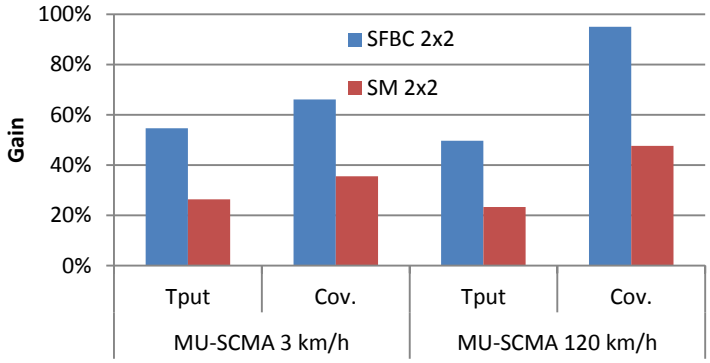
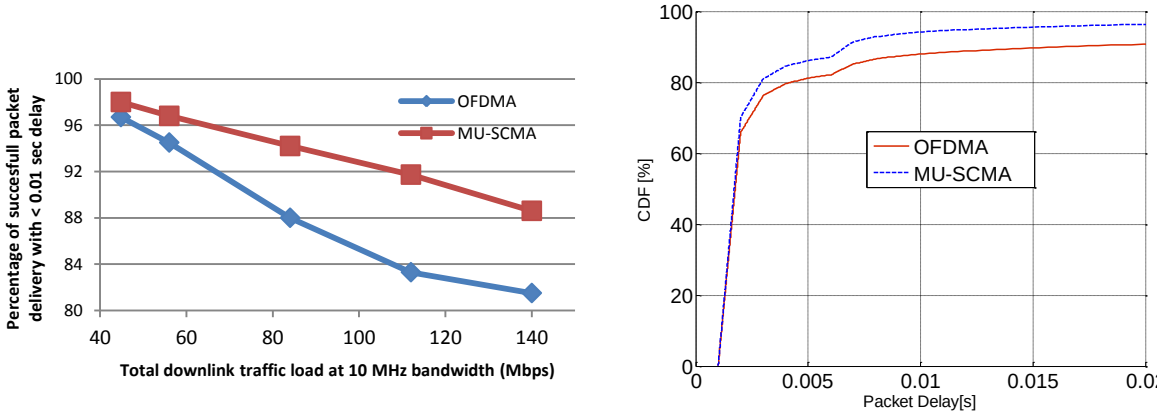


Figure A.1-17. Performance gain of DL MU-SCMA over DL OFDMA for low and high speed scenarios with full buffer traffic

The system-level simulation results with non-full buffer traffic are shown in Figure A.1-18 and Table A.1-13 for high speed scenario (120km/h) and SM transmission mode. The relative gain of DL MU-SCMA over OFDMA in high speed scenario is shown in Figure A.1-18(a) in terms of the total supported load for the required percentage of successful packet delivery with 10 ms delay constraint. As illustrated in this figure, assuming 10 ms delay requirement for 95% of packets, MU-SCMA can support more than 41% higher load compared to OFDMA. The packet delay distribution for DL MU-SCMA and DL OFDMA is shown in Figure A.1-18(b) for a total DL traffic load of 84 Mbps at 10MHz bandwidth in high speed scenario. As can be observed, MU-SCMA benefits from lower packet delay compared to OFDMA.



(a) **(b)**

Figure A.1-18. Non-full buffer simulation results for high speed scenario and SM mode: a) Performance gain of MU-SCMA over OFDMA for with non-full buffer traffic, (b) Delay distribution of DL MU-SCMA vs. DL OFDMA for a total DL traffic load of 84 Mbps at 10 MHz bandwidth in high speed scenario

Table A.1-13 shows the performance comparison between DL OFDMA and DL MU-SCMA in terms of the 90% and 95%tile of perceived throughput per user. Perceived throughput is defined as the ratio of packet size over the total delay of the packet delivery. As can be observed, MU-SCMA provides 495% perceived throughput gain at 95%tile throughput CDF, which means that among the packets that satisfy 10 ms delay constraint, MU-SCMA delivers packets much faster than OFDMA.

Table A.1-13. System-level simulation results comparing OFDMA and DL MU-SCMA for high speed users and SM mode with non-full buffer traffic: 90% and 95%tile of perceived throughput

Perceived Tput (Mbps)	At 95% CDF	At 90% CDF
OFDMA	0.2	0.92
MU-SCMA	1.19	2.25
Gain	495%	145%

The gain of MU-SCMA significantly improves if it is combined with CoMP techniques. The relative gain of MU-SCMA/CoMP is shown in Figure 2-14.

A.2 Mobility Management and Robustness Enhancements

A.2.1 Efficient Service to Layer Mapping and Connectivity in UDN

For efficient service to layer mapping and connectivity in UDN, three main approaches were considered: (a) purely UE autonomous action based on prediction whether a better connection opportunity will become available in time, or (b) with information exchange between the UE and the network, or (c) the cloud can also assist the UE by indicating when there is more capacity available. For the UE autonomous approach, the main idea is to use alternate connectivity options in UDN, based on a longer term prediction of the available connectivity options. For the second approach, the network provides assistance information to the UE to help decide when to initiate the tasks that are delay tolerant in nature, possibly requiring transfer of significant amount of data. In the third approach, UE is assumed to provide information regarding amount of data and delay tolerance, and wireless connectivity characteristics to the network or cloud. The cloud server could in turn inform the UE regarding the order of synchronization tasks that needs to be done.

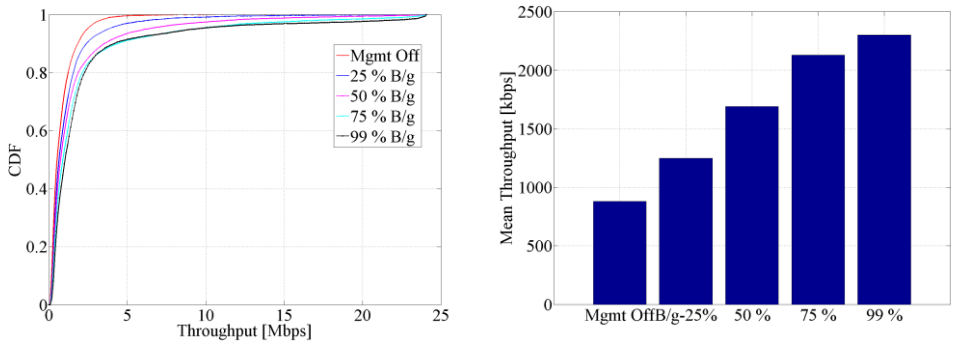


Figure A.2-1. (a) User throughput [Mbps] distribution, and (b) Mean user throughput [kbps] values for various background traffic rates, with and without traffic management

Simulations were conducted for this component using METIS test case 2 [MET13-D61], with 40 UEs / macro cell, deployed outdoors, moving in random directions at 3 km/h. The macro/pico/UE transmit powers are 46/30/21 dBm respectively, with 2 GHz operating frequency and 10 MHz bandwidth. METIS antenna model was used with full buffer traffic, with three small cells deployed in the scenario. Varying number of UEs with low-priority background (B/g) traffic ratios were simulated, with a file size of 100 MB, and packet data unit (PDU) size of 2 MB.UE power consumption based on transmit power to units/subframe linear mapping done in [3GPP13-36843] is also presented. The model basically assumes an increase from 0 – 15 units per subframe of UE power consumption, for an increase of 0 – 23 dBm of UE TX power. The evaluations were done using a generic connectivity management approach of delay all the background data until UE is connected to a small cell, which is common to all the three approaches considered.

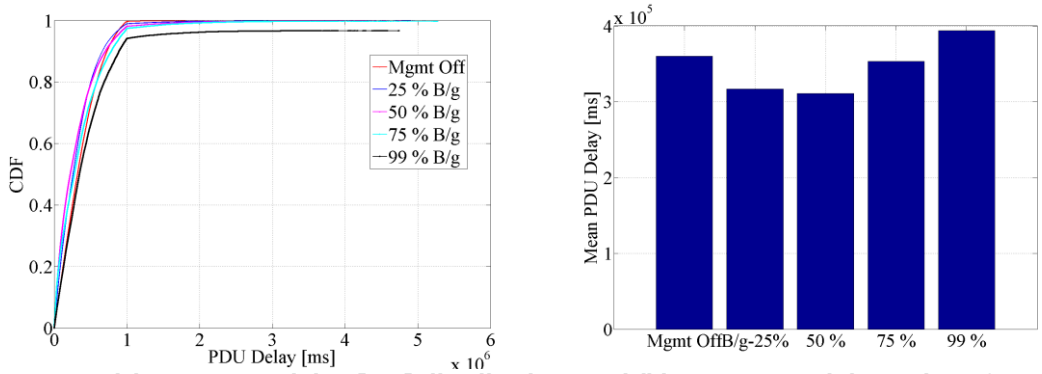


Figure A.2-2. (a) User PDU delay [ms] distribution, and (b) Mean PDU delay values for various background traffic rates, with and without traffic management

The user throughput distribution and mean user throughput values with connectivity management off and with management, having different percentages of background traffic, is shown in Figure A.2-1. From the figure, we can observe that there are significant gains in user throughput by using the connectivity management approach, mainly due to the user throughput gains for users with non-background traffic in the macro cell. Due to the use of full buffer model, the resources made available by delaying low-priority traffic in macro cells are utilized for sending high-priority traffic by the other users. From the figure we can also observe that the throughput gains are also linked to the amount of low priority traffic present in the scenario as well.

The main impact of delaying the low priority traffic can be observed in the PDU delay results presented in Figure A.2-2 as well. From the CDF distribution and mean values, we can observe that while the PDU delay initially decreases with low amount of background traffic (<50%), the delay eventually increases as the amount of background low-priority traffic exceeds the amount of high-priority traffic, leading to a net total delay increase. The increased delay is due to the buffering/delayed transmission of low priority data, with the high-priority data being sent as soon as it arrives in the network.

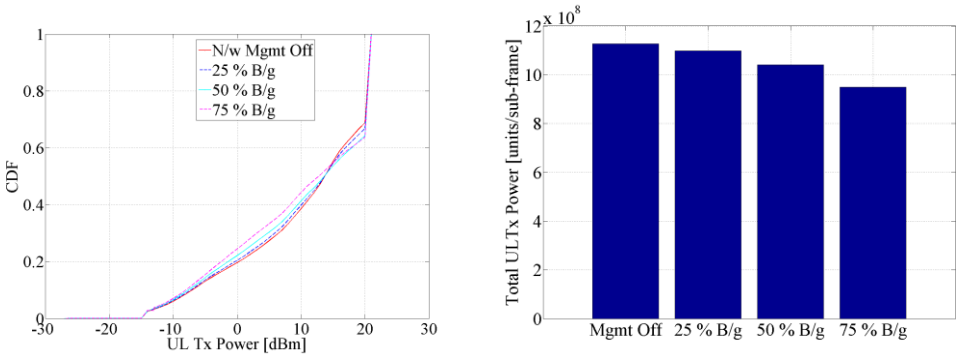


Figure A.2-3. (a) Uplink transmit power distribution, and (b) Total UL transmit power in units/subframe

The uplink transmit power distribution and total transmit power in units / subframe results is shown in Figure A.2-3. From the figure we can observe that, by delaying the transmission of low-priority data until the UE connects to a small cell significantly reduces the UL transmit power, and consequently reduces the total amount of UE battery power consumption as well. From the total transmit power consumption values, it can be observed that, depending on the background traffic ratio, up to 16% total power consumption reduction can be obtained.

The considered connectivity management approaches have several advantages such as improved offloading of low priority traffic to the ultra-dense small cell network, balanced

network load with reduced peak network capacity due to low-priority traffic, improved UE power efficiency, and enabling cloud control of UE synchronization issues, with UE assistance information. The approaches would also require the ultra-dense deployment of small cells, with possibly slightly higher signalling load for exchanging information between UE, the radio access network and the cloud server.

A.2.2 Small Cell Mobility Enhancements in Multicarrier and mmW Small Cell Network - Network Assisted Small Cell Discovery

The currently available mechanisms in LTE-A for finding offloading opportunities in inter-frequency small cells are somewhat limited. The network can configure UE with inter-frequency measurements for those carriers where the small cells have been deployed. However, if this number of carriers becomes large, blindly searching all of them consumes a lot of energy and causes additional delay as the UE does not always search the correct carrier first. Moreover, as the small cells have limited coverage, the UE may end up spending lot of energy searching unnecessarily in areas where small cells have not been deployed. Therefore we propose a network assisted small cell discovery solution that addresses this problem. The UE is configured with radio fingerprints (e.g. RSRP measurements of macro cells) to limit the search area to only those locations and carriers where small cells are operating.

We have carried out static system simulations to assess the feasibility and performance of the proposed solution. The simulations were conducted using METIS Test Case 2 urban outdoor scenario [MET13-D61], where a 3-sector macro cell was deployed in the middle and 12 small cells provided high data rate hotspots. We have assumed that the network wants to offload UE to small cell layer when there is a small cell with RSRP above -75 dBm available, which in this scenario means that the small cells cover approximately 20% of the outdoor area. The macro and small cells were deployed on different frequencies. Figure A.2-4 illustrates the scenario and shows the small cell coverage (RSRP). UEs were assumed uniformly distributed in the outdoor simulation area. It is assumed (as in LTE-A) that UE performs inter-frequency measurements periodically and thus the energy consumption due to the inter-frequency cell search is proportional to simulation area where the UE searches for inter-frequency small cells.

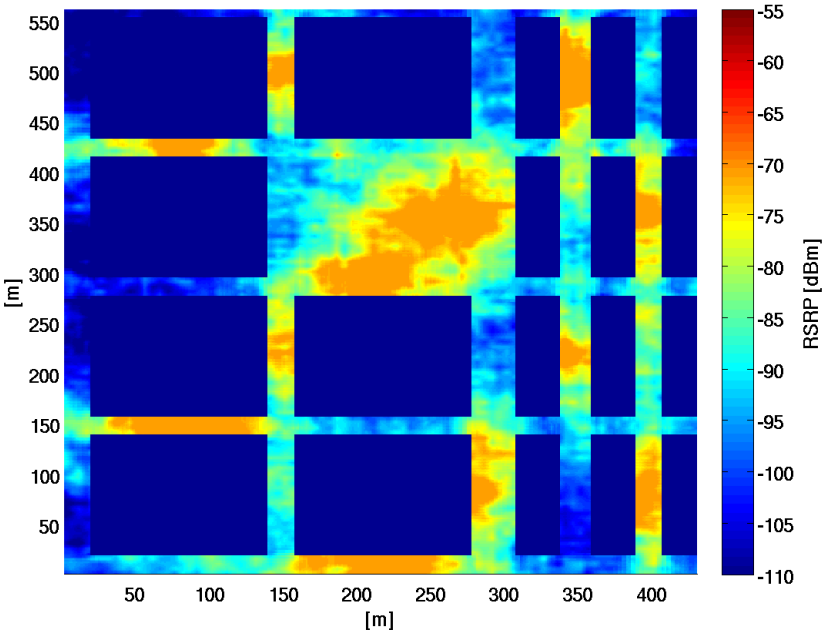


Figure A.2-4. Simulation scenario. Small cell RSRP outdoors (in the simulations, UEs were restricted to outdoors, so the indoor coverage is not considered) is shown

The network signals UE set of macro cell radio fingerprints (consisting of macro cell ID + RSRP for 3 strongest macro cells). Each of these corresponds to a small cell location so that there are N fingerprints per small cell. We have simulated values N = 1, 3, 5, 9 and 12. The fingerprints are obtained by k-means clustering the UE measurement samples within small cell coverage area. A fingerprint is considered to match when the 2-norm of the difference between current measurements and fingerprint is below a threshold T. Different values for threshold T ranging from 1 dB to 15 dB have been considered to obtain the trade-off curve between fingerprint coverage and measurement power saving as shown in Figure A.2-5.

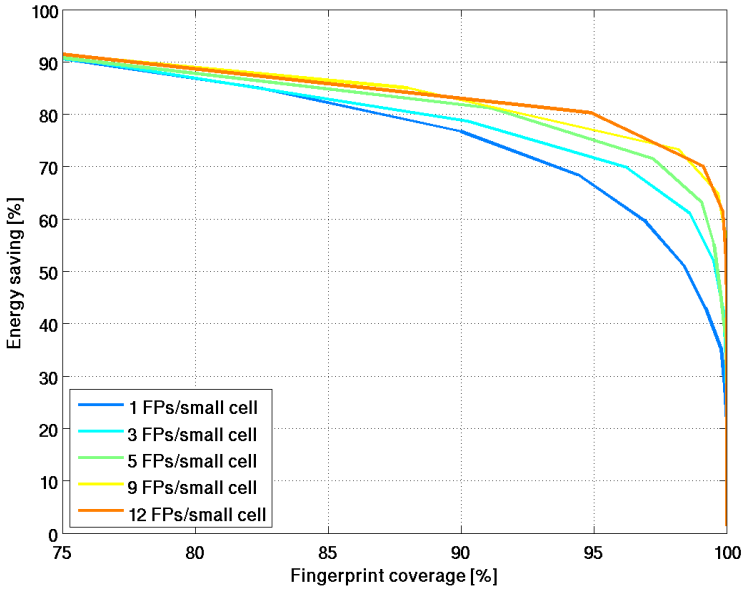


Figure A.2-5. Trade-offs between UE power saving (of power consumption due to the inter-frequency measurements) and fingerprint coverage (area of small cell coverage where there is a matching fingerprint) for different number of radio fingerprints per small cell

Taking a closer look at the results, Figure A.2-5 shows that increasing the number of fingerprints per small cell allows more accurate mapping of small cell coverage and thus a better trade-off between fingerprint coverage and energy saving. However, smaller number of fingerprints would be preferred to keep the signalling overhead to the minimum. Thus 1 to 3 fingerprints samples per small cell seem perhaps the most attractive option. Even with just a 1 fingerprint sample per small cell we can obtain up to 65% energy saving with 95 % accuracy of small cell coverage, when compared to UE searching for inter-frequency small cells in the entire outdoor area of the scenario. With a 12 fingerprints sample per small cell a more accurate mapping is achieved and the energy saving goes up to 80%.

In a deployment as in Figure A.2-4, LTE-A UE would need to be configured to search for inter-frequency small cells pretty much continuously, to avoid missing any offloading opportunities. This would cause a significant energy consumption burden for a UE: With current LTE-A assumptions (according to [3GPP14-36133] and [3GPP14-36331]) inter-frequency measurements mean 5 ms of measurements every 40 ms, i.e. approx. 12.5% energy consumption compared to having receiver on continuously. From this we can roughly estimate that the impact on UE power consumption is in the order of 10 % of DL energy consumption (if UE has full traffic activity, but more than this with a more realistic bursty traffic pattern and DRX where the share of measurements of all receiver activity is increased). Saving 80% of this would be a substantial improvement. UL power consumption depends on the traffic assumptions as well, but with typical DL-heavy traffic it is below DL power consumption.

The simulation results showed that the proposed network assisted small cell discovery based on macro cell radio fingerprints can reduce this by 65-80% with a reasonable initial signalling overhead (sending the fingerprint data to the UE). Additionally, if there are small cells deployed on several different carriers, UE's measurements can be directed towards the correct carrier and thus also the latency of offloading can be reduced, as UE needs to search fewer carriers. Measuring more than one inter-frequency carrier does not consume more power in LTE-A, but slows down the search proportionally to the number of carriers [3GPP14-36133].

A.2.3 Handover Optimization for Moving Relay Nodes

This work is an extension of the work in [MET15-D33], the study of moving relay nodes (MRNs). In [MET15-D33], we showed that the use of half-duplex MRNs can improve the QoS at the VUEs in a noise limited system or in a system with limited amount of interference. In considered studies, we take the mobility of the VUE into account, and proposed a framework to optimize the handover parameters for the VUEs. Similar to work in [MET15-D55], half-duplex decode-and-forward (DF) MRNs are considered in this study.

The handover is triggered by the A3 event defined in the 3GPP LTE standards as "Neighbor cell becomes better than an offset relative to the serving cell" [STB11, Ch. 3]. The offset in the A3 event is also referred to as the HO hysteresis margin. In order to avoid the so called ping-pong effect, i.e., a UE is being handed over back and forward between source and targeted BSs, the A3 event needs to be observed for a given period, TTT, before the handover can be performed.

Two metrics are used to evaluate the handover performance, i.e., the average power OP, and the average number of ping-pong handovers. In order to lower the OP at the VUEs, a shorter TTT and smaller handover margin is preferred, as the VUE can be quickly handed over to a targeted BS. However, due to fading and shadowing, this might result in excessive number of ping-pong handovers. Therefore, a balance needs to be achieved.

For such a multi-objective optimization problem, we usually focus on one of the objectives while keeping the other below a given threshold. The technique is given in details in [SRS+13], where the average power OP is minimized while the average number of ping-pong handovers is kept below a given number. Different handover parameters are obtained at different VPL, and presented in [SRS+13].

Table A.2-1. Simulation Parameters

Parameter	Value
Inter-site distance	1732 meters
Average BS transmit power	46 dBm
Average FRN transmit power	30 dBm
Average MRN transmit power	20 dBm
Carrier frequency	2.0 GHz
System bandwidth	10 MHz
Receiver noise figure for both the RN and the VUE	9 dB
Normalized minimum required rate at the VUE	R= 1 bit/s/Hz
VUE velocity	120 km/h
Handover critical time	5 seconds
Handover performance observation time	200 seconds
Shadowing de-correlation distance	50 meters

The performance of the MRN assisted transmission is evaluated through system level simulations, and the simulation parameters are given in Table A.2-1. The BS-to-VUE direct transmission is used as the baseline, and the FRN assisted transmission is also used for

comparison. The FRN position is optimized through exhaustive search, which minimized the end-to-end average power OP at the VUEs. Detailed simulation results are given in [SRS+13], and the results show that the benefits of using MRNs to serve the VUEs can be observed from a VPL of 10 dB and above.

A.3 Context Awareness Approaches

A.3.1 Context Awareness Through Prediction of Next Cell

Mobility of commuters is not purely random but rather direction oriented and may be learned after monitoring user movements for a couple of business days. Exploiting movement data and context information of diurnal user movements (public transportation, vehicular users, etc.) allows for predicting cell transitions and lays the basis, e.g., for designing efficient resource reservation schemes or smart resource mapping approaches. In real life scenarios, several mobile users co-travel in public transport forming data intensive moving user clusters or moving networks. Various load balancing solutions exist to manage congestion situations that could arise. However, the crucial trigger for these solutions is timely prediction of arrival of moving user clusters or moving networks into a cell. This paper presents prediction and detection schemes that exploit context information for predicting user cell transitions and resulting congestion. These schemes are utilized to anticipate the arrival of data intensive moving user groups/moving networks, which are also referred to as “hotspot-situation, into a cell. Simulation results demonstrate robust and timely prediction of these events and their applicability for handover optimization and smart resource management even at high velocities.

In case of a random walk mobility model a user can travel in all six directions with equal probability from its current cell. However, in a diurnal mobility model, user group direction is probable in only two directions (e.g. streets, train tracks) and zero in other directions. Hence, a user group can transit into one of the two adjacent cells from its present cell. If ϕ_1 and ϕ_2 are angles made by user’s direction with respect to the closest directions leading towards the center of neighbouring cells, then the probability of the user transition into those neighbours are, P_1= (1- ϕ_1/60) andP_2= (1- ϕ_2/60).These are probabilities based on user angle.

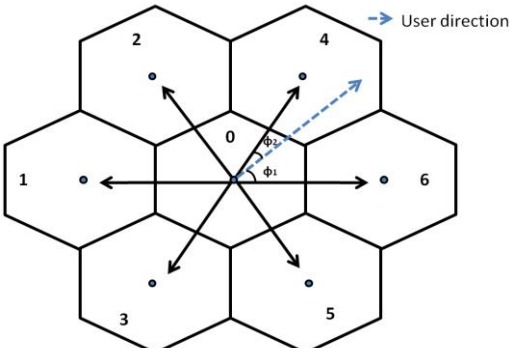


Figure A.3-1. Diurnal mobility model

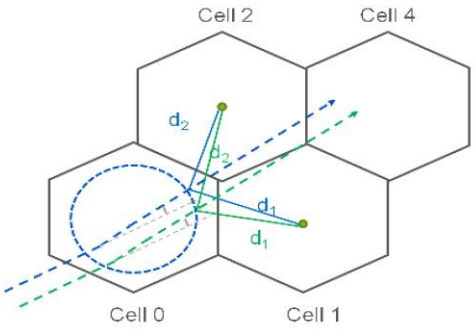


Figure A.3-2. Distance-based approach

The user motion indicated in green and blue lines in Figure A.3-2have the same estimated user direction (angle). But the next cell depends on the position of the user trajectory at the

circumference of the circle. At the point of prediction, if d_1 and d_2 are the distances of user from the centers of cell1 and cell2, respectively, then the probabilities of transition to these cells based on distance are,

$$P_1 = 1 - \frac{d_1}{d_1 + d_2} \text{ and } P_2 = 1 - \frac{d_2}{d_1 + d_2}.$$

It is possible to combine the probabilities from the above two approaches to obtain a combined probability. If there are three possible candidates to be the next cell, then the probabilities of transition based on user distances are:

$$P_1 = \frac{2}{3} - \frac{d_1}{d_1 + d_2 + d_3}, \quad P_2 = \frac{2}{3} - \frac{d_2}{d_1 + d_2 + d_3}, \quad P_3 = \frac{2}{3} - \frac{d_3}{d_1 + d_2 + d_3}.$$

Probabilities based on combined approach are:

$$P_1 = \frac{(3+2\alpha)}{3(1+\alpha)} - \frac{\phi_1}{60(1+\alpha)} - \frac{\alpha}{(1+\alpha)} \frac{d_1}{(d_1 + d_2 + d_3)}$$

$$P_2 = \frac{(3+2\alpha)}{3(1+\alpha)} - \frac{\phi_2}{60(1+\alpha)} - \frac{\alpha}{(1+\alpha)} \frac{d_2}{(d_1 + d_2 + d_3)}$$

$$P_3 = \frac{2\alpha}{3(1+\alpha)} - \frac{\alpha}{(1+\alpha)} \frac{d_3}{(d_1 + d_2 + d_3)}$$

Where, distance based component is weighed by $\alpha > 1$.

It is possible to derive probabilities of transition into next cells based on geometry (dB) of the user with respect to its neighboring cells. If geo_1, geo_2 and geo_3 are the geometry of the user w.r.t. potential next cells, then probabilities of transition are given by,

$$P_1 = \frac{geo_1}{geo_1 + geo_2 + geo_3}, \; P_2 = \frac{geo_2}{geo_1 + geo_2 + geo_3}, \; P_3 = \frac{geo_3}{geo_1 + geo_2 + geo_3}$$

Table A.3-1. Comparison between the angle-based, distance-based, combined and geometry based approaches

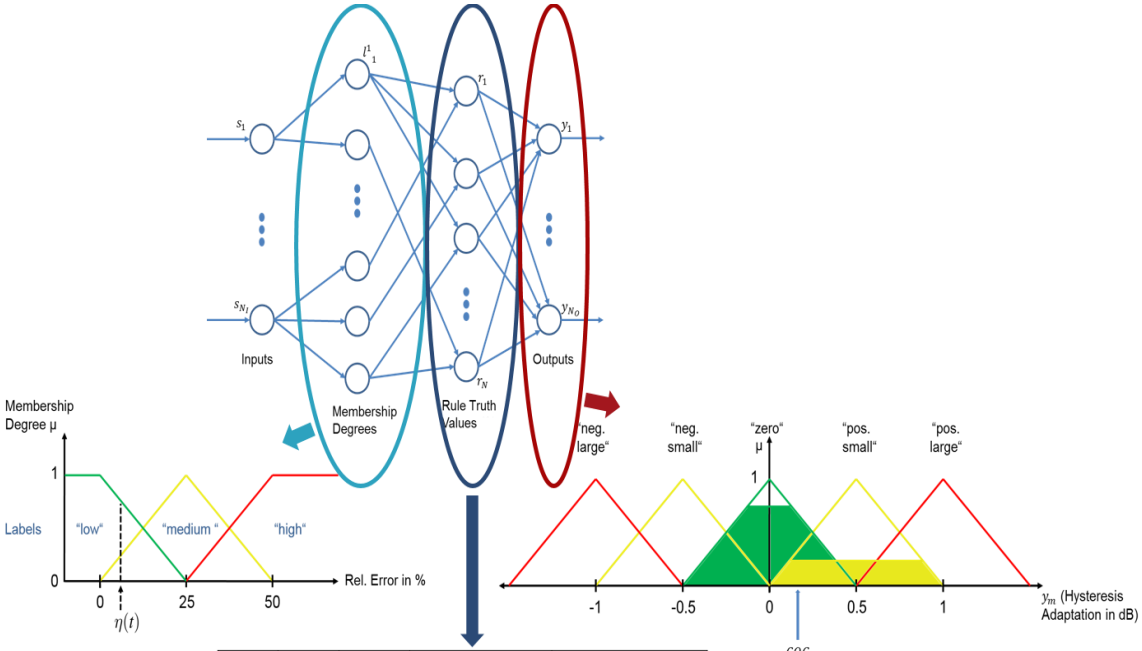
Present Cell	Set of Next Cells	P_{angle}	$P_{distance}$	$P_{combined}$	$P_{geometry}$
Cell 3	Cell 5	0.25	0.346	0.327	0.084
	Cell 0	0.75	0.410	0.478	0.847
	Cell 1	0	0.243	0.194	0.067
Cell 0	Cell 6	0.25	0.447	0.408	0.807
	Cell 4	0.75	0.353	0.433	0.161
	Cell 2	0	0.198	0.158	0.032
Cell 6	Cell 18	0.25	0.202	0.211	0.001
	Cell 15	0.75	0.332	0.416	0.004
	Cell 4	0	0.465	0.372	0.995
Cell 4	Cell 15	0.25	0.468	0.424	0.978
	Cell 13	0.75	0.311	0.399	0.002
	Cell 11	0	0.219	0.175	0.020

The entries marked in blue indicate probabilities of transition into actual next cell. Geometry based approach has a higher probability of prediction compared to rest of the schemes. Distance based method is also capable of predicting next cell but with probability lesser than geometry based. Angle based approach fails to predict transition into third next cell and this affects combined approach as well.

A.3.2 Context Aware Mobility Handover Optimization Using Fuzzy Q-Learning

Robust mobility support that is able to adapt to local conditions is a challenging but important feature of current and future RANs. Due to user mobility and thereby triggered handover processes that are subject to RAN-specific delays, which may severely affect users' end-to-end performance, autonomous controlling of system behaviour and service provisioning are crucial for Mobile Network Operators (MNOs) in order to mitigate OPEX and to improve mobile users' QoE.

In order to reduce the degree of human intervention in network optimization processes, a fuzzy Q-learning based self-learning and self-tuning mechanism has been developed that aims at optimizing the robustness of mobility support according to locally observed conditions, thus establishing context awareness. In particular, optimal handover parameter settings are determined and enforced autonomously that yield an appropriate trade-off, e.g., with respect to connection drops, handover failures, and ping-pong handovers. Figure 5-4 exemplarily depicts the different processing steps of the developed scheme.



Rule	HYS	CDR	HFR	PHR	Actions
1	low	low	low	low	No action
2	low	low	low	≠ low	{HYS ↑, HYS ↑ & TTT ↑, TTT ↑}
3	low	low	≠ low	-	{HYS ↑ & TTT ↑, TTT ↑, HYS ↑}
4	low	≠ low	-	-	{HYS ↓, HYS ↓ & TTT ↓, TTT ↓}
5	high	low	low	low	No action
6	high	low	low	≠ low	{TTT ↑, HYS ↑ & TTT ↑, HYS ↑}
7	high	low	≠ low	-	{TTT ↑, HYS ↑ & TTT ↑, HYS ↑}
8	high	≠ low	-	-	{HYS ↓ & TTT ↓, HYS ↓, TTT ↓}

Figure A.3-3. Exemplary fuzzy Q-learning scheme (including fuzzy input classification, adaptation rule base, and output calculation)

First, relative deviations $\eta(t)$ of the respective KPI $\rho(t)$ from the pre-defined performance targets ρ^* are classified using fuzzy logic, where$\eta(t) = \frac{\rho(t)-\rho^*}{\rho^*}$. For that purpose, membership functions of different shape and width have to be defined for assessing severity of KPI deviations. Further, the fuzzy inference system checks based on the membership degrees of classified inputs and its fuzzy rule base which rules apply to the given situation or system

state. The rule truth values activate the respective output membership functions that are used to determine the corresponding parameter adaptation, e.g., using the centre-of-gravity (COG) method.

At the next time step $t+1$, a reinforcement signal $r(t+1)$ is received by the learning entity, which allows to assess the effectiveness of the previously applied adaptation and which is defined as follows:

$$r(t+1) = \sum_{m=1}^{N_{KPI}} w_m \left(\rho_m^* - \rho_m(t)\right),$$

where ρ_m^*denotes the respective KPI target value and $\rho_m(t)$the observed and measured KPI at time t, respectively. $(w_m)_{m=1,\dots,N_{KPI}}$ represent the weights assigned to the respective KPIs specified by the MNO policy. The reinforcement signal is further normalized to yield feedback values in the range of [-1,1]. The overall fuzzy Q-learning scheme is listed in Figure 5-4.

1. Initialize Q-value LUT: $\forall i \in \{1,2,\dots,N\}, \forall k \in \{1,2,\dots,N_{\mathcal{A}^i}\}$
$q(i,k) = 0$ and set time $t = 0$.
Repeat:
2. Receive and classify system state
$s(t) = \left(H_s(t), TTT(t), O_s(t), O_t(t), \rho_{\mathrm{CDR}}(t),\ \rho_{\mathrm{HFR}}(t),\ \rho_{\mathrm{PHR}}(t)\right)$.
3. For each rule i select an action k following the EEP:
$k = \arg\max_j q(i,j)$ with probability ε
or k = random $\{j,\ j = 1,2,\dots,N_{\mathcal{A}^i}\}$ with probability $1-\varepsilon$.
4. Calculate inferred output using COG method.
5. Determine its corresponding quality $Q(\boldsymbol{s},\boldsymbol{a})$.
6. Execute action $a(\boldsymbol{s})$ at time t that leads the system to state $\boldsymbol{s}' = \boldsymbol{s}(t+1)$. Receive reinforcement $r(t+1)$.
7. Calculate truth values $\alpha_i(\boldsymbol{s}(t+1))$ for $i \in \{1,2,\dots,N\}$.
8. Determine the value of the new state:
$V_t(\boldsymbol{s}(t+1)) = \sum_{i=1}^{N} \alpha_i(\boldsymbol{s}(t+1))\ \max_k q(i,k)$.
9. Calculate the variation of the quality $Q(\boldsymbol{s},\boldsymbol{a})$:
$\Delta Q = r(t+1) + \gamma\ V_t(\boldsymbol{s}(t+1)) - Q(\boldsymbol{s},\boldsymbol{a})$.
10. Update the elementary quality $q(i,k)$ of each rule i and action k:
$\Delta q(i,k) = \kappa(i,k)\ \Delta Q\ e(i,k)$.
11. Save the elementary quality $q(i,k)$ in the Q-value LUT.
12. If convergence is obtained, stop learning.
13. $t = t+1$.

Figure A.3-4. Fuzzy Q-learning scheme

An Overall Performance Indicator, denoted by $\rho_{\mathrm{rel},i}$, aggregates the impact of various KPIs and represents the relative improvement of a considered optimization scheme i with respect to the reference scheme (REF). It is calculated as follows:

$$\rho_{\mathrm{rel},i} = \frac{w_{\mathrm{CDR}} \frac{N_{\mathrm{CDR}}^{\mathrm{REF}} - N_{\mathrm{CDR},i}}{N_{\mathrm{CDR}}^{\mathrm{REF}}} + w_{\mathrm{HFR}} \frac{N_{\mathrm{HFR}}^{\mathrm{REF}} - N_{\mathrm{HFR},i}}{N_{\mathrm{HFR}}^{\mathrm{REF}}} + w_{\mathrm{PHR}} \frac{N_{\mathrm{PHR}}^{\mathrm{REF}} - N_{\mathrm{PHR},i}}{N_{\mathrm{PHR}}^{\mathrm{REF}}} + w_{\mathrm{sat}} \frac{N_{\mathrm{sat},i} - N_{\mathrm{sat}}^{\mathrm{REF}}}{N_{\mathrm{sat}}^{\mathrm{REF}}}}{\sum_m w_m},$$

where $N_{\mathrm{CDR},i}$,$N_{\mathrm{HFR},i}$, $N_{\mathrm{PHR},i}$, and $N_{\mathrm{CDR}}^{\mathrm{REF}}$, $N_{\mathrm{HFR}}^{\mathrm{REF}}$, $N_{\mathrm{PHR}}^{\mathrm{REF}}$denote the overall numbers of connection drops, handover failures, ping-pong handovers of the considered and the reference scheme, respectively. In the denominator, $\sum_m w_m$ is used to normalize the weighted sum with respect to the sum of all employed weights w_m.

For a considered approach i and the reference scheme, $N_{\text{sat},i}$ and $N_{\text{sat}}^{\text{REF}}$, respectively, represent the number of time instances the users are considered as satisfied, i.e., the users are served with the corresponding amount of radio resources that meet the respective service requirements.

A.3.3 Long-Term Context-Aware Scheduling for UDN

System model

In this research work, we consider an indoor scenario comprising of multiple cells multicasting on-demand high definition video streaming. In such a scenario the uplink traffic is a very small comparing to downlink. To evaluate proposed scheme and for the sake of simplicity, we assume an air interface based on SISO-OFDMA (downlink only) with no power control and a flat power spectral density.

Due to small cell size of ultra-densely deployed access points, interference in such a system is a big issue which affects the overall performance in terms of throughput and delay. As mentioned, we address the problem of minimizing the number of dropped packets using a centralized scheduling. The best solution to the problem that achieves lowest average delay among all possible solutions is chosen.

The scheduling problem for the objective of packet drop ratio minimization subject to traffic demand and SINR constraints is formulated as follows:

$$\min_{x_{p,n}} \sum_{p=1}^{P_{curr}} step\left(d_p - \sum_{l=1}^{L} \sum_{n=n_{curr}}^{n_{d_p}} r_{p,n,l} \right) ; \quad \forall n_{curr} \tag{1}$$

Subject to

$$r_{p,n,l} \le BT_s \log_2 \left(1 + \frac{x_{p,n,l}\alpha_{p,l}\left|H_{p,p,n,l}\right|^2}{\sigma^2 + \sum_{p' \ne p} x_{p',n,l}\alpha_{p',l}\left|H_{p',p,n,l}\right|^2} \right) \tag{2}$$

variables

$$x_{p,n,l} \in \{0,1\} ; \quad \forall p \in 1,...,P_{curr} \ , \forall n, \forall l \tag{3}$$

where n_{curr} represents the current time unit index, P_{curr} is the number of packets available in the system at the current time unit, d_p is the available size of the packet p at n_{curr}, the largest deadline among all available packets is represented by n_{d_p}, $r_{p,n,l}$ is achievable throughput of packet p at time unit n and on resource block l and step function is used as a penalty function, i.e. If a packet cannot be served completely before its individual deadline, the penalty function's output is 1 for that packet. Equation (2) denotes the Shannon upper bound on achievable throughput for the link serving packet p, where α_p is the fixed downlink transmit power, *H* represents the channel coefficients of the link between any two arbitrary devices exchanging a packet. $x_{p,n,l}$ is the schedule decision variable, which has the value of 1 if the packet is decided to be served at time unit *n* and 0 if it is not scheduled. The solution matrix $x_{p,n,l}$ contains scheduler decisions over time and frequency resource. For this type of problems the solution is not unique which means that at every time unit optimizer outputs a number of solutions to the problem. We choose the solution where the average delay for the served packet is minimized. For this purpose for each packet the time difference between packet's

creation and serving the last segments are calculated and the solution which give the shortest time difference in average for all packets will be chosen as the best solution.

Performance evaluation

To conduct a system level study we setup a dense indoor simulation scenario with access points and users in rooms of size 5 m x 5 m. The mentioned scenario consists of 4 rooms with uniformly placed small cell base stations in the center of each room and each room contains 3 randomly dropped users. It is also assumed that a cluster of users (context users) are moving toward a coverage in the building (basement, elevator, etc.) and spend an arbitrary time there (cf. Table A.3-2). The scheduler is long-term CSI-aware if it perfectly knows future channel of context users, and non long-term-CSI aware otherwise (in this case CSI only at current time unit is known and replicated over prediction time window to be used by scheduler). This assumption enables us to investigate to which extend the quality of experience for context users and for the whole system improves if the scheduler is aware of CSI predictions. Other simulation assumptions are described in Table A.3-2.

Table A.3-2. Simulation parameters

Parameter	Default value
Carrier frequency	2600 MHz
Resource block (RB) size	1.8 MHz
Guard band	200 KHz
Path loss model	3GPP small cell path loss model, Indoor fast fading Jakes model
Wall Penetration loss, UE noise figure	3 dB
Total SC power	-3 dBm in downlink over 20MHz bandwidth.
UE speed	3 kmph
Antenna configuration	SISO
Receiver type	Interference Rejection Combining
Traffic	Poisson with 0.01 sec inter-arrival time
Packet size	100 / 500 KB
Packet deadline	60 ms
System load (%)	65,90,98
Overall simulation time	1 sec
Sliding prediction time window length	Largest deadline – current time <= 60 ms
RB set	[1:10] , [1:5] , [1 3 5 7 9]
Context information	Ideal knowledge of future CSI
Prediction time window length	1 sec
User mobility	Random movements
UE's under coverage holes	4
Time to spend in coverage holes	0.2 – 0.9 sec

A.3.4 Context-Based Device Grouping and Signalling

This technology component is intended for reducing the signalling overhead for MTC traffic and mitigating the potential congestion in the signalling channels. The idea of this concept is to exploit the correlated user behaviours from MTC devices and effectively remove the redundancy in the transmitted messages, which is typical during the congestion period for machine originating traffics. The correlated signalling messages are classified into two types:

- ***Fully redundant messages***, which are repeated messages sent by different devices. Transmission is unnecessary if the data has been transmitted by others, e.g., multiple handover requests sent by a group of tracking tags delivered on the same trunk.
- ***Diverse but correlated messages***, which may have correlation in two dimensions
  - Messages triggered by the same event but with distinct contents.

- Messages contain (partially) redundant information, e.g., a group of devices with consecutive IDs, or location information from a group of proximate UEs.

The rest of the transmitted messages are defined as ***diverse messages***.

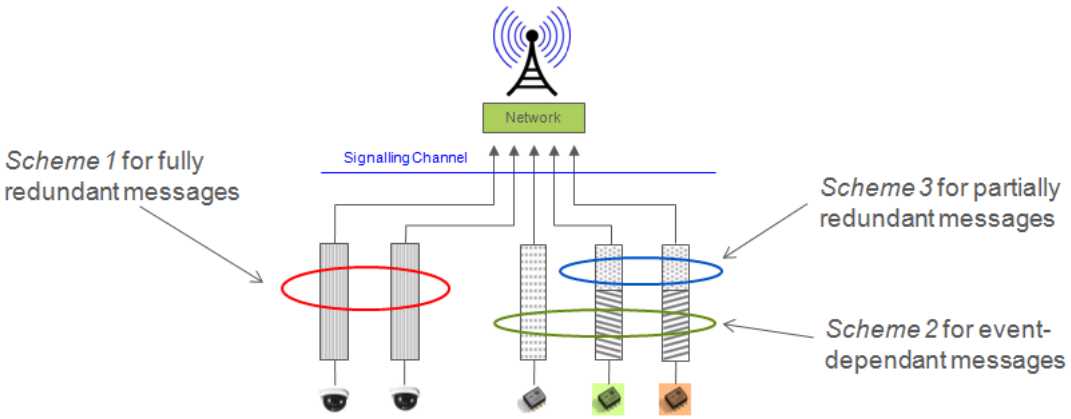


Figure A.3-5. A unified concept to remove the redundancy of correlated messages

The redundancy in the signalling messages is then removed by the investigated scheme, which provides a unified concept composed by three sub-schemes to target at different types of correlated messages, as shown in Figure A.3-5.

- Scheme 1: Inhibition of the fully redundant messages: The BS summarizes and analyzes the received messages. If the BS identifies a message that has high potential to be repeated by other MTC devices, it broadcasts/multi-casts the information to the rest of the MTC devices. Upon listening to the broadcast information, any receiving MTC device will be aware if it is necessary to transmit its own message or not.
- Scheme 2: Coordination and scheduling of RACH resources for the event-dependent messages: If a series of actions are invoked by an event, diverse messages from a subset of the associated MTC devices will be expected by the network. The BS then allocates a certain amount of RACH resources for the transmissions of these diverse messages. And the associated MTC devices are scheduled and send their messages consecutively following the priority order.
- Scheme 3: Cross-device compression for partially redundant messages: The compression is achieved by exploiting the partially redundant information between the associated devices. Firstly, a dictionary for the signalling messages is built up based on the previous messages from associated devices sent in the recent past, which is available both at the device side and at the base station. Thereafter, the new signalling message can be compressed into a shorter message based on the dictionary, which is then de-compressed and reconstructed at the receiver side reversely.

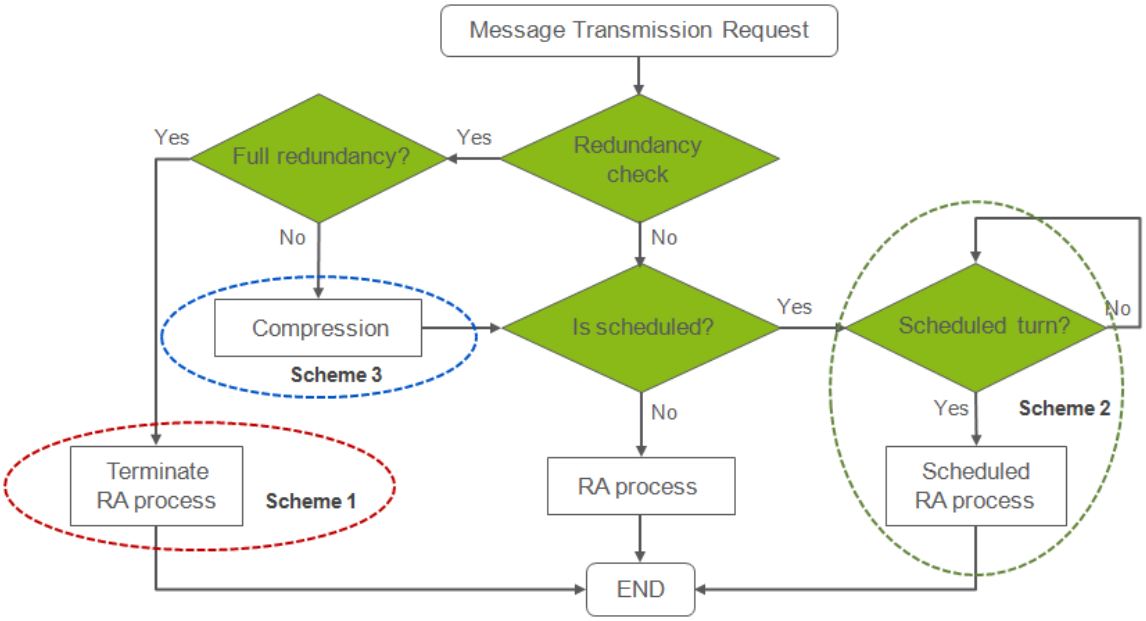


Figure A.3-6. Work flow of group-based signalling

The whole flow is illustrated in Figure A.3-6. Once a message transmission request is to be transmitted, the device first listens to the broadcast information from the BS and goes through a redundancy check. If the message it is about to send is a fully redundant message, the device will immediately terminate its attach request to the network and keep silent until triggered by the next event. If the message is a partially redundant message, then it can be compressed according to the common dictionary. Afterwards, the device checks if it has been scheduled by the BS. If the answer is positive, then the device merely needs to wait for its scheduled turn and use the assigned resources for its RA procedure. Otherwise it proceeds with the RA process as normal diverse messages.

A.4 D2D and V2X Mechanisms

A.4.1 Device Discovery in Emergency Communication

In the absence of network coverage, we propose the clustering concept to help building the network and support D2D connections. The key element is the CH node which will take over some of the functionalities of a cellular BS and help providing network coverage extension, synchronization and RRM for a group of attached devices. This will facilitate the smooth integration of the clusters in the infrastructure in case of partial network coverage. In the case of an emergency communication scenario as described in METIS test case 10 [MET13-D11], we define two types of UEs:

- **Type-1 UEs**: this category includes the remaining BS after a natural disaster, temporary emergency BS or the devices carried by first responders. We call these devices as CH capable nodes and we assign for each of them a pre-computed capability metric taking into account many aspects such as availability of network coverage, UE capabilities, maximum power allowed and remaining power level. This metric is defined using a weighing function in such a way that it gives the priority to BS to take the CH role. Only the nodes belonging to this category are initially allowed to transmit discovery beacons.

- **Type-2 UEs**: this type represents normal devices that can only connect to CHs.

Our clustering procedure consists of three phases: metric exchange, cluster formation (grouping) and slave-CH association.

1) Phase 1 - Metric exchange: During this phase, each type-1 UE broadcasts beacons containing its identifier and its predefined capability metric using a peer discovery resource (PDR) in each sub-frame. Alternatively, we could define a certain beacon transmission probability per sub-frame to reduce the overall beacon load in the system. Every device, when receiving and successfully decoding a beacon, stores the identifier of the sender and the corresponding metric. At the end of this phase, the type-1 UEs are able to build knowledge about their neighbors' metrics.
2) Phase 2 - Cluster head election and cluster formation: In this phase, each non-CH device selects the appropriate CH and associates with it. For this stage, we evaluate three clustering schemes.
    a. CH-driven (CH-based): Only type-1 nodes with highest metrics continue broadcasting their beacons. This approach has the advantage of reducing the number of devices competing for PDRs as compared to approach where every device is transmitting its beacons. Hence, it improves the SINR of received beacons and reduces the collision probability. However, as we show in the performance evaluation, using this approach some devices may stay out of coverage of a BS or a CH, thus in an emergency situation some victims might be in the proximity of a CH-capable UE (first responders), but still remain undiscovered.
    b. Hybrid: Every type-1 UE continues broadcasting beacons and so each type-2 UE that receives beacons from type-1 UEs in proximity selects the one with strongest metric and send a notification. Upon receiving such a notification signal, a type-1 becomes a CH. The advantage of such scheme is that it increases the probability of a device being covered by a CH. However, the drawback is the higher beacon load which increases the PDR collisions and thereby the discovery time and cluster formation in addition to higher energy consumption.

c. Threshold-based: this alternative takes into account the quality of the received beacon signals to estimate ifa type-1 UE is located at a edge of a cluster formed by a type-1 UE with higher metric. In fact, if the maximum signal strength received from any of the type-1 UEs with higher metrics falls below a predefined threshold(e.g., SNR = 20, 40, or 60 dB), a type-1 UE considers itself to be located at a cluster edge and continues sending out beacons. A type-2 UEs will connect to the type-1 UEs with highest metrics and SINR above the threshold. The advantage of this approach is that it prevents slaves with limited capabilities from remaining isolated and provides better coverage by ensuring high-quality links between each CH and its slaves. It also guarantees a trade-off between coverage and energy efficiency.

In our concept we have also defined mechanisms that trigger the cluster reconfiguration. For instance, in the CH-driven approach, if a slave type-1 UE does not hear its CH for a time period, it triggers a reconfiguration by sending its beacons and if no notification is sent back by a neighbouring CH it considers that it is out of coverage and continues transmitting its beacons. While in the threshold-based approach, if the received signal from the CH falls below the threshold during a period of time, the type-1 UEs that were initially slaves will trigger a reconfiguring procedure and start sending beacons again.

The concept was evaluated in an emergency communication scenario as in METIS Test Case 10 using a static system level simulator and Monte Carlo simulation. The simulation where performed for different number of UEs(20-500) in an area equivalent to cell area in a cellular system with ISD of 500 m, in which equal number of type-1 UEs and type-2 UEs are uniformly dropped. The propagation models are based on [MET13-D61], where the propagation model of cellular UEs is the O2I model (PS#2) and the propagation model of D2D UEs is the O2I model (PS#10). We considered one building per sector and assuming that 30% of the users are indoor. No fast fading effect was considered. 20 PDR per sub-frame were used for the discovery and we assumed a user beacon transmission probability per sub-frame equal to 50%. The discovery evaluation was done using a simplified link-to-system interface, according to which a discovery beacon will be detected if SINR is larger than 0dB. We simulated 10 drops of user and used 25s time length for each drop. For each device the right CH is pre-known based on the path-loss, SINR threshold and UE metric and they are used at the end of simulation to indicate if a user is under coverage or not (a UE is considered under coverage if it has connected to the right CH and is considered out of coverage if at the end of simulation it is already connected to the right CH).

Two KPIs were used to evaluate the concept: the discovery ratio representing the percentage of the users that connect to at least one CH, the discovery time and the energy consumption

The simulation results show that both hybrid and threshold-based approaches achieve 99.9% discovery ratio when a moderate number of CH-capable devices are in the network. While the CH-based scheme requires a higher number of devices to converge to higher discovery ratio and the main reason for that is that only few devices with the highest metrics are allowed to act as CH and provide coverage for the rest of the users. Finally, it can also be observed that the threshold-based achieve trade-offs between number of CHs and coverage ratio.

When looking to the time needed by the slaves to connect to a CH, the figure shows that in the hybrid approach all CHs capable devices transmit their beacons which leads to a higher collision rate and hence a higher discovery time which increases with the number of the devices in the network. The CH-based approach achieves lower discovery time at the expenses of a lower discovery ratio. The threshold-based uses the measurements about the beacon signal quality to optimize the selection of the CHs and thus handles the trade-offs between number of CHs and coverage ratio.

A.4.2 Distributed Channel State Information (CSI) Based Mode Selection for D2D Communications

Distributed CSI based mode selection for D2D communications has been described and analysed in [FSS14] and [MET14-D42]. In the Annex we discuss how the single hope mode selection algorithm can be extended to multi-hop D2D communication scenarios and present numerical results obtained by system level simulations.

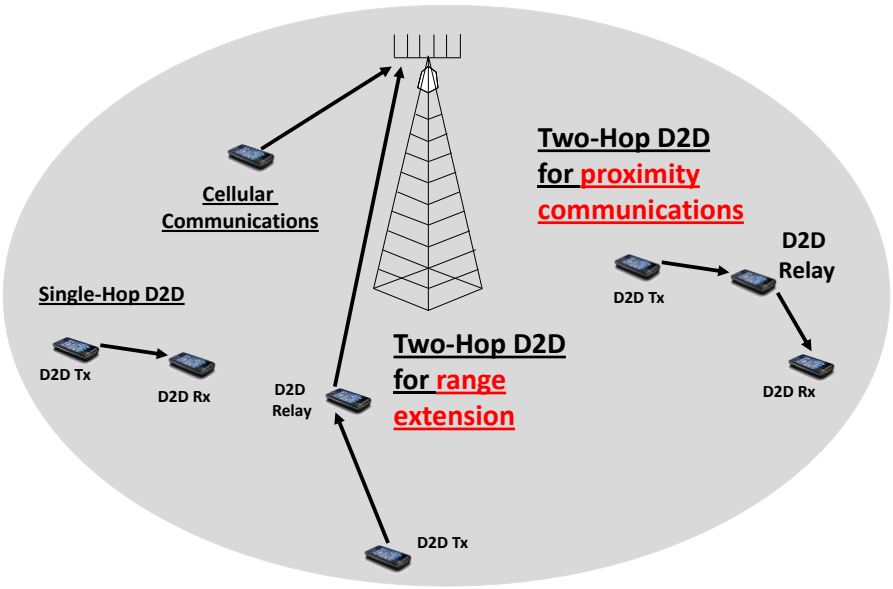


Figure A.4-1. An example of a cellular network supporting single- and multi-hop D2D communications in cellular spectrum. Two-hop D2D communication scenarios include range extension and proximity communications. In range extension, a UE at the cell edge or in outage is helped by a relaying UE. In proximity communications two devices communicate via a relaying UE

As shown in Figure A.4-1, D2D communications can help an out of coverage or cell edge UE to boost its link budget and thereby allow the UE to communicate with a cellular BS. Also, D2D communications can use two hops between a D2D transmitter and D2D receiver pair in the so called proximity communications scenario. The distributed CSI based mode selection algorithm is applicable in these two-hop scenarios to select between traditional cellular communications, single hop D2D communications and two-hop D2D communications mode. Specifically, the CSI based mode selection scheme discussed in [MET14-D42] is applied to the two-hop range extension and two-hop proximity communication scenarios of Figure A.4-1 as follows.

Algorithm 1 Harmonic Mode Selection (HMS) for Proximity Communication

1: **if** $G_{eq} \geq \max\{G_{TxRx}, G_{TxBS}\}$ **then**
2: Choose D2D two-hop communications
3: **else if** $G_{TxRx} \geq G_{TxBS}$ **then**
4: Choose D2D single-hop communications
5: **else**
6: Choose cellular mode, that is D2D Tx and Rx communication through the BS.
7: **end if**

Figure A.4-2. CSI based mode selection applied to the two-hop D2D Proximity Communication scenario of Figure A.4-1. G_{TxRx} and G_{TxBS} refer to the path gain between the D2D Tx and D2D Rx nodes and the D2D Tx and BS respectively. G_{eq} refers to the equivalent path gain defined as the harmonic mean of G_{TxRx} and G_{TxBS}.

Algorithm 2 Harmonic Mode Selection (HMS) for Range Extension

1: **if** $G_{eq} \geq G_{TxBS}$ **then**
2: Choose D2D relay assisted communication
3: **else**
4: Choose cellular mode that is D2D Tx transmits directly to the BS.
5: **end if**

Figure A.4-3. CSI based mode selection applied to the two-hop D2D Range Extension scenario of Figure A.4-1. G_{TxBS} refer to the path gain the D2D Tx and BS. G_{eq} refers to the equivalent path gain defined as the harmonic mean of G_{TxRe} and G_{ReBS}, G_{TxRe} denotes the path gain between the D2D Tx and the D2D relay (helper) device, while G_{ReBS} denotes the path gain between the relay (helper) device and the BS

The distributed CSI based mode selection scheme is naturally extended to the two hop D2D scenarios of Figure A.4-1 (i.e. proximity communications and range extension) as described in Figure A.4-2 and Figure A.4-3. Below we consider the performance of the proposed mode selection scheme in a 7 cell system, whose parameters are specified in Section A.4.9.

Figure A.4-4 is the scatter plot showing the energy consumption and throughput performance of the mode selection algorithm (Algorithm 2) in the range extension scenario. The utility maximizing power control reaches the highest average throughput, although it does not show the best SINR for some users, with a gain of approximately 29 % over the fixed power control scheme. This behaviour explains how utility maximizing PC achieves the highest throughput. Further details of the operation of the mode selection scheme in two hop D2D scenarios are available in [SFM14].

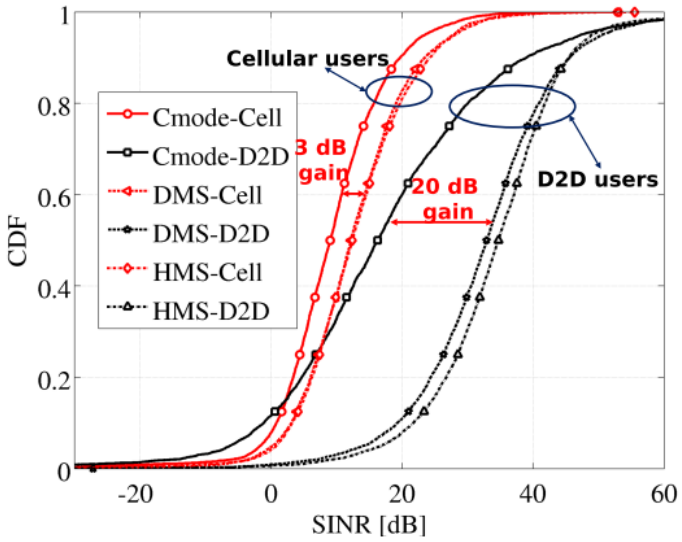


Figure A.4-4. Performance of the distributed CSI based mode selection algorithm in the two-hop D2D Proximity communication scenario: CDF of the SINR for both cellular UEs and D2D candidates with forced cellular mode (no D2D communication, Cmode) Direct (single hop) D2D mode (DMS) and mode selection according to the CSI based mode selection for two-hop D2D (harmonic mode selection, HMS)

As Figure A.4-4 clearly indicates, the distributed CSI based mode selection (HMS) is superior for both the cellular UEs (denoted '-Cell') and the D2D candidates and considering all the modes. The cellular UEs benefit somewhat (3 dB) from D2D communications. For the D2D candidates, the mode selection gain is much more pronounced (22 dB) with the proposed CSI based mode selection scheme (HMS) as described in Figure A.4-2. More results, including the range extension scenario are available in [SFM14].

A.4.3 Location-based mode selection for D2D communication

Location-based mode selection method for D2D enables seamless transition between indirect and direct transmission modes for cellular users. In this approach the mode selection procedure is centralized and is transparent to the users. The criterion for mode selection decisions is the transmit power of the UE, which implies that mainly UE's gains are taken into account. However, selection of the direct mode for some UEs also brings benefits to the network side, e.g., traffic offload. The proposed mode selection algorithm is based on distance estimation and its main advantage is that it does not require any CSI to be exchanged in order to determine appropriate transmission mode.

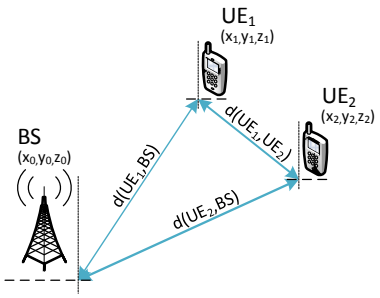


Figure A.4-5. Distance requirements for transmission mode selection

The user location information is used to estimate distances between considered UEs and the base station (Figure A.4-5). Since the distance has the biggest impact on path-loss between nodes, it is justified to assume that the transmit power is directly related to distance. Based on

location and distance knowledge a decision is made on transmission mode selection for *i*-th and *j*-th UE with the following criteria in mind:

$$\begin{aligned} &\mathrm{d}(\mathrm{UE_i}, \mathrm{UE_j}) < \mathrm{d}_{\mathrm{D2D_{max}}} \\ &d(UE_i, UE_j) < d(BS, UE_i) \\ &d(UE_i, UE_j) < d(BS, UE_j) \end{aligned} \quad (1)$$

Where $\mathrm{d}(.)$ is an Euclidean distance between two nodes and $\mathrm{d}_{\mathrm{D2D_{max}}}$ is an arbitrary distancelimit that can be adjusted depending on the scenario and it is limited by the maximum transmit power of the UE.

Evaluation

The proposed location-based approach to D2D mode selection was evaluated in a simplified version of Test Case 2 environment according to the description in [MET13-D61]. Only macro BSs and outdoor users were considered.

In the evaluation scenario, 400 users were uniformly placed outside the buildings either on pavements (150 CUEs) or in cars (170 CUEs). Among these users 40 D2D pairs were deployed (80 DUEs which constitutes around 20% of overall UE number) with the distance between each DUE from D2D pair chosen from a uniform distribution $\mathrm{U}(0\mathrm{m}, 100\mathrm{m})$. Additionally, 70% of D2D pairs were pedestrian users, whereas 30% were in-car users. The mobility of the users was modeled according to the description in [MET13-D61].

The most relevant parameters of the simulations are presented in Table A.4-1.

Table A.4-1. Simulations parameters

Parameter	Value
Carrier frequency	800 MHz
Bandwidth	20 MHz
Max. BS TX power	46 dBm / 10MHz
Cellular scheduler type	Round Robin
Max. users per TTI	10
TTI duration	1 ms
Max. UE TX power	24 dBm
Power control mechanism	OLPC: $P_{TX} = \min(-75 + 0.8 \cdot PL(d), P_{TX}^{\max})$
Max. D2D distance	100 m

In the simulations two resource allocation schemes were evaluated. These are the proposed location-based uplink resource allocation and a random resource allocation, where CUEs for resource sharing are selected randomly. For both resource allocation methods the proposed location-based mode selection procedure was used. Additionally, the evaluation included the comparison with a system where D2D communication was not allowed. The system level simulations were carried out for 1000 drops consisting of 30 TTIs repeated 5 times with different random generator seeds.

In the context of mode selection the following aspects were analyzed:

- Level of interference experienced by CUEs introduced by the DUEs (Figure A.4-6a),
- Transmit power level of UEs (Figure A.4-6b),

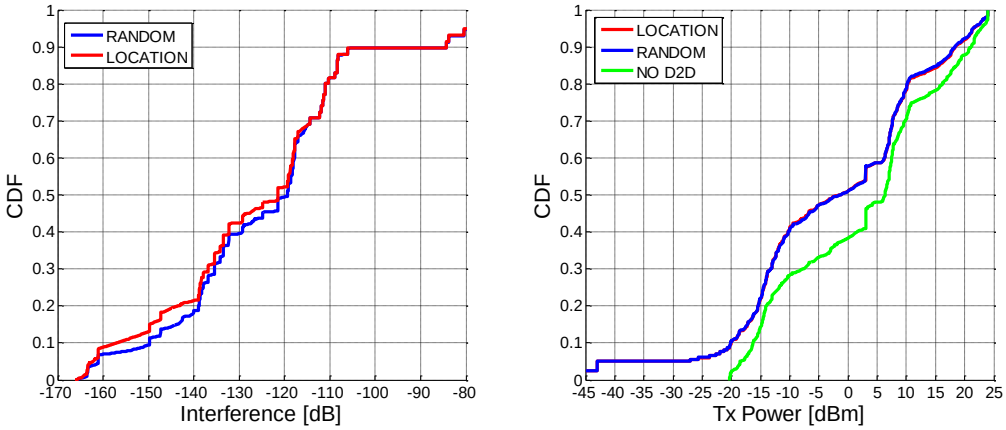


Figure A.4-6. D2D introduced interference CDF (a) UE transmit power CDF (b)

Analyzing the CDF of interference introduced by DUEs it can be noticed that for high level of interference both investigated cases give similar results. For low interference region the location-based approach achieves better results. However, these results show that the introduction of more complex method of resource allocation does not necessarily lead to a reduced interference from DUEs. Looking at the UE transmit power CDF one can see that with a location-based mode selection a significant reduction of transmit power can be achieved by using direct transmission mode regardless of the resource allocation method used for D2D communication.

A.4.4 Multi-cell Coordinated and Flexible Mode Selection and Resource Allocation for D2D

We focus on an ultra-dense multi-cell indoor scenario, where D2D links can extend over multiple cells because of short cell radii and differentiate between the following modes of communication.

- Direct only D2D: all D2D traffic is only served through direct links between devices. Reuse of frequency resources is allowed between the cellular and D2D users.
- Indirect only D2D: all D2D traffic is routed through the network. A D2D communication involves two hops, i.e. UL transmission and a subsequent DL transmission.
- Mode selection: D2D traffic can be routed either directly or through the network according to different criteria e.g. link length, pathloss, intra-cell or inter-cell communication, instantaneous SINR conditions, network load, etc.

In time slot *t* the scheduler goes through all possible transmitter/receiver combinations of D2D devices (respecting a half duplex constraint) and calculates a sum delay-weighted throughput using the SINR conditions in the system based on scheduling decisions made in time slot *t-1.* For each frequency block of resources, the following metric is calculated:

$$\sum_{i \in C_k} w_i x_i R_i + \sum_{i \in D_k} w_i y_i R_i \quad (1)$$

Where *wi* is the maximum packet delay in link *i*, *Ri* is the achievable data rate in link *i*, *xi* and *yi* are boolean variables that allow the muting option of resources, *Ck* is the set of cellular users and *Dk* is the set of D2D UEs, which can reuse resources with the cellular UEs.

A brute force approach is used for each frequency resource block, assigning it to every possible UL, DL or D2D link (with the option of muting that resource block and reusing it between cellular and D2D users) and the link combination that gives the highest metric computation as defined in (1) will be chosen by the scheduler. This implies deciding which links to schedule (mode selection) and the resource allocation for each link. Finally, simple

LTE-like power control with fractional pathloss compensation is applied toward the receiver D2D UE for direct links and the same power control mechanism, but towards the serving access point, is applied for indirect links We consider a decentralized or stand-alone variant, with the metric in (1) being computed independently by every cell, whereas a centralized variant benefits from global knowledge and makes coordinated decisions. The centralized scheduler uses the following metric:

$$\sum_{k \in G} \left(\sum_{i \in C_k} w_i x_i R_i + \sum_{i \in D_k} w_i y_i R_i \right) \quad (2)$$

with *G* being the cluster or coordination group of cells. In the centralized variant, the amount of overhead, computational complexity and signalling needed to achieve practical gains should be considered, especially with all the D2D links CSI required in addition to the conventional cellular links. The time scale or granularity in which mode selection should operate is an interesting and open research question.

The following considerations and assumptions are made for the simulation results: flexible (dynamic) UL/DL TDD; half duplex; multi-link D2D, by which a D2D device could transmit to multiple D2D devices in a scheduling slot by utilizing parallel frequency resource blocks; slow power control applied on D2D links based on path loss between the devices; 25 indoor cells of 10x10 m with 3dB wall attenuation and 3GPP small cell path loss model and indoor fast fading; 8 users per cell with 30% of them being D2D and multi-cell D2D across small cells allowed up to 8/16 m range; IRC receiver and 4x4 MIMO assumed in all links; traffic with Poisson distribution, 1 second inter-arrival time and fixed file sizes of 960 kB/ 240 kB/ 480 kB for DL/UL/D2D, respectively.

The performance indicator of interest in our studies is the 99th percentile of the packet delay distribution, focusing on optimizing the performance of applications for which all packets should be served under certain delay constraints. Clearly, this generally implies a performance trade-off for lower percentiles, which is especially visible in centralized results for which the improved performance of very degraded users comes at the expense of some degradation in average user experience.

Joint mode selection and resource allocation obtains a 15% reduction in packet delay over direct D2D communication only at 99th percentile for D2Dtraffic (Figure A.4-7) and 14% gain for overall (cellular + D2D) traffic (Figure A.4-8). The gains come from smartly choosing which mode of communication to use. With moderate D2D ranges it is a good idea to favour direct D2D communication as it implies lower delays, especially if these direct connections are coordinated as in the centralized variant. This can be confirmed by observing that only direct (black curves in both figures) performs considerably better than only indirect (green curves in both figures).

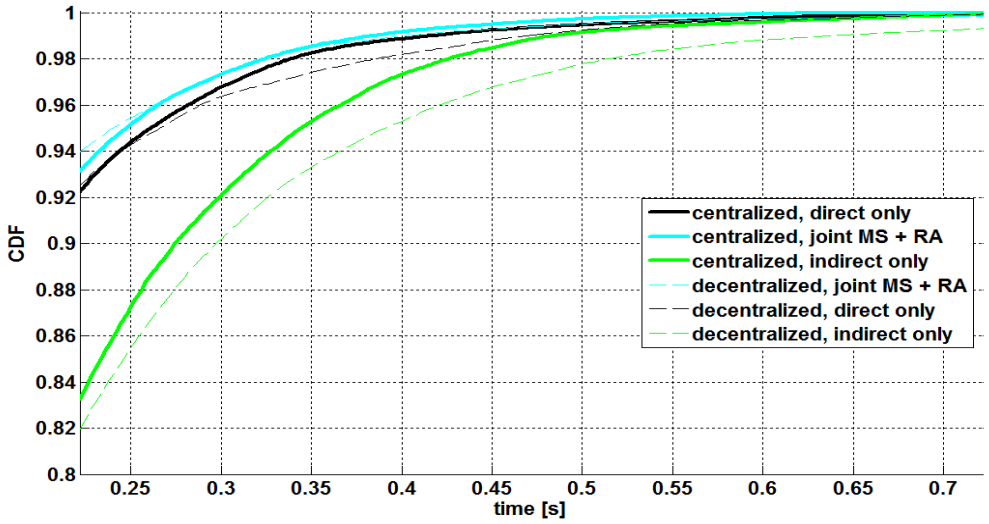


Figure A.4-7. Comparison of decentralized and centralized D2D packet delay distribution for different communication schemes D2D links up to 8 m

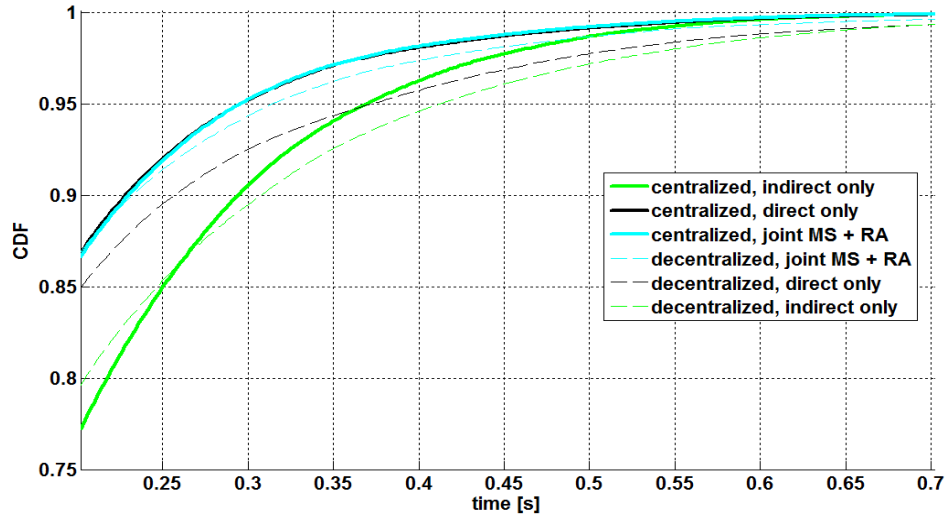


Figure A.4-8. Comparison of decentralized and centralized overall (cellular + D2D packet delay distribution for different communication schemes. D2D links up to 8 m

In case of using a centralized scheduler (solid lines in Figure A.4-7 and Figure A.4-8), joint MS + RA obtains 10% packet delay reduction over direct only for D2D traffic and no visible gain when considering both the cellular and D2D traffic. Comparing centralized versus decentralized performance (solid versus dashed lines) for the different communication variants, it is observed that overall knowledge and coordination provides gains in terms of packet delay reduction over decentralized scheduling decisions. For the overall traffic in the system, 11% packet delay reduction for joint MS + RA, 24% for direct only and 19% for indirect only are achieved (Figure A.4-8).

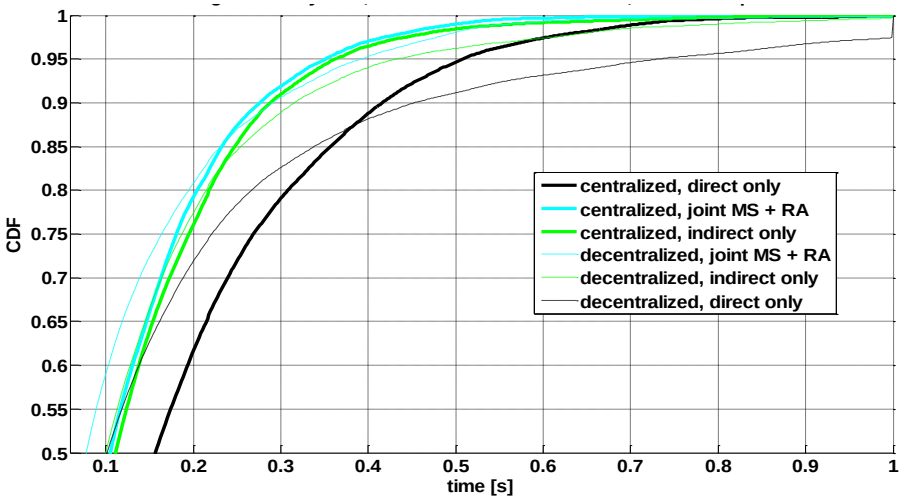


Figure A.4-9. Comparison of decentralized and centralized D2D packet delay distributions for different communication schemes. D2D links up to 16 m

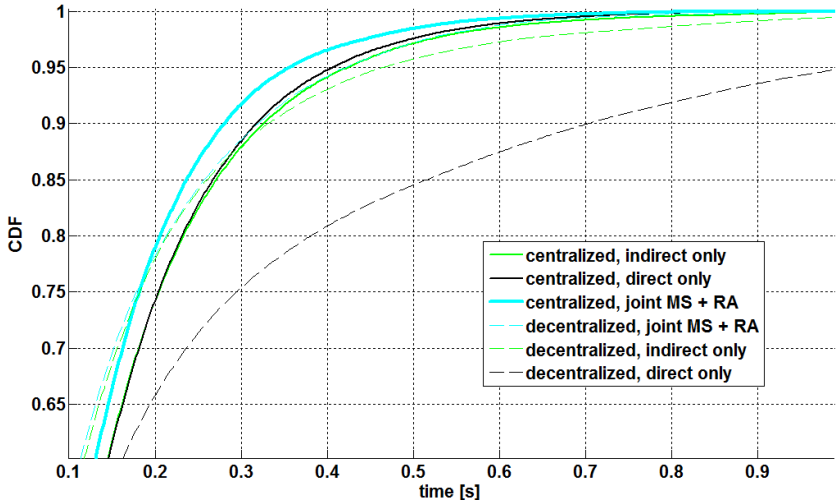


Figure A. 4-10. Comparison of decentralized and centralized overall (cellular + D2D) packet delay distributions for different communication schemes. D2D links up to 16 m

The significant amount of multi-cell D2D long links in the case that direct links up to 16 m are allowed, direct only performs badly if done in a decentralized or uncoordinated manner. There is huge improvement through centralization for direct only. Joint MS + RA is able to achieve some gains over indirect only by scheduling direct D2D communication when the conditions are appropriate, especially for shorter links. Gains around 30% packet delay reduction are achieved by joint MS + RA over indirect only both for D2D and overall traffic for the decentralized case and around 12% for the centralized case, respectively.

A.4.5 Location-Based Resource Allocation for D2D Communication

In the context of resource allocation for D2D communication two main approaches can be distinguished. First approach is an orthogonal resource allocation where D2D communication utilizes dedicated resources. Second approach is to allow for resource sharing between D2D DUEs and cellular users CUEs. In this case DUEs are allowed to use resources that are already used by CUEs. The second approach is more demanding in terms of interference management and D2D communication coordination, however it promises more benefits to the network performance by harvesting the so called reuse gain inherent to D2D communication.

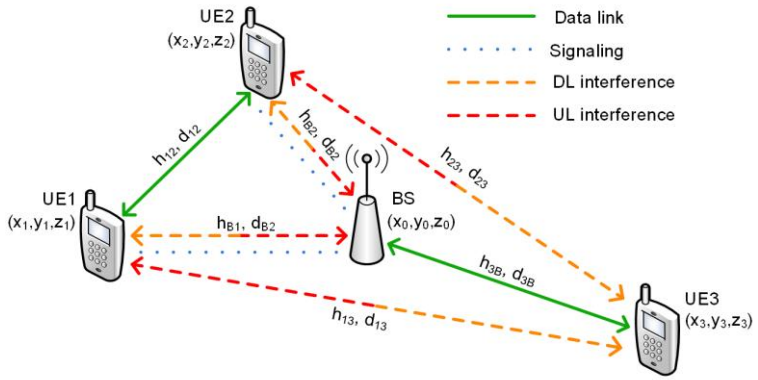


Figure A.4-11 Interference pattern in D2D enabled cellular network

Location-based resource allocation scheme where D2D communication shares resources with CUEs is proposed. The aim of the method is to enable D2D communication in the network while minimizing the interference related with the introduction of D2D communication. Figure A.4-11 shows the typical interference pattern in D2D enabled cellular network. In the location-based method users positions are used to estimate the distances between nodes in the network. The selection of the best candidate for resource sharing is performed based on the distance maximization, which with the assumption that the path-loss has the biggest impact on interference is a mean to minimize the interference. More detail on sharing candidate selection is given in Figure A.4-12 for either downlink or uplink resource sharing.

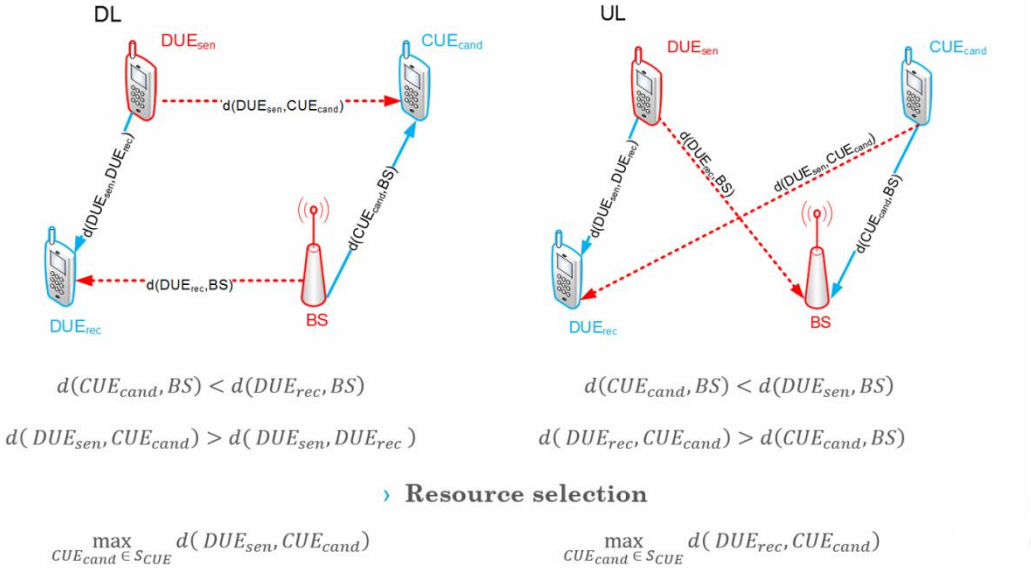


Figure A.4-12. Resource sharing candidate selection criteria

The advantage of this simple resource allocation method is that it requires no CSI from the D2D users due to the fact that location information is used. Moreover, centralized approach gives more control of D2D links and as a result allows for a better control of interference related to D2D communication.
The proposed location-based approach to resource allocation for D2D communication was evaluated in the set-up described in Section A.4.3.

The evaluation investigated the following aspects:

- SINR (Figure A.4-13)
- Spectral efficiency (Figure A.4-14)

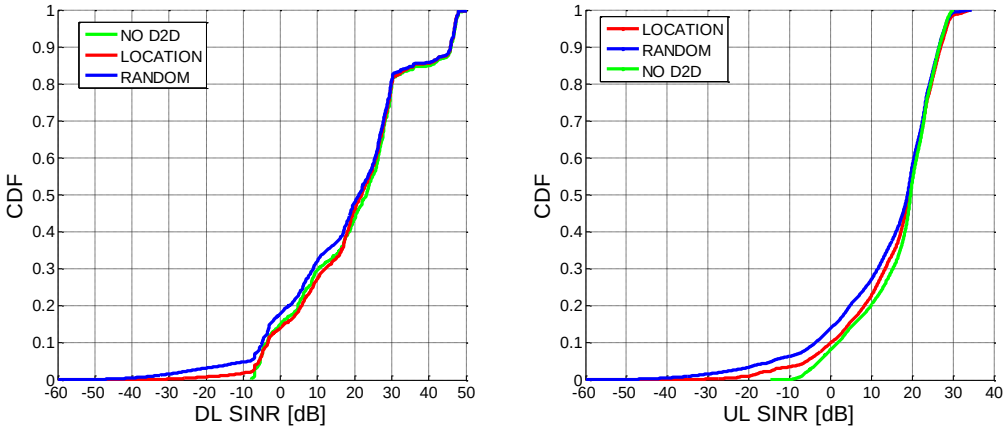


Figure A.4-13. DL SINR CDF (a) UL SINR CDF (b)

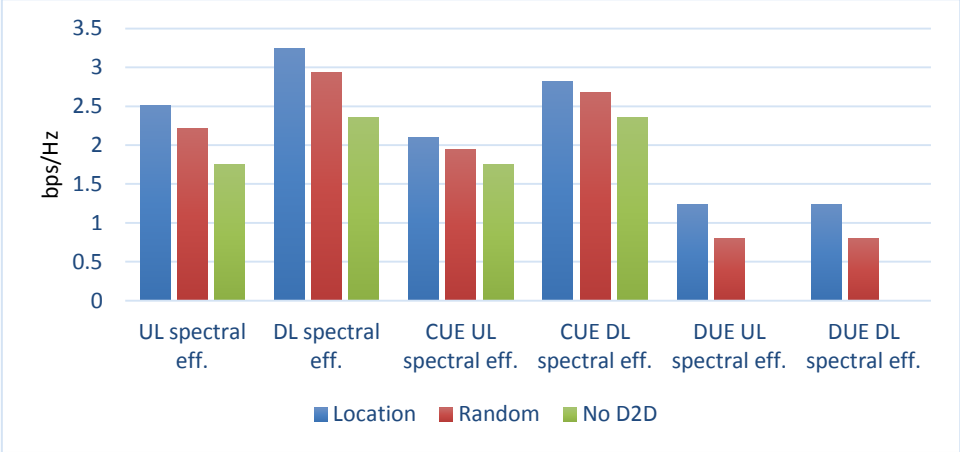


Figure A.4-14. Spectral efficiency comparison.

Figure A.4-13 contains the CDF curves of DL and UL SINR in the network. A reduced SINR for both directions of transmission can be observed when D2D communication is present in the system. Nevertheless, the location-based resource allocation outperforms the random allocation approach. Figure A.4-14 show the spectral efficiency achieved by the network. An increase in spectral efficiency can be seen for both investigated D2D resource allocation methods, as compared to the no D2D case. Comparing the two D2D resource allocation methods' performance one can see that location-based approach achieves better spectral efficiency for both overall spectral efficiency as well as for the spectral efficiency of DUEs.

A.4.6 Context-Aware Resource Allocation Scheme for Enabling D2D in Moving Networks

In regards to the context-aware resource allocation scheme for enabling D2D in moving networks, it is assumed that the D2D underlay is utilized only by a certain class of UE, i.e., VUE running Vehicle-to-Device (V2D) applications. Many future V2D services (in particular, safety services) have very strict QoS and reliability requirements. For example, a beacon marking a road hazard (e.g., accident site) needs to be periodically (typically every 10 ms) and reliably transmitted after such an event is detected. Therefore, the aim of the proposed heuristic LDRAS is to guarantee continuous (or periodic) transmission opportunities for such services in the D2D underlay, while reducing the control overhead and the interference within the primary network. The proposed scheme reuses the available UL and DL resources for data transmissions in the underlying D2D network. In order to fulfill the service characteristics

described above, LDRAS relies on persistent resource allocation. Each cell sector is divided into a number of spatially disjoint zones, where in each zone a set of RBs is appointed for D2D communication. The same sets of RBs are then also reused within the primary network. However, in order to guarantee a certain maximum interference level caused by CUEs in the D2D underlay, the reuse of a certain RB set is not allowed in specific zones in the neighborhood of the one where the same RBs are used for D2D transmissions. An example zone layout and a corresponding example RB reservation assignment are illustrated in Figure A.4-15. The zone layout and RB sets are then fixed and do not change over time. We assume that this data can be stored in the memory of each VUE, which uses its built-in positioning system in order to track its location. Upon entering a new zone, a VUE signals the zone index to the eNB which then assigns a subset of RBs from the respective set to the VUE. In the primary network the resources are assigned to CUEs according to the network operator's scheduling policy, with the additional constraint that only RBs which are not prohibited for reuse in the respective zone for CUEs can be allocated. In this manner, the need for full CSI knowledge at the eNB is eliminated and thus the necessity for extensive channel measurements and signaling overhead.
Further details on the basic concept of the proposed scheme are available in [BKK+14], while [tbd] outlines how it can be applied to multicell deployments.

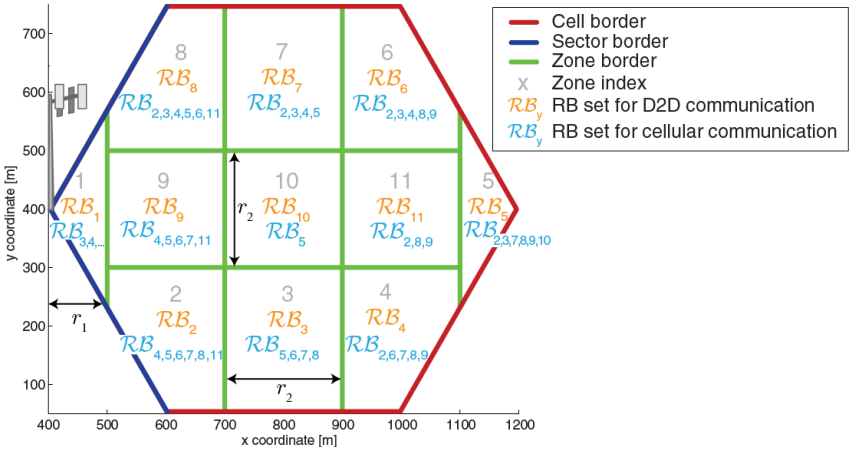


Figure A.4-15. Example zone topology with resource reservation for a single cell sector

In addition to the packed delay evaluation presented in Section 5.7, Figure A.4-16 shows the achieved throughput with LDRAS and a SOTA reference scheme. Here, the UL and DL resource allocation algorithms due to Zulhasnine et al. [ZHS10] have been chosen for their comparable heuristic nature. A mean throughput of approximately 8.5 Kbps is achieved with LDRAS, corresponding to the traffic demands in the simulated scenario (i.e., packet size of 1600 bytes with inter-arrival time of 1.5 s). In contrast, the reference scheme shows poor reliability (or equivalently, availability) of the D2D links. Due to the opportunistic resource allocation strategy and the prioritization of cellular links, around 40% of the potential vehicular transmissions are dropped as no resources are allocated on time.

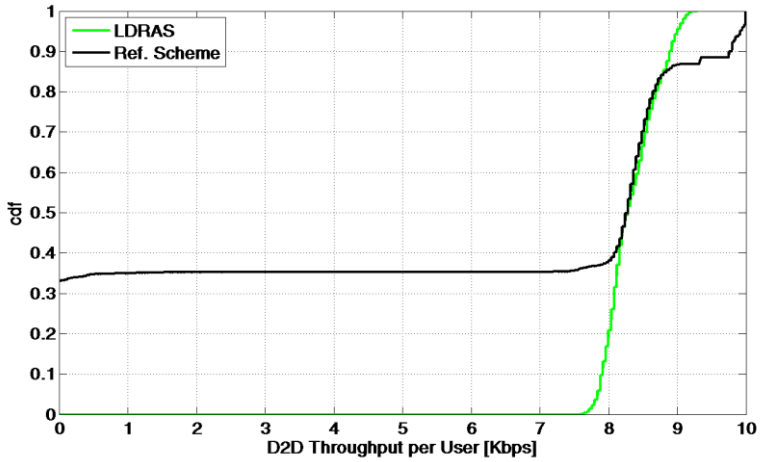


Figure A.4-16. Cumulative distribution function of the throughput per VUE in the D2D underlay

Another important aspect of the performance of RRM schemes for D2D communications is the impact on the primary network. The CDFs of the corresponding cellular throughput are shown in Figure A.4-17. Here the performance achieved with LDRAS and the reference scheme is additionally compared to a legacy LTE system in order to assess the impact of resource reuse. It can be seen that the reuse of DL resources has virtually no impact on the cellular links as LDRAS as well as the reference scheme show almost identical performance with the legacy LTE system. In the UL, on the other hand, both LDRAS and the reference scheme display some degradation. Hereby, due to the prioritization of cellular links, the performance of the reference scheme is closer to the legacy LTE system where there is no interfering D2D underlay. This, however, comes at the price of poorer reliability of the D2D links as discussed previously. In the case of LDRAS, the mean UL throughput is decreased by 15% as compared to the LTE baseline and by 11% as compared to the reference scheme. The reason for this is two-fold. On the one hand, the exclusive utilization of a portion (10% in the simulated example) of the UL resources for D2D communication in close proximity to the eNB leaves fewer resources for UL transmissions. On the other, the resource reuse and the prioritization of D2D links in order to meet the reliability requirements of V2D applications leads to cross-interference which degrades the UL performance.

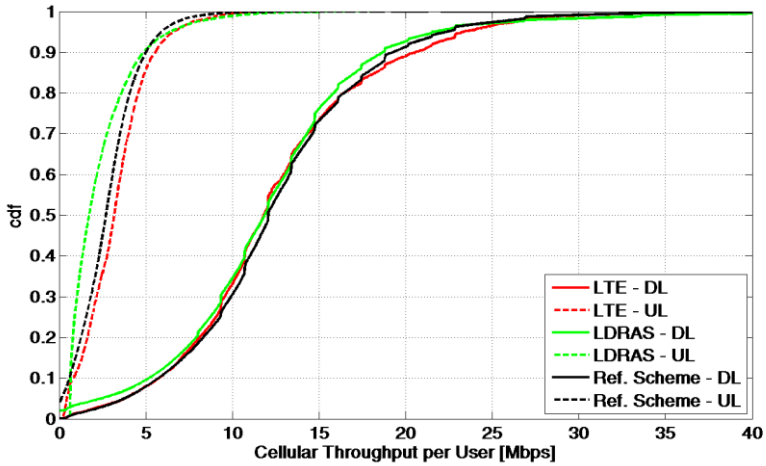


Figure A.4-17. Cumulative distribution function of the throughput per CUE in the primary network

A.4.7 Network Assisted Resource Allocation for Direct V2V Communication

In order to increase the reliability of the safety message transmission, the proposed concept relies on network assistance to assign separate resources to different clusters located in the vicinity of each other. Such resources can either be time and/or frequency resource; however for the sake of simplicity we only consider time resources for the rest of this paper. Additionally, the network control enables efficient spatial reuse of resources when the density of nodes exceeds the available resources based on location information collected from different CHs. However, to maintain a reasonable signaling overhead even for a large number of vehicles, the assignment of resources is organized in two levels: one network level where a set of resources is assigned by the network node (base station) to each cluster head and one cluster level where each CH uses its respective set of resources to allocate time slots to the individual vehicles within the cluster. Specifically, the BS utilize the context information, such as the size of the cluster, geographical location and the movement of the cluster, regularly provided by the CH to adjust the network level resource assignment accordingly. Finally, to allow the spatial reuse of the resources among clusters, we used the greedy coloring algorithm that tries to maximize the minimum distance between the clusters using the same resources.

The proposed solution was implemented using a system level simulator in order to verify the impact of the network-assistance on the reliability of the safety message transmissions and to gain an insight into the complexity of clustering and signalling procedures. For the evaluation, we consider an 8 km highway scenario with 4 lanes in each direction and covered by a cellular system with a cell radius of 1km. The vehicles arrive at each end of the highway according to a Poisson distribution with a speed modelled as a Gaussian random variable. Different vehicular densities and speed scenarios are simulated by varying the mean arrival-time of the vehicles, the mean value and the standard deviation of the Gaussian distribution. The path loss between vehicles is modelled using the Nakagami-model [CSB+07] which is based on channel sounding performed at 5.9 GHz in an outdoor environment. We assume a bandwidth of 10 MHz dedicated for safety message broadcasting. The performances of the proposed solution are measured for different traffic loads, different vehicular speeds and densities. CAM transmission frequency, range (defined by the transmit power), and size.

The reliability of the proposed solution is measured based on the message delivery ratio which reflects the percentage of vehicles that successfully receive Cooperative Awareness Message (CAM) among all the vehicles present within the transmission range of the transmitting vehicle. To investigate the effect of these parameters on the CAM delivery rate we consider four different CAM configurations. Figure A.4-18 below depicts the results obtained for a density of 11 vehicles per lane per km and an average speed and a standard deviation of 110 km/h and 33 km/h, respectively.

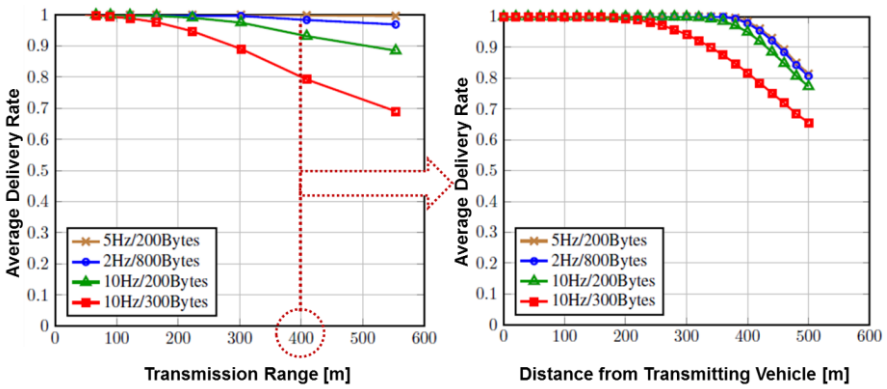


Figure A.4-18. Distribution of the CAM delivery rate

As shown in the left graph, when the transmission range increases, the average delivery rate

decreases due to higher interference received from vehicles using the same resources even at higher distances which results in higher collision rate. This effect becomes clearer for high CAM load where the number of available time slots is lower than the required resources and hence increasing the number of vehicles sharing the same resources. However, it can be observed that when considering low CAM frequency and small size (for example 5 Hz and 200 bytes) the network assisted is able to ensure high delivery ratio even for high transmission range, increasing thereby the safety awareness region.
Furthermore, when analysing the distribution of the delivery ratio over the distance for a fixed transmission range of 400 meters, it can be observed that, even at high traffic load, the proposed solution guarantees high reliability within the area close to the transmitter, referred to as awareness area, i.e. 100 percent of the vehicles within a distance of 300 meters around the transmitter successfully receive the CAM messages (see the right graph of Figure A.4-18). This is mainly due to the use of context information to maximize the minimum distance between the vehicles using the same resources. When comparing these results to those obtained for IEEE 802.11p and STDMA in [BUS+09], the analysis shows that the network assistance increases the CAM delivery rate by 20% and 16% compared to IEEE 802.11p and STDMA, respectively for medium traffic load (see Table A.4-2). Although the improvement is lower for high CAM load, we believe that one of the advantages of the network assistance is to use information collected about the vehicle topology and density in order to adapt the transmission parameters and efficiently reallocate the resources.

Table A.4-2. Average delivery rates in the network-assisted solution and the exiting ad hoc schemes

CAM Configuration	Network-Assisted V2V	802.11p	STDMA
2Hz/800Byte	0.98	0.78	0.82
10Hz/300Bye	0.8	0.73	0.77

One additional metric used to measure the efficiency of the architecture is the overhead generated by the control signalling both for the cluster maintenance and the radio resource management. Depending on the spectrum used for the exchange of the control messages, we divided the total overhead into a cellular overhead which includes all the communications between the CHs and the NNs, and an ad hoc overhead created by the control messages exchanged between the CHs and their CMs, namely the advertisement message and the cluster joining requests. Figure A.4-19 shows an estimation of the overhead generated within a cell for different vehicle densities assuming a worst case scenario where the network node broadcasts the cluster topology information every super-frame (i.e. 200 ms).

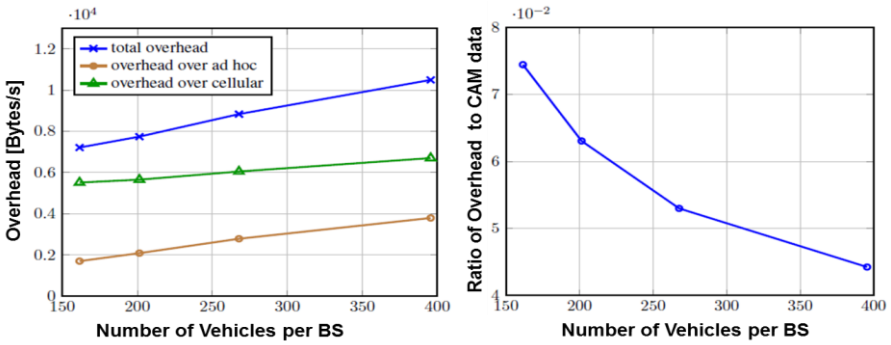


Figure A.4-19. Estimation of the overhead generated over cellular links and V2V links

The results show that the major overhead is generated over the cellular links due to the frequent topology updates by NN as well as the clustering maintenance. It also increases for

high density of vehicles as the number of clusters increases consequently and resulting thereby in more topology changes. Nevertheless, the overhead remains quite low compared to the traffic that could be handled by the new generation of cellular networks (e.g. LTE) and this could be even reduced by adjusting the cellular broadcast according to the topology changes. In other words, high frequency value should be chosen for highly dynamic vehicular network, whereas low update frequency should be selected for less changing environment. Furthermore, it can also be clearly observed from the right graph that the ratio between the total overhead and the traffic generated by the CAM transmissions (useful data) remains very low and even decreases when the density of vehicles increases. This proves the scalability of the proposed architecture and emphasizes the role of the network assistance at highly dense scenarios.

A.4.8 Resource Allocation and Power Control Scheme for D2D-based V2V Communications

Due to the similarity between the QoS requirements of safety critical V2X application and the benefits of D2D communication, the direct D2D link is a promising enabler for V2X communication as long as the RRM is conducted in a careful way. Even though extensive researches are carried on in the context of traditional D2D systems which usually prioritize cellular links but consider the D2D underlay as opportunistic, few of the existing D2D approaches can be directly applied to safety critical V2X communication which has stringent requirements on latency and reliability.

This technology component studies RRM schemes, i.e., RB allocation and power control algorithms, when applying the direct (uplink resource sharing) D2D underlay network to safety critical V2X communications. Under the condition satisfying vehicular users' strict QoS requirements on latency and reliability, it aims at maximizing cellular users' sum rate with fairness consideration. The main contributions are as follows.

1. It is proposed to transform the latency and reliability requirements of safety critical V2X communication into optimization constraints that are computable with only slowly varying CSI [SSB+14][SYS+15]. Specifically, via this method, the original V2X requirements on latency and reliability (e.g., METIS considers that a maximum end-to-end latency of 5 ms, with transmission reliability of 99.999% of 1600 bytes packets should be guaranteed to deliver traffic safety and efficiency applications) can be transformed into the needed number of RBs and the required 'average' SINR constraint on each used RB. The method allows us to extend certain existing D2D RRM algorithms to cater also for V2X communication with strict latency and reliability requirements and still maintain good performance. In contrast, a naive modification of existing algorithms performs poorly.
2. Two RRM problem formulations are proposed for allocating multiple RBs and transmit power to a set of CUEs and VUEs. The problems are both stated as optimization problems with the objective to maximize the CUE sum rate with proportional bandwidth fairness under the constraint of satisfying the VUEs' requirements on latency and reliability. The difference lies in that in the first problem formulation the VUEs are required to use orthogonal RBs [SSB+14], whereas the non-orthogonality among VUEs is allowed in the second problem formulation [SYS+15]. Specifically, the problems are stated as follows.

   Objective: maximize the sum-rate of all CUEs

   Constraints:
   - Orthogonal CUEs;
   - Max sum power constraint for each VUE and each CUE;
   - Required number of RBs for each VUE and each CUE;
   - The 'average' SINR constraint for each VUE on each of its used RBs;
   - **For problem formulation 1**: orthogonal VUEs;

   where the outputs are RB and power allocation results for both VUEs and CUEs.

3. It can be proved that both problems are NP-hard. Therefore, three heuristic algorithms are proposed to approximately solve the RRM optimization problems.
    a. Separate Resource Block allocation and Power control (SRBP)[SSB+14] for **problem formulation 1**: firstly, assuming equal power allocation, the eNB allocates RBs to both VUEs and CUEs in an optimal and time efficient way by transforming the RB allocation problem into an maximum weight matching (MWM) problem for bipartite graphs. Secondly, based on the RB allocation results from the first stage, the eNB further optimally adjusts the transmit power for each VUE and CUE. This is realized via transforming the power control problem into convex optimization which can be solved by certain efficient algorithms.
    b. Separate resOurcebLock and powEr allocation (SOLEN) for **problem formulation 1**: Firstly, by replacing the max sum power constraint for each UE with the max power constraint on each RB, the eNB allocates RBs to both VUEs and CUEs in an optimal and time efficient way by transforming the RB allocation problem into an MWM problem for bipartite graphs. Secondly, based on the RB allocation results from the first stage, the eNB further optimally allocates the transmit power for each VUE and CUE when taking the max sum power constraint into account. This is realized via transforming the power allocation problem into convex optimization and then solving it with a dual decomposition method which can be efficiently computed.
    c. RB Sharing and Power Allocation (RBSPA) [SYS+15] for **problem formulation 2**: firstly, ignoring the max sum power constraints, the eNB allocates RBs to VUEs in a sequential manner according to a specific metric which is designed from the Perron-Frobenius theorem. Secondly, based on the RB allocation results from the first stage, the eNB further optimally adjusts the transmit power for each VUE and CUE. This is realized via transforming the power control problem into convex optimization which can be solved by certain efficient algorithms.

   The common feature to all of the three proposed RRM algorithms is that the RB and power allocation is conducted in a long-term and centralized manner at the eNB.

1.2 Performance evaluation

We assume a single cell outdoor system with a carrier frequency of 800 MHz and that each RB has a bandwidth of 180 kHz for the uplink communication. In particular, we consider a simplified version of test case 2 defined by METIS, which describes an urban environmental model similar to the Manhattan grid layout. In this topology, the entire region is a 444 m*444 m square and the size of each building is 120 m*120 m. See Figure A.4-20 below.

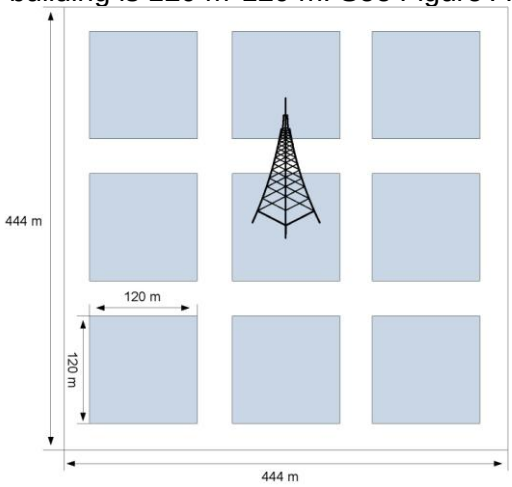


Figure A.4-20. System topology

Regarding channel models, we refer to propagation scenario 3 for the links connected to the eNB, and propagation scenario 9 [MET13-D61] for the links between UEs. Moreover, the max transmit power of each UE is 24 dBm. The intended broadcast range of each vehicle is 18m. The small scale fading of the channels is assumed to have Rayleigh distribution with unit power gain. Further details of simulation parameters can be found in [SSB+14]. Besides, we compare the proposed schemes with the following baseline methods.

1) Ref. in [ZHS10]: to fit the scheme into V2X communication with stringent QoS requirements, we simply decrease the transmit power of the corresponding CUE until the SINR constraint of VUE is satisfied.
2) SRBP-Ext in [SYS+15]: it is a straightforward extension of the SRBP scheme when being applied to non-orthogonal V+UEs.

Firstly, we would like to mention that, as shown by [SSB+14][SYS+15], all the schemes can satisfy the outage probability requirement from METIS on safety critical V2X communication. This result demonstrates the validity of our proposed method for requirement transformation. Also, when small scale fading effects are also involved in the simulated channels, these long-term RRM schemes do not incur big difference on the average performance of the cumulative distribution function of CUEs' sum rate, which verifies the effectiveness of the long-term RRM consideration.

Moreover, Figure A.4-21 evaluates CUEs' sum rate with respect to different number of VUEs, where each VUE is assumed to use 2 RBs and 5 RBs in Figure A.4-21 (a) and Figure A.4-21 (b), respectively. Compared to other schemes, the Ref. exhibits significantly degraded performance, which illustrates the ineffectiveness of a naive modification to certain existing D2D RRM approach when being applied to V2X network. Besides, in the orthogonal VUE model (i.e., problem formulation 1), SOLEN outperforms the Ref. and SRBP. Also, in problem formulation 2 where VUEs are allowed to share the same RBs, RBSPA reveals clear superiority to SRBP-Ext. On the other hand, as shown by Figure A.4-21 (a), SOLEN is slightly better than RBSPA when the number of VUEs is less than 50. This is resulted from the perfect MWM in the first stage of the SOLEN algorithm. Nevertheless, the approaches proposed for the orthogonal VUE model are only feasible for a strictly limited range of VUE numbers, e.g., 50 VUEs in Figure A.4-21 (a) and 20 VUEs in Figure A.4-21 (b).

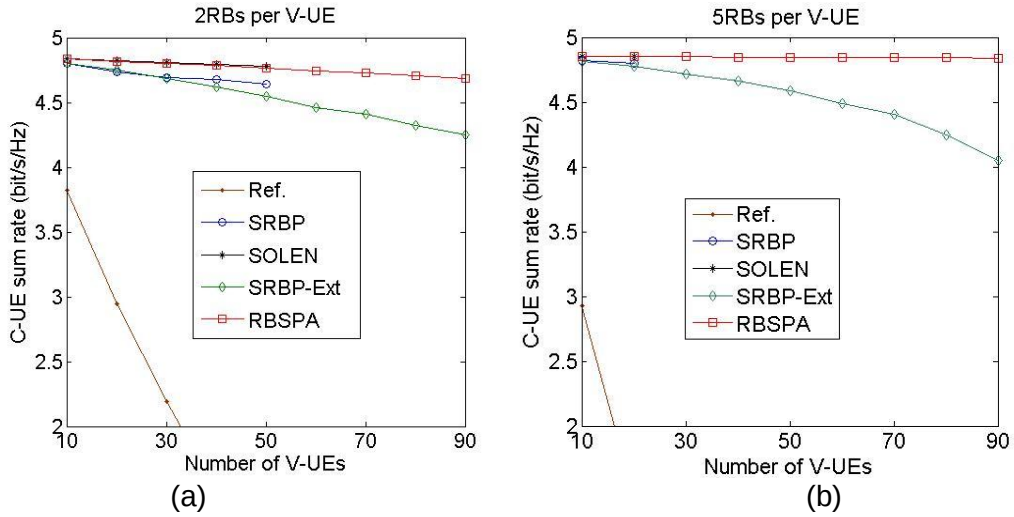


Figure A.4-21. CUE sum rate versus number of VUEs

More simulation results and detailed analyses can be found in [SSB+14][SYS+15].

A.4.9 Joint Methods for SINR Target Setting and Power Control for D2D Communications

For single hop D2D communications, [FSS14] and [MET14-D42] proposed a distributed SINR target setting and power control algorithm that balances between energy efficiency and spectral efficiency. In this Annex we present power consumption and throughput results of that technology component when applied to two-hop D2D communications, such as the range extension scenario illustrated in Figure A.4-22.

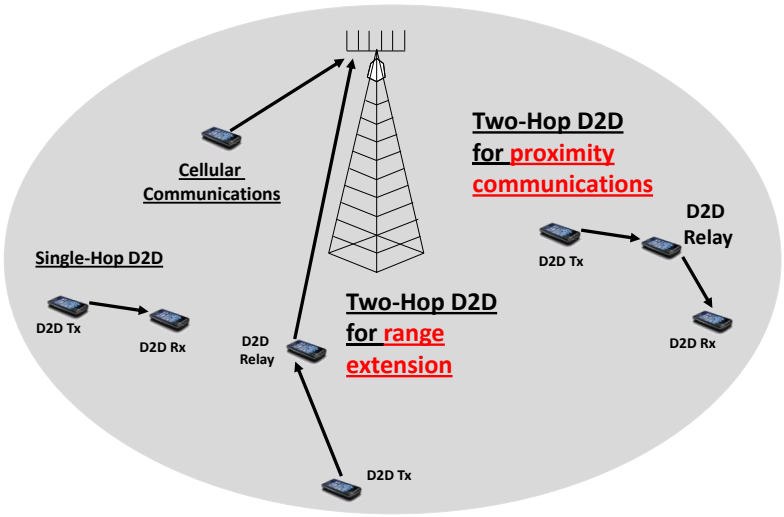


Figure A.4-22. An example of a cellular network supporting single- and multi-hop D2D communications in cellular spectrum. Two-hop D2D communication scenarios include range extension and proximity communications. In range extension, a UE at the cell edge or in outage is helped by a relaying UE. In proximity communications two devices communicate via a relaying UE

As shown in Figure A.4-22, D2D communications can help an out of coverage or cell edge UE to boost its link budget and thereby allow the cell edge or out of coverage UE to communicate with a cellular BS (range extension). Also, D2D communications can use two hops between a D2D transmitter and D2D receiver pair in the so called proximity communications scenario.

Utility maximizing power control is useful to meet spectral and energy efficiency targets by balancing between the inherent trade-off between spectral and energy efficiency by tuning a simple scalar parameter (here referred to as ω):

$$\begin{aligned} \underset{\mathbf{P},\mathbf{s}}{\text{maximize}} \quad & \textstyle\sum_i u_i(s_i) - \omega \sum_{i=1}^{I} \sum_{h=1}^{f(i)} P_{\mathbf{t}(i,h)} \\ \text{subject to} \quad & \tilde{\mathbf{R}}\mathbf{s} \preceq \textstyle\sum_{q=1}^{Q} \mathbf{c}_q(\mathbf{P}), \quad \forall i,h, \\ & \mathbf{P},\mathbf{s} \succeq 0 \end{aligned}$$

The joint method for SINR target setting and power control for D2D communications as formulated above maximizes the sum of utility functions *u(.),* where the summation runs through all communicating D2D pairs (D2D Tx and D2D Rx) and UE-to-BE pairs (cellular communication links). **P** is a vector that represents all transmit power levels and **s**is the vector of achieved rates along all communication routes (where a route can comprise a single hop or a two-hop connection). The index *q* runs through all resources (e.g. frequency channels of a frequency division system) [SFM14]. Finally, the matrix **R** is the so called *routing matrix* that associates (single or two hop) routes with links and resources and is assumed to be constructed by some appropriate relay and mode selection algorithm, which is outside the scope of the present technology component. The function *f(i)* returns the number of hops

along the route *i,* and I denotes the total number of routes (that is the total number of communicating pairs, including D2D and cellular UE-BS pairs) in the system.

Table A.4-3. The investigated power control algorithms. Note that we use the LTE open loop (OL) and fixed power control (Fix) schemes as benchmarks for the proposed utility maximizing (UM) scheme

POWER CONTROL ALGORITHMS

Name	Cellular UE power control	D2D power control
Fix	LTE Open Loop	Fixed Power
Open Loop (**OL**)	LTE Open Loop	LTE Open Loop
Utility Maxim. (**UM-**ω)	Utility maximizing PC with parameter ω	Utility maximizing PC with parameter ω

Figure A.4-22 is the scatter plot showing the energy consumption and throughput performance of the distributed utility maximizing power control scheme in the range extension scenario. As a benchmark, we use the fixed power control scheme ("Fix") and the LTE open loop fractional path loss compensating power control scheme ("OL") and compare their performance with utility maximizing power control scheme ("UM"), when setting the parameter of the UM scheme to ω = 0.1 (high throughput, but high energy consumption) and to ω = 100 (lower throughput energy conserving mode), where ω denotes a degree of penalty for invested transmission power.

Table A.4-4. Simulation parameters for the performance evaluation of a 7 cell system

SIMULATION PARAMETERS

Parameter	Value
Number of BSs	7
Cell radius	500 m
Minimum distance BS-UE	50 m (Scen. 1)/400 m (Scen. 2)
Minimum distance UE-UE	10 m
Mean distance D2D Tx-D2D Rx	100 m
Number of cellular UEs per cell	6
Number of D2D triplets per cell	6 (Scen. 1)/18 (Scen. 2)
Monte Carlo iterations	100
Central carrier frequency	2 GHz
System bandwidth	5 MHz
Number of RBs	18 RBs
Gain at 1 m distance	−37 dB
Thermal Noise power per RB	−116.4 dBm
Path Loss coefficient	3.5
Shadowing standard deviation	8 dB
BS transmit power	40 dBm
UE min/max transmit power	−23 dBm/23 dBm
Fixed Power for LTE PC	−10 dBm
Path loss compensation factor (α)	0.8
SNR/SINR target	15 dB
Number of outer-loop iterations	70
Number of inner-loop iterations	10
ϵ for the outer-loop	0.05
Initial power for the inner-loop	10 dBm
Initial γ^{tgt} for the outer-loop	0 dB
ω of Eq. (2)	[0.1 1 10 100]

We study the power consumption and throughput performance of a 7 cell system with the parameters listed in Table A.4-4. As illustrated in Figure A.4-23, the utility maximizing ("UM") power control scheme reaches the highest average throughput, although it does not show the best SINR for some users, with a gain of approximately 29% over the fixed power control scheme. This behaviour explains how utility maximizing PC achieves the highest throughput.

Further details of the operation of the application of the joint SINR target setting and power control scheme in two hop D2D scenarios are available in [SFM14].

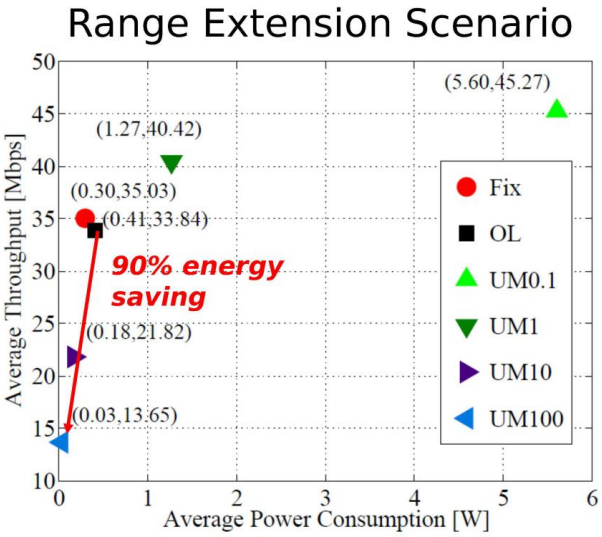


Figure A.4-23. Range extension scenario: Scatter plot of the total power consumption and average throughput achieved by the examined power control algorithms. The utility maximizing PC can reach the highest throughput (with lower values of ω) or the lowest power consumption (with higher ω values). LTE OL provides a reasonable engineering trade-off

A.4.10 Further enhanced ICIC in D2D Enabled HetNets

For further enhanced ICIC in D2D enabled HetNets solution we have conducted system simulations to assess the performance gain. A quasi-static TDD system simulator is used. Both downlink and uplink directions are simulated at the same time during one simulation run.

METIS test case 2 [MET13-D61] is used with full buffer traffic in both directions. 50% of the UEs are assumed to be D2D users and 50% use cellular communication. Users forming a D2D pair must be within 50 meters from each other. The content of the D2D transmission is assumed to be local, i.e. either generated locally in the UE or cached opportunistically earlier. Four small cells are deployed in the park limits (yellow, orange, pink and brown markers in Figure A.4-24). In addition there are three macro cells. 60 UEs are simulated. Figure A.4-24 shows the distribution of cells and an example of the distribution of users over the simulated area. Macro cell, small cell and UE transmit powers are 43 dBm, 30 dBm and 21 dBm, respectively. No power control is applied for the UE.

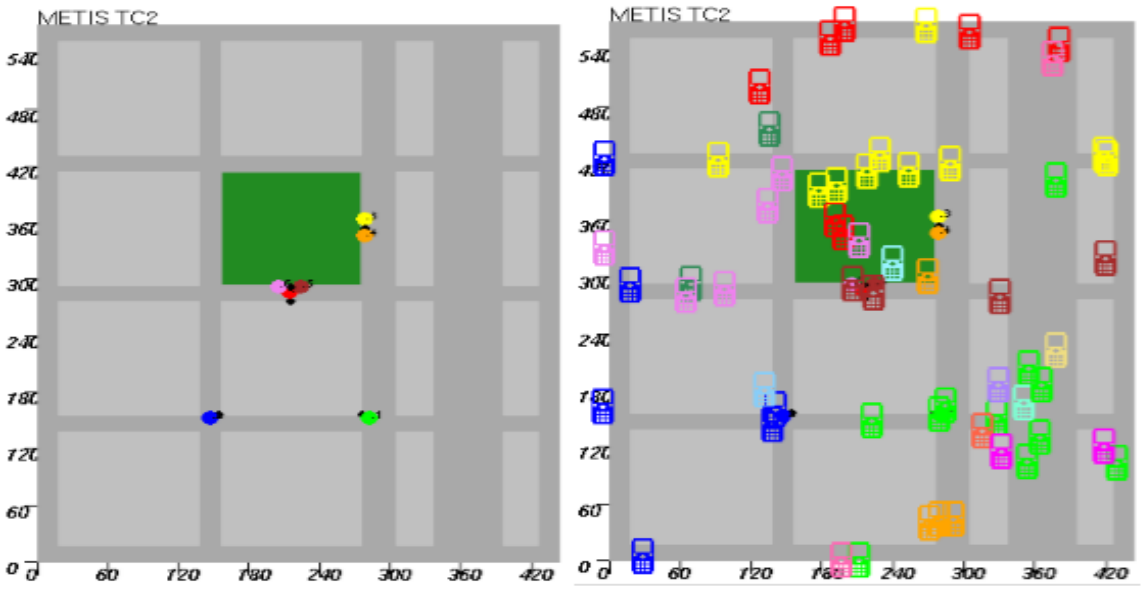


Figure A.4-24. Macro and small cells with UEs (right figure) attached (colour of the UE shows the cell it is attached to)

In the reference case, D2D communication uses macro cell UL resources only, and we compare that to the case where D2D communication uses also small cell UL or DL resources provided that the D2D UEs are far enough from the nearest small cell. In the simulations this configured "safety distance" from nearest small cell was ranging from -120 to -80 dBm (in terms of RSRP).Additionally there was another reference case where the small cell resources were used without restriction for D2D UEs.

Macro and small cell resources are separated in time as shown in Table A.4-5.

Table A.4-5. The separation of macro and small cell resources

Macro	D	U	D	U	-	-	-	-
Small Cell	-	-	-	-	D	U	D	U

The D2D can use three different resource configurations specified in Table A.4-6.

Table A.4-6. The specified D2D resource configuration

D2D #1	-	X	-	X	-	-	-	-
D2D #2	-	X	-	X	-	X	-	X
D2D #3	-	X	-	X	X	-	X	-

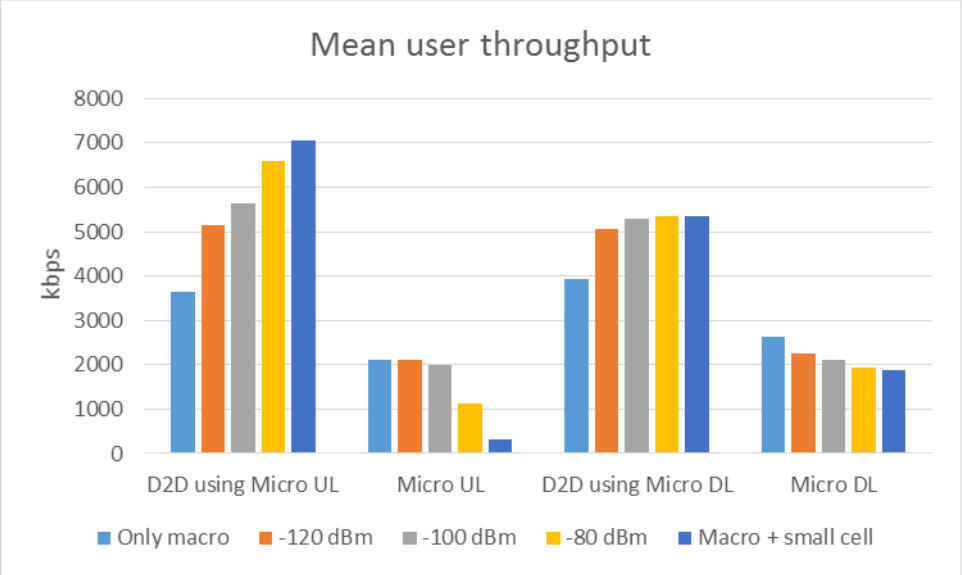


Figure A.4-25. Mean user throughput for different allocation configurations and RSRP thresholds

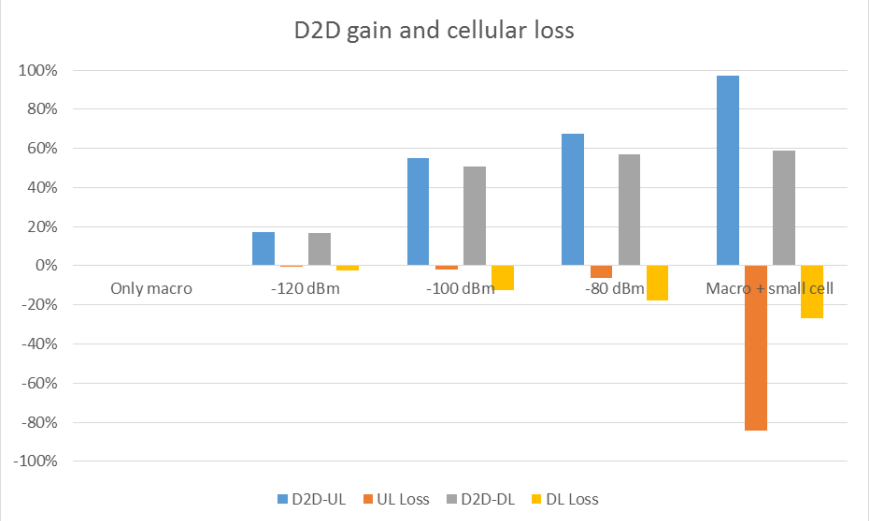


Figure A.4-26. Gains and losses when using D2D with UL resources (D2D-UL) or with DL resources (D2D-DL)

Figure A.4-25 shows the mean user throughput of D2D users and small cell (micro cell) users. As seen from the results, using the resources reserved for small cell can lead to a substantial throughput/capacity gain for D2D UEs. With properly selected safety distance, this can be done with negligible or minor impact on small cell throughput. The exact gains depend on the scenario (in particular the density and coverage of the small cells) as well as the selected trade-off between D2D and small cell capacity. With -80 dBm RSRP threshold there is approximately 60% D2D throughput gain. Figure A.4-26 further illustrates the trade-off between D2D UE throughput gain and the corresponding loss in mean cellular UE throughput in small cells (due to the increased interference from D2D users).

The results indicate that the needed safety distance is bigger in UL than in DL. The interference is more severe in the UL because the D2D interference power is the same as the cellular UL transmit power, whereas in the DL the small cell BS transmit power is assumed 9 dB higher than the interference power of D2D UEs. Optimally a different RSRP threshold would be applied for UL and DL resources.

A.4.11 D2D Handover Schemes for Mobility Management

Simulations were conducted for this component using METIS test case2 [MET13-D61] scenario, with 192 UEs, deployed outdoors, and moving in random directions at 3 km/h, always in a group of two users. The macro/pico/UE transmit powers are 46/30/21 dBm respectively, with 2 GHz operating frequency and 10 MHz bandwidth. METIS antenna model with calibration case 3 was used [MET13-D61], with small cell deployments as shown in Figure A.4-27. Regular handover based on RSRP signal strength satisfying the A3 offset measurement conditions has been used as the reference case for evaluations. The D2D aware mechanisms tries to keep both UEs in a D2D group connected to the same cell, by possibly delaying the handover of one of the UEs when it crosses the cell boundary, as long as the SINR of the UEs in the group is > -6 dB.

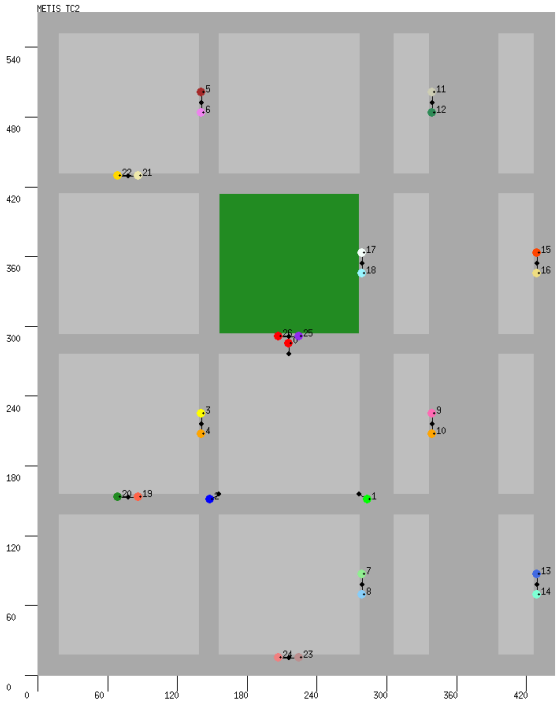


Figure A.4-27. METIS test case 2 scenario.

Simulations were done to compare the performance of the D2D aware handover mechanism with the reference LTE-A Release 10 mechanism, in terms of same cell connected time distribution, also emphasizing of the values for users which are closer and further away from the cell center. The evaluations were also done for D2D group connection distance of 20 m and 100 m. The CDF results for the mechanisms and the mean values are as shown in Figure A.4-28. From the figure, we can observe that there is significant improvement (> 28% for 20 m D2D distance), in same cell connected time using the D2D aware mobility management mechanism. The CDF curves also indicate that the connection times when UEs are in the same cells is clearly higher than the time spent on different cells.

The 50th and 95th percentile results are also shown in Figure A.4-29. It can be observed from the figures that the gains for 20 m D2D connection distance are much higher for the 95th percentile users, as compared to the 50th percentile ones. This also indicates the impact of cell geometry on mobility performance, since with a macro cell overlay, even with the ultra-dense deployment of small cells considered in the scenario, the probability of staying with the same cell is considerably less for the 100 m D2D connection distance, as compared to 20 m.

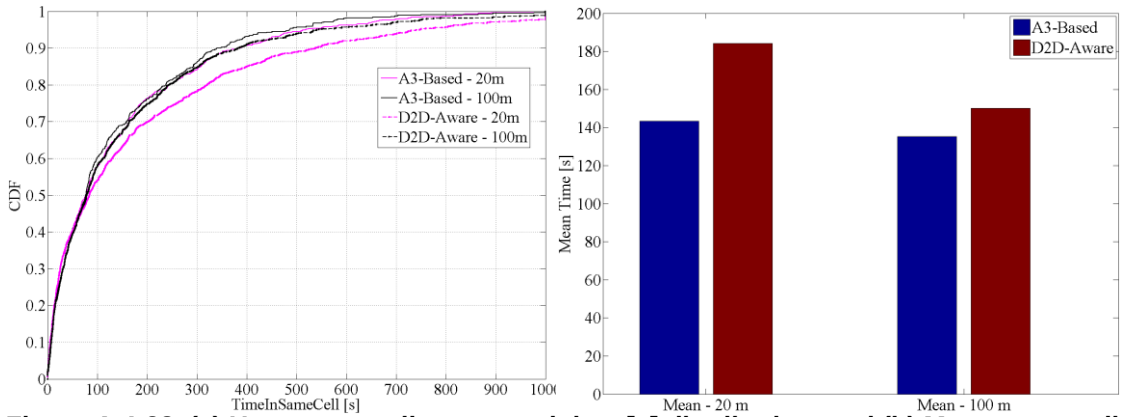


Figure A.4-28. (a) User same cell connected time [s] distribution, and (b) Mean same cell connected time [s] values.

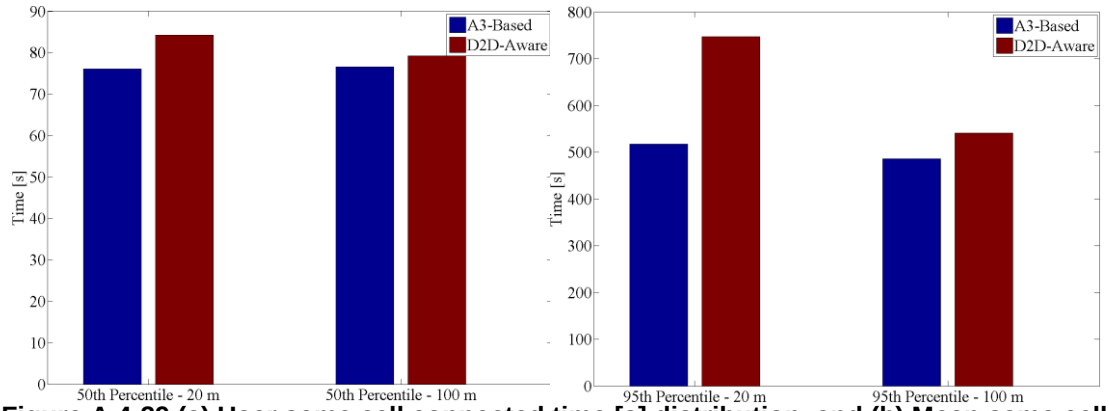


Figure A.4-29.(a) User same cell connected time [s] distribution, and (b) Mean same cell connected time [s] values.

From the results, we can conclude that the D2D aware mechanism improves the mobility performance significantly by delaying the handover of one of the UEs in the group to a different cell, thereby avoiding the added complexity involved in such an operation. While the additional cost of implementing the mechanism is the delayed or early handover, which might affect the UE-eNB link quality, the SINR threshold used ensures that significant performance degradation in terms of radio link failure, etc., is avoided.

A.5 Dynamic Reconfiguration Enablers

A.5.1 Energy Savings Schemes for Phantom Cell Concept Systems – Macro-assisted UE-Small Cell Connection Procedures

This section describes in details the macro-assisted UE-small cell connection flow present in the various energy savings schemes for PCC systems, as introduced in Section 6.2 of this document.

Uplink Signalling-based Energy Savings Scheme

When using the UL signalling-based energy savings scheme, small cells are woken up by UEs, which send wake-up signals on the small cell air interface. A typical UE-small cell connection flow, illustrated in Figure A.5-1, goes as follows:

1. The macro cell sends a *wake-up signal configuration* message to the UE to trigger the UE-small cell connection procedure and to potentially configure the wake-up signal from the UE to have a limited scope (e.g. affecting only small cells belonging to a certain closed subscriber group or offering a certain bandwidth);
2. The UE sends the wake-up signal on the small cell air interface so that small cells in sleep mode are turned on. Small cells are able to detect this signal since their radio-frequency receiving chain is turned on even when they are in sleep mode;
3. The UE receives pilot symbols from all surrounding small cells and performs small cell channels estimations;
4. The UE sends a report of the channel quality measurements to the macro cell;
5. The macro cell determines the best small cell for the UE to connect to, based on the UE's measurement report and on additional information reported from small cells (such as the current small cell load), and informs the UE of the decision;
6. The UE starts the connection procedure with the best selected small cell, while remaining connected to the macro cell.

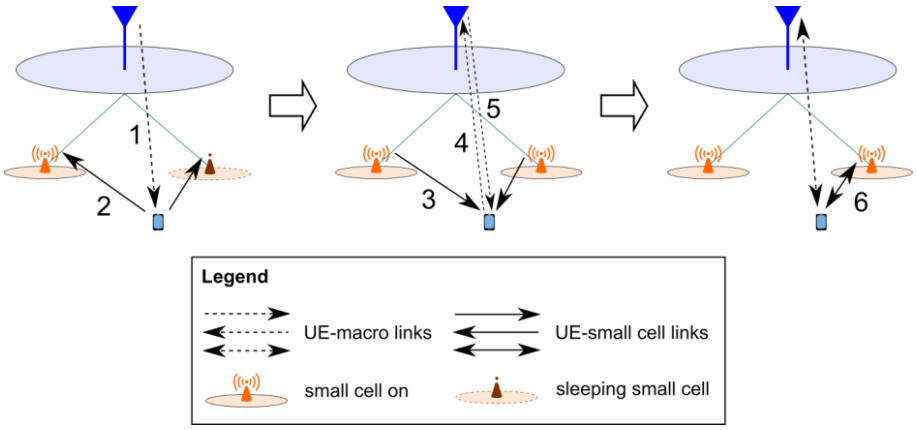


Figure A.5-1. Illustration of the UE-small cell connection procedure in the UL signalling-based energy savings scheme

Downlink Signalling-based Energy Savings Scheme

In the DL signalling-based energy savings scheme, it is assumed that small cells send pilot symbols to UEs even when in sleep mode, albeit at a reduced rate compared to small cells in on mode. A typical UE-small cell connection flow, illustrated in Figure A.5-2, goes as follows:

1. The UE receives pilot symbols from all surrounding small cells, both in on mode and sleep mode, and performs small cell channels estimations;
2. The UE sends a report of the channel quality measurements to the macro cell;
3. The macro cell determines the best small cell for the UE to connect to, based on the UE's measurement report and on additional information reported from small cells (such as the current small cell load), informs the UE of the decision, and turns on the best selected small cell through the backhaul link (if needed);
4. The UE starts the connection procedure with the best selected small cell, while remaining connected to the macro cell.

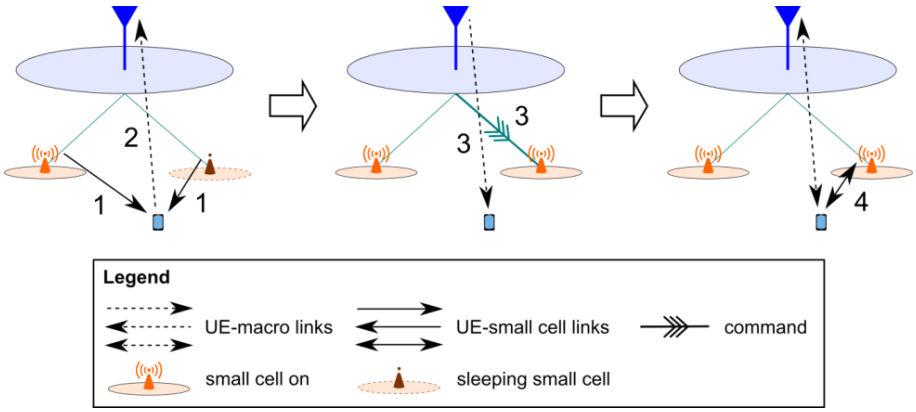


Figure A.5-2. Illustration of the UE-small cell connection procedure in the DL signalling-based energy savings scheme

Database-aided Energy Savings Scheme

The database-aided energy savings scheme relies on the principle of storing UE-small cell links quality data (such as SNR values) on a database at the macro cell, so that this data is available even when the small cell channel is not measurable by UEs (e.g. when small cells are in sleep mode and therefore do not transmit pilot symbols). A typical UE-small cell connection flow, illustrated in Figure A.5-3, goes as follows:

1. The UE receives pilot symbols from all surrounding on mode small cells and performs small cell channels estimations;
2. The UE sends a report of the channel quality measurements to the macro cell, as well as information regarding its geographical position, and the macro cell retrieves the stored sleep mode small cell channel quality information from the database;
3. The macro cell determines the best small cell for the UE to connect to, based on the UE's position and channel quality measurement report, the retrieved database values (using the UE's position as input), and on additional information reported from small cells (such as the current small cell load), informs the UE of the decision, and turns on the best selected small cell through the backhaul link (if needed);
4. The UE starts the connection procedure with the best selected small cell, while remaining connected to the macro cell.

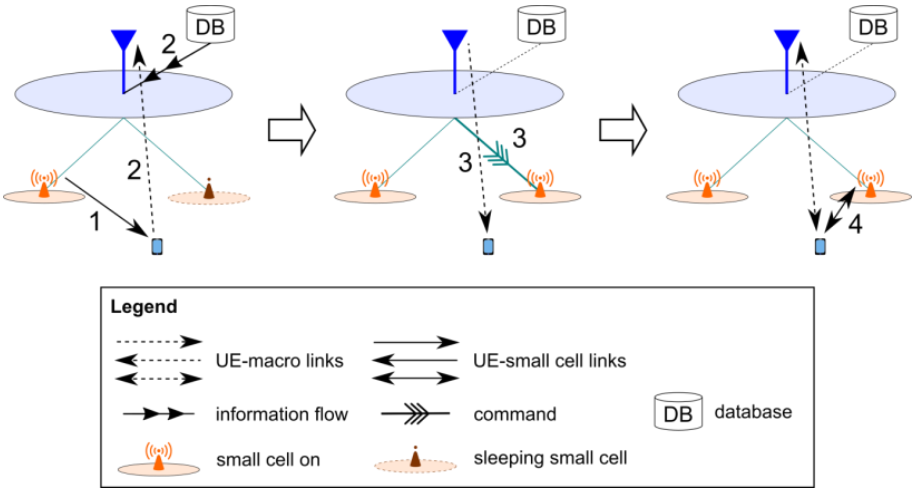


Figure A.5-3. Illustration of the UE-small cell connection procedure in the database-aided energy savings scheme

A.6 Dependencies between investigated technology components

This section contains the analysis of interaction between technology components presented in this deliverable. It is not based on architectural dependencies (which is captured in [MET15-D64]), but rather reflects possible information flow between investigated solutions. Numbers in brackets refer to a specific section in D4.3 where given mechanisms are explained.

Analysis was done for technology components grouped into four different categories, with some overlapping between them.

1. Radio resource and interference management
2. D2D and V2V operations
3. Mobility management including context awareness
4. Dynamic reconfiguration of the network

Arrow's direction in Figure A.6-1 to Figure A.6-4 shows output of which solutions can be used as an input to another.

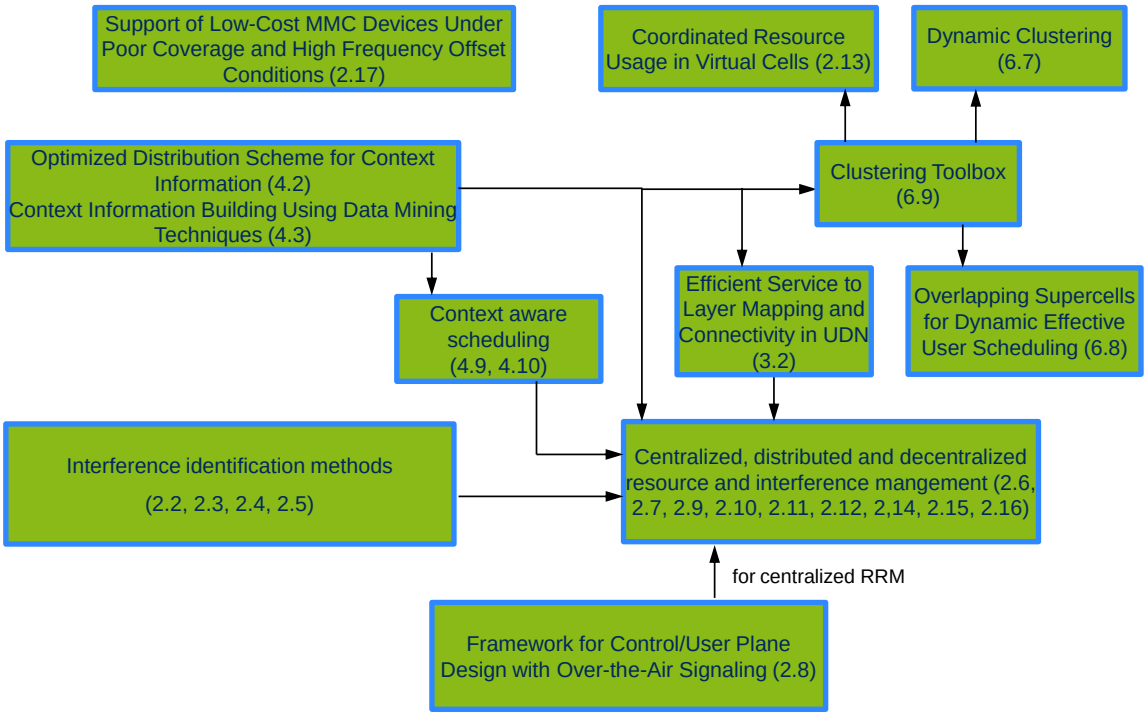


Figure A.6-1. Radio resource and interference management dependecies

Several solutions in RRM area depicted in Figure A.6-1 use context information as an input. Information on the location and prediction of user's trajectory, availability of a cell with a large capacity or a warning about expected coverage hole is used as an input for a proactive scheduling (e.g., delaying/prioritizing a certain data transmission). Context information can be also used as an input to the clustering mechanisms, which in turn feed schemes based on grouping of the cells for interference coordination. Finally, both context information and interference identification methods can be used for appropriate radio resource allocation for MMC operations. Interference identification solutions can also be used to support centralized and distributed resource and interference management schemes. The performance of these schemes can be also dependant on the control/user plane split solutions used for conveying reliable control information. Some of the radio resource management concepts investigated in this document can also benefit from clustering toolbox mechanisms.

There are several approaches proposed for centralized and distributed/decentralized RRM (more information on centralization/decentralization for RRM can be found in [MET14-D42]). These schemes can operate in a complementary manner. If there is a centralized controller available, such as a macro station that facilitates RRM in small cells under its coverage, then centralized schemes can be used. If there is no wide area coverage, distributed/decentralized schemes can be pursued. Investigated solutions differ also in the time scale of their operations. As an example RRM mechanism developed in Section 2.6 operates with sub millisecond resolution (if air interface for UDN is assumed [MET15-D24]), while concept from Section 2.10 provides a long time scale strategies with the resolutions of minutes or hours.

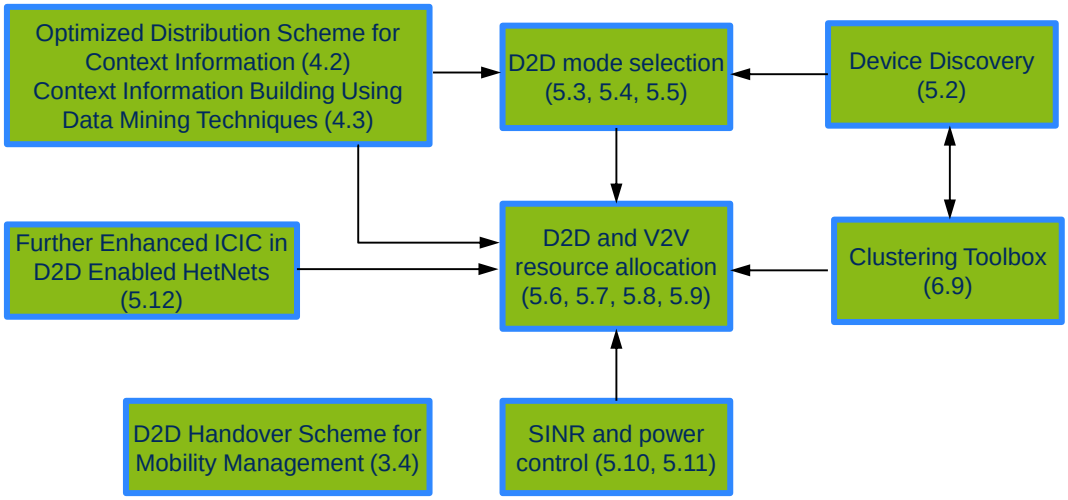


Figure A.6-2. D2D and V2V solutions dependecies

METIS investigations provide a holistic system level concept for solutions based on direct D2D communications, as depicted in Figure A.6-2. Proposed mechanisms cover the initial phase (device discovery), resource allocation, mode selection and SINR/power control. To provide a high performancing V2V communication also context and clustering information may be needed.

There are different options proposed for device discovery, depending, e.g. on the availability of the network coverage. Also for the RRM and mode selection several alternative solutions were investigated. One division line is the usage of location information or channel state information. Alternatively those solutions can operate with long (seconds) or short (milliseconds) resolutions for the decisions process. Trade off investigations related to these operations are treated in [MET14-D42].

Context awareness solutions (including mobility management enhancements) shown in Figure A.6-3 are heavily dependent on the technology components related to obtaining and distribution of context information. The most common context information is the location and trajectory prediction or their derivatives (e.g. availability of connection to a high capacity cell in the nearest future). The performance of considered schemes is dependent on the accuracy of the context information, but often update of accurate context information may result in an increased signalling overhead. Studies of complexity vs. performance improvement and information interflow vs. throughput/mobility enhancement are part of [MET14-D42].

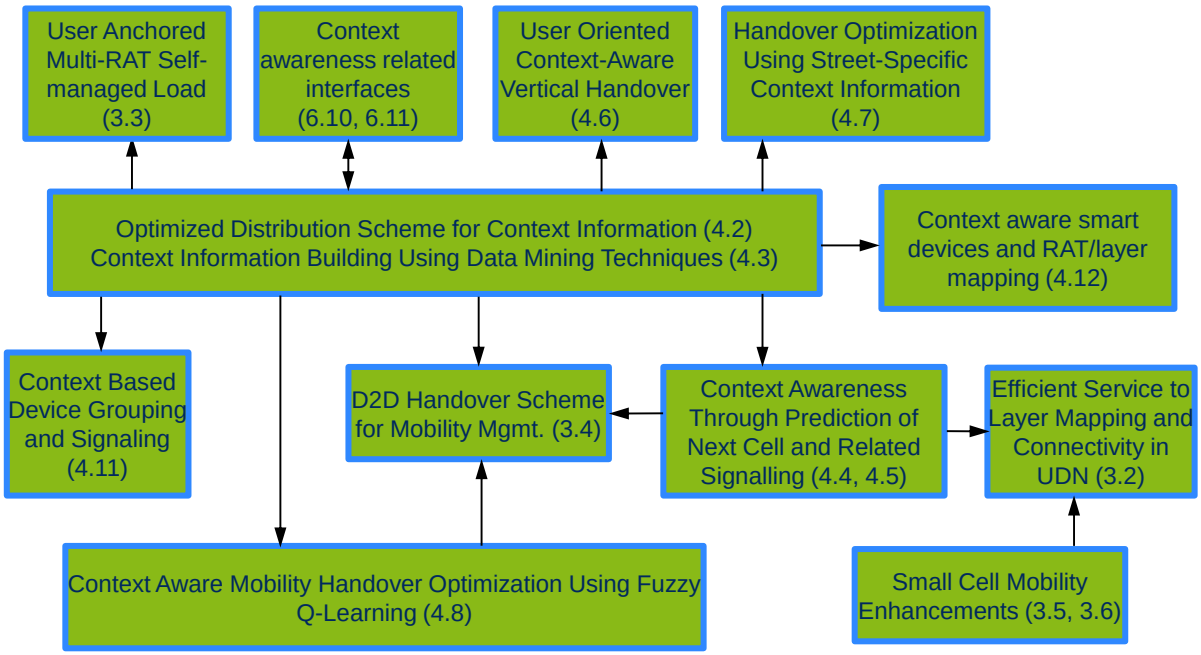


Figure A.6-3. Mobility management and context awareness solutions dependencies

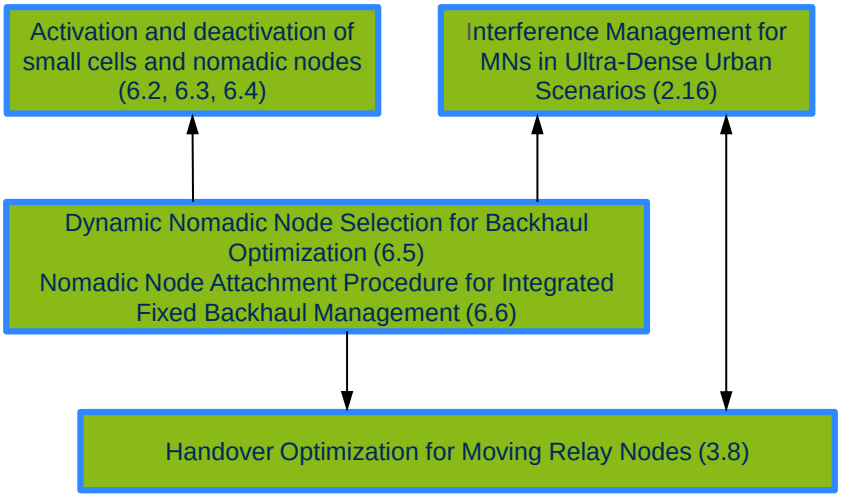


Figure A.6-4. Dependencies between solutions for dynamic reconfiguration of the network

The main focus of dynamic reconfiguration solutions from Figure A.6-4 is activation or deactivation of network nodes or nomadic cells. An important enabler for the latter group is the wireless backhaul optimization and the efficient resource sharing between access and backhaul link. This enabler is also important for the solution focusing on RRM and mobility management of moving relay nodes.