VoIP Performance of LTE Networks
VoLTE versus OTT

Von der Fakultät für Elektrotechnik und Informationstechnik
der Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines Doktors der
Ingenieurwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Ingenieur
Maciej Grzegorz Mühleisen

aus Wrocław

Berichter: Universitätsprofessor Dr.-Ing. Bernhard Walke
Universitätsprofessor Dr.-Ing. Andreas Timm-Giel

To my beloved wife Patrycja...
Long Term Evolution (LTE) standardized by 3GPP operates fully packet switched and does not have a circuit switched domain as older systems HSPA, UMTS and GSM. It must support a minimum number of voice calls to be recognized as IMT-Advanced system by ITU. Therefore ITU has published the IMT-Advanced Evaluation Guidelines describing evaluation scenarios and minimum quality of service demands for voice calls to be evaluated by system level simulation. Those guidelines require transmission of VoIP packets within 50 ms over the air interface and packet loss rate below 2 %.

3GPP came to the conclusion Voice-over-LTE (VoLTE) with semi-persistent scheduling satisfies the ITU demands. This was verified by independent evaluation groups. Semi-persistent scheduling reserves radio resources for VoIP packets for the duration of a talk spurt while other packets, e.g. retransmissions of lost packets, individually receive resources (dynamic scheduling). Advantage is significantly decreased signaling overhead but the method is more complex than the classical approach individually assigning radio resources for each packet. Individual assignment of radio resources is also used for data traffic of so called Over-the-Top (OTT) VoIP applications.

This thesis compares VoIP capacity of VoLTE and OTT focusing on the impact of the radio link. Radio resource assignment optimized for delay critical traffic like VoIP is assumed for OTT to allow a fair comparison. A concept developed by me to design complex VoIP schedulers from basic resource assignment algorithms is described. It is applied for VoLTE and OTT. The openWNS/IMTAPhy simulator developed at RWTH Aachen University ComNets Research Group and extended at TU Munich Chair of Communication Networks is enhanced for VoIP evaluation based on that concept. Methods to improve
VoIP capacity are evaluated separately for up- and downlink. Different channel prediction strategies are applied for each direction and based on that transmission rate adaptation is performed.

Results show improved capacity through aggressive rate adaptation and resource assignment resulting in higher data rates and therefore reduced resource demand while at the same time increasing packet error rate. Too aggressive parameters significantly reduce VoIP capacity while moderate adjustments increase it by ca. 15% in the downlink and 10% in the uplink for VoLTE. OTT cannot benefit from these improvements if the number of control channels is limited and becomes a bottleneck. If more control channels were available than defined by the current standard OTT VoIP capacity could be improved for downlink. High uplink interference variance makes channel prediction more difficult and prevents capacity gains for OTT as achieved with VoLTE.

Besides simulation results an analytical discrete time Engset queueing model was described and solved for its waiting time distribution to determine VoIP capacity in scenarios without and with adjacent channel interference from neighbor cells.

In the Urban Macrocell scenario lowest downlink VoIP capacity results 43 users / cell / MHz for OTT and 41 users / cell / MHz for VoLTE were obtained with unlimited number of control channels. The Downlink is the bottleneck and therefore determines the overall VoIP capacity. Allowing users with bad channel condition to transmit at higher data rates than suggested by the rate adaptation algorithm reduces resource occupation but at the same time increases packet error rate. VoIP capacity is increased to 48 and 47 users / cell / MHz this way for OTT and VoLTE, respectively.

From this it is concluded that dynamic scheduling used by OTT is capable to provide sufficient VoIP capacity without the complexity introduced by SPS given that sufficient control channels are available, as expected in reality.

Die 3GPP ist zum Schluss gekommen, dass Voice-over-LTE (VoLTE) unter Ausnutzung von semipersistenter Funkressourcenzuteilung die Mindestanforderungen der ITU erfüllt, was auch von unabhängigen Gutachtergruppen bestätigt wurde. Bei der semipersistenteren Funkressourcenzuteilung werden Funkbetriebsmittel für VoIP Pakete für die Dauer eines Talk-Spurts reserviert während anderen Pakete, beispielsweise wiederholten Übertragungen fehlerhaft empfangener Pakete, einzeln Ressourcen zugewiesen werden. Der Vorteil ist ein deutlich reduzierter Signalisierungsaufwand, doch das Verfahren ist komplexer als der klassische Ansatz jedem Paket einzeln Funkbetriebsmittel zuzuweisen. Letzteres wird auch für den Datenverkehr von so genannten Over-the-Top (OTT) VoIP Anwendungen eingesetzt.

Diese Arbeit vergleicht die VoIP Kapazität von VoLTE mit der von OTT, wobei der Schwerpunkt auf dem Einfluss der Funkschnittstelle liegt. Für einen fairen Vergleich wird angenommen, dass die Funkressourcenzuweisung für OTT für


Neben Simulationsergebnissen wurde auch ein analytisches zeitdiskretes Engset Wartemodell beschrieben und die Verteilung der Wartezeitverzögerung von diesem bestimmt, um die VoIP Kapazität im interferenzfreien Fall und bei Gleichkanalinterferenz durch Nachbarzellen zu bestimmen.
Im Urban Macrocell Szenario ergibt sich die geringste VoIP Kapazität. Diese beträgt 43 Benutzer / Zelle / MHz für OTT, wenn man von einer unbeschränkten Anzahl Signalisierungs kanäle ausgeht, und 41 Benutzer / Zelle / MHz für VoLTE in der Abwärtsstrecke, die den Engpass darstellt und somit die Gesamtkapazität bestimmt. Lässt man Benutzer mit schlechter Kanalqualität mit höheren Datenraten übertragen als vom Ratenanpassungsalgorithmus bestimmt, so dass der Ressourcenverbrauch sinkt, dafür aber die Paketfehllrate steigt, so erhöht sich die Kapazität auf 48 und 47 Benutzer / Zelle / MHz für OTT und VoLTE.

Die Schlussfolgerung ist, dass dynamische Ressourcenzuweisung, wie sie von Over the Top (OTT) verwendet wird, eine ausreichende Voice over IP (VoIP) Kapazität sicherstellt unter der realistischen Vorrausetzung, dass genügend Signalisierungs kanäle verfügbar sind.
Contents

Abstract v
Kurzfassung vii

1 Introduction 1
1.1 Motivation 2
1.2 Objectives 3
1.3 Contribution of this Thesis 4
1.4 Outline 4

2 3GPP Long Term Evolution 7
2.1 OFDMA Resource Grid 7
2.2 User Plane Data Transmission Procedure 11
   2.2.1 Packet Data Convergence Protocol (PDCP) Sublayer 12
   2.2.2 Radio Link Control (RLC) Sublayer 13
   2.2.3 Medium Access Control (MAC) Layer 14
2.3 Scheduling 20
   2.3.1 Signaling 20
   2.3.2 Semi-Persistent Scheduling 22

3 Related Work 25
3.1 Voice Call Performance Evaluation 25
   3.1.1 VoIP Capacity State-of-the-art 28
   3.1.2 Scheduling Algorithms for VoIP 31
   3.1.3 Link Adaptation (LA), Power- and Admission Control 32
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.4</td>
<td>Analytical Evaluation of VoIP Capacity</td>
<td>33</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Activities of the 3rd Generation Partnership Project (3GPP) Consortium Prior to IMT-Advanced (IMT-A) Evaluation</td>
<td>35</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Long Term Evolution Advanced (LTE-A) IMT-A Evaluation Results</td>
<td>37</td>
</tr>
<tr>
<td>4</td>
<td>IMT-Advanced Evaluation Guidelines</td>
<td>43</td>
</tr>
<tr>
<td>4.1</td>
<td>Scenarios</td>
<td>43</td>
</tr>
<tr>
<td>4.2</td>
<td>Channel Model</td>
<td>44</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Large Scale Channel Model</td>
<td>44</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Small Scale Channel Model</td>
<td>49</td>
</tr>
<tr>
<td>4.3</td>
<td>System Model Baseline Assumptions</td>
<td>50</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Cell Assignment</td>
<td>50</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Protocol Overhead</td>
<td>50</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Error Modeling</td>
<td>52</td>
</tr>
<tr>
<td>4.4</td>
<td>Traffic Source Models</td>
<td>52</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Full Buffer Traffic Model</td>
<td>53</td>
</tr>
<tr>
<td>4.4.2</td>
<td>VoIP Traffic Model</td>
<td>53</td>
</tr>
<tr>
<td>5</td>
<td>LTE VoIP Queueing Model</td>
<td>55</td>
</tr>
<tr>
<td>5.1</td>
<td>Related Work</td>
<td>55</td>
</tr>
<tr>
<td>5.2</td>
<td>Model Description</td>
<td>57</td>
</tr>
<tr>
<td>5.3</td>
<td>Numerical and Simulation Results</td>
<td>64</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Interference Free Scenario</td>
<td>64</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Scenario with Interference</td>
<td>67</td>
</tr>
<tr>
<td>6</td>
<td>Mathematical Analysis of Indoor Hotspot (InH) Scenario</td>
<td>73</td>
</tr>
<tr>
<td>6.1</td>
<td>downlink (DL) Throughput Capacity in InH Scenario</td>
<td>73</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Signal to Interference and Noise Ratio (SINR) Distribution</td>
<td>74</td>
</tr>
</tbody>
</table>
**Contents**

6.1.2 Throughput Distribution and Cell Spectral Efficiency ........................................... 76
6.1.3 Numerical Evaluation ......................................................................................... 79

6.2 Uplink (UL) SINR Distribution in InH Scenario ...................................................... 80
6.2.1 SINR Distribution .............................................................................................. 81

7 System Level Simulation Model and Results .............................................................. 87
7.1 VoIP Scheduler for Multi-Cell Scenarios ................................................................. 88
7.1.1 Queues and Scheduling Strategies .................................................................. 88
7.1.2 Call State Estimation ......................................................................................... 89
7.1.3 Resource Assignment and Link Adaptation (LA) ........................................... 91
7.1.4 Queue Management ......................................................................................... 100
7.1.5 Preempting Persistent Reservations ............................................................... 101

7.2 Scenario .................................................................................................................. 103
7.3 VoIP Capacity Results ............................................................................................ 107
7.3.1 Optimal Queue Sizes and Discard Timers ....................................................... 107
7.3.2 Relocation of Persistent Reservations ............................................................. 110
7.3.3 DL with OTT Scheduling of VoIP Packets ...................................................... 112
7.3.4 DL with SPS of VoIP Packets ......................................................................... 117
7.3.5 Uplink ............................................................................................................... 119
7.3.6 Comparison of the Impact of Resource Assignment Algorithms on VoIP Capacity ................................................................................................................. 123
7.3.7 Preemption of Persistent Reservations ............................................................ 129
7.3.8 Soft Frequency Reuse (SFR) and VoIP Capacity in the UL ............................. 132
7.3.9 Single Input Multiple Output (SIMO) ............................................................... 135
7.3.10 Individual Contribution of the Algorithms and Parameters .......................... 137

8 Conclusions ................................................................................................................ 141

A Equation Proofs ....................................................................................................... 143

List of Symbols ............................................................................................................. 145
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bibliography</td>
<td>155</td>
</tr>
<tr>
<td>Acronyms</td>
<td>173</td>
</tr>
<tr>
<td>Curriculum Vitae</td>
<td>181</td>
</tr>
<tr>
<td>Acknowledgment</td>
<td>183</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

Content

1.1 Motivation .................................................. 2
1.2 Objectives .................................................. 3
1.3 Contribution of this Thesis .......................... 4
1.4 Outline ...................................................... 4

There is no doubt about high and growing demand for mobile data services with video services currently being the main contributor to this trend resulting in almost 60% average annual growth expected for 2013-2019 [1]. Despite this the total duration of circuit switched mobile voice calls experienced 13% average growth in 2008-2012 with declining growth rate reaching 4% in 2012 [2]. Still further growth at moderate rate is expected [3] with high variance depending on the level of development of the respective geographic region. The trend for SMS is different showing a decrease by 20% in 2013 in North America [3]. SMS is clearly being replaced by OTT messaging services using Internet Protocol (IP) data transmission service at significantly lower cost.

The most prominent OTT messaging service by 2015 is WhatsApp used and installed on every fourth Internet enabled cell phone worldwide [4]. Plans to add VoIP call functionality to the application were recently announced and other OTT voice applications have been available for several years. Different than for traditional circuit switched voice services those calls are charged by data volume and not by duration and regardless of location of the person being called. Still an decline as
Chapter 1 – Introduction

for SMS is currently not happening for voice calls. One potential reason are the more stringent Quality of Service (QoS) demands with regard to delay and packet loss of voice compared to messaging services. OTT application usually only get access to best effort data services and could be unable to reach the quality of voice calls transmitted through dedicated circuit switched resources.

4G radio standard LTE-A does not include a circuit switched domain and uses VoIP to offer voice call services. The term Voice over LTE (VoLTE) was established in order to distinguish it from previously described OTT voice services. VoLTE traffic receives highest priority through the IP Multimedia Subsystem (IMS). As of today 16 operators from 7 different countries including Germany, USA, and Japan have commercially launched VoLTE in their networks and 90 other operators from 47 countries are preparing it [5]. They are jointly developing Rich Communication Services (RCS) [6] defining standard interfaces and data formats for multimedia applications in competition to OTT application developers. RCS compliant VoLTE applications are therefore compatible what is usually not true for OTT voice applications from different vendors.

While OTT offers VoIP services at lower cost than VoLTE the later might offer better perceived performance by the user. An important role plays the VoIP performance of the radio link which is often the bottleneck of the end-to-end connection. That is why Long Term Evolution (LTE) radio link performance is assessed in this thesis using well accepted evaluation guidelines from ITU.

1.1 Motivation

VoIP capacity results for LTE under similar modeling assumptions were published by 3GPP as well as different independent evaluators. The author himself was involved in one of the independent evaluation studies. While final results with regard
to LTE VoIP capacity are publicly available it remains unclear which radio resource algorithms and parameter settings lead to the particular results and the portion each of those contributed to the overall result. Semi-Persistent Scheduling (SPS) persistently reserves resources for VoIP packets every 20 ms while still scheduling Hybrid Automatic Repeat Request (HARQ) retransmissions dynamically was identified as the preferable scheduling algorithm in this context. Currently OTT applications for voice calls receive increased popularity. Those applications only have access to the default bearer providing best effort data transmission services. Respective VoIP packets are scheduled dynamically on a per packet basis. The performance of the two different voice call services (SPS and OTT) is compared in this thesis under fair assumptions allowing to primarily focus on the impact of scheduling and radio resource management on VoIP capacity.

1.2 Objectives

Objective of this thesis is to compare VoIP capacity in LTE achieved using SPS as done in VoLTE against dynamic scheduling as used by OTT voice applications. Radio resource assignment algorithms, especially ones known to improve cell spectral efficiency (CSE) and cell edge user spectral efficiency (CEUSE), are evaluated with regard to their impact on VoIP capacity. Also the impact of limited control channel resources is assessed. Besides presenting capacity results for different evaluation scenarios also the individual contribution of different algorithms and parameters is presented. The key question is whether evaluated methods are capable to improve QoS for the users experiencing worst performance since their packet loss and delay determines overall system VoIP capacity.
1.3 Contribution of this Thesis

In this thesis an IMT-A compliant VoIP capacity evaluation for LTE is presented. The most important contributions of this thesis are explicitly named in Chapter 8, Conclusion. It is shown by event driven simulation and analytical calculation that the capacity is much higher than requested by International Telecommunication Union Radiocommunication Sector (ITU-R) for IMT-A systems. It is also shown what impact scheduler parameters and algorithms have on VoIP capacity. A comparison of VoIP service capacity when realized OTT without radio link specific support to VoIP capacity of LTE under SPS gives evidence that LTE signaling capacity is the limiting resource under OTT scheduling. VoIP capacity for OTT is further evaluated without considering signaling capacity limitations motivated by the fact that in reality systems are not fully loaded with VoIP users as defined in the IMT-A Evaluation Guidelines [7] used in this thesis. In this case VoLTE and OTT show similar VoIP capacity results.

1.4 Outline

This thesis is organized as follows: Chapter 2 gives an overview of Layer 1 (L1) and Layer 2 (L2) of the 3GPP LTE standard focusing on header and control traffic overhead for small packets as used for VoIP. In Chapter 3 the state of the art is presented. Key concepts to increase VoIP capacity are described and VoIP capacity results from different studies are compared. The evaluation guidelines and modeling assumptions used in this thesis are explained in Chapter 4. This includes the IMT-A evaluation scenarios and the VoIP traffic model together with definition of the satisfied user criterion determining VoIP capacity. In Chapter 5 an Engset queueing model to calculate the VoIP capacity under SPS is described and VoIP capacity results for scenarios with and without interference are derived from
it. Chapter 6 presents an analytical model used to validate simulation results for a two cell scenario. Results presented there give strong evidence the model also creates valid results in the multi-cell scenarios evaluated in the following Chapter 7. The structure of the VoIP scheduler is presented in Chapter 7.1. The scheduler is capable of SPS as well as dynamic scheduling used for VoIP traffic originating from OTT voice call applications. The evaluated scenarios are described in Chapter 7.2 together with derivations from the evaluation guidelines presented in Chapter 4. Chapter 7.3 presents and compares simulation results for UL and DL for SPS used by VoLTE and dynamic scheduling used by OTT voice call applications. The work is summarized in Chapter 8 and an outlook on future work is given.

In Appendix A proof for mathematical equations used is provided.
For better understanding of the problems resulting from providing VoIP in LTE, first the relevant system components are introduced here. This includes a detailed description of all overheads having a significant impact on VoIP performance due to the small VoIP packet size. Further LTE control signaling is described since limited control channel capacity is potentially limiting VoIP capacity of LTE.

### 2.1 OFDMA Resource Grid

Figure 2.1 shows the Orthogonal Frequency Division Multiple Access (OFDMA) radio frame of the LTE standard. A configuration with 1.4 MHz channel bandwidth $BW_{Ch}$ is shown with 72 Orthogonal Frequency Division Multiplexing (OFDM) subcarriers divided into 6 Physical Resource Blocks (PRBs) each comprising 12 subcarriers. Table 2.1 lists other possible configurations.
channel bandwidths $BW_{Ch}$ and the according number of PRBs as specified in LTE Release 8 [9, 10]. Ten subframes in time domain form one radio frame. A radio frame describes the periodicity of the Broadcast Control Channel (BCH) and, in Time Division Duplex (TDD) mode, not further considered in this thesis, the switching points between DL and UL. The Synchronization Channel (SCH) is transmitted on the center 72 subcarriers (six center PRBs) in the first and sixth subframe of the radio frame. For $BW_{Ch} > 1.4$ MHz systems more PRBs are provided below and above the six center PRBs in frequency domain. The BCH and SCH are only transmitted at the six center PRBs, so the overhead is reduced as the channel bandwidth $BW_{Ch}$ is increased.

A PRB is a logical resource spanning 12 subcarriers in frequency- and 7 or 6 symbols in time domain for normal and extended cyclic prefix, respectively. An OFDM symbol is
Table 2.1: Number of PRBs for different system bandwidths.

<table>
<thead>
<tr>
<th>$BW_{Ch}$ [MHz]</th>
<th>1.4</th>
<th>3</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{RB}$</td>
<td>6</td>
<td>15</td>
<td>25</td>
<td>50</td>
<td>75</td>
<td>100</td>
</tr>
</tbody>
</table>

the data transmitted during one symbol duration on the entire frequency bandwidth according to [8]. A radio resource spanning one symbol duration in time- and one OFDM sub-carrier in frequency domain is called Resource Element (RE) in the LTE standard. The grouping of REs to PRBs results from MAC scheduler constraints, where a pair of Virtual Resource Blocks (VRBs), which is mapped to a pair of PRBs (two slots in Figure 2.1), is the smallest resource unit that can be assigned as radio resource to a User Equipment (UE) [11]. There are two types of PRBs specified, localized and distributed [8]. With localized PRBs a pair of VRBs is mapped to a pair of PRBs with same index that are transmitted consecutive in time. With distributed PRBs a pair of VRBs is mapped to PRBs consecutive in time but not in the same frequency to increase frequency diversity. In this thesis only localized pairs are considered and denoted PRBs following publications on LTE like [12].

If multiple transmit antennas are present multiple spatial layers are used each representing the transmitted PRB on one of the antennas.

Figure 2.2a shows the format of a PRB not transmitting BCH or SCH data on DL. The first one to three symbols across all subcarriers form the Control Channel (CCH) region. Its size can be set individually for each subframe and is signaled in the Physical Control Format Indicator Channel (PCFICH) that is part of the CCH [8]. The PCFICH is transmitted on 16 REs of the first OFDM symbol. The remaining REs are reserved for cell-specific reference signals (RSs) for different antenna ports. The numbers in Figure 2.2a and 2.2b show the REs used to transmit RSs for the respective antenna port.
The six central PRBs of every fifth subframe transmit the SCH on two symbols as shown in Figure 2.2b. The SCH enables the initial synchronization of a UE to a cell and synchronization to surrounding cells to be able to assess their signal quality during handover decision. The BCH is transmitted once per radio frame with 10 ms period. It is transmitted on the six center PRBs on four symbols. It carries essential configuration and identification information of a cell to support attachment and
Table 2.2 summarizes the number of REs available in a PRB depending on the number of antennas operated and transmission of BCH or SCH if three symbols are assumed for the CCH region.

Figure 2.2c shows the structure of an UL PRB transmitting user data in the Physical Uplink Shared Channel (PUSCH). A configurable amount of PRBs at the edges of the frequency band is reserved exclusively for the Physical Uplink Control Channel (PUCCH) [8]. No user data is transmitted in these PRBs.

A UE transmits Demodulation Reference Signal (DM RS) in the fourth and eleventh OFDMA symbol of a PRB that serve to estimate the channel parameters at Enhanced Node-B (eNB). The last symbol that carries user data may occasionally be used to transmit Sounding Reference Signals (SRSs) to eNB to support scheduling decisions.

### 2.2 User Plane Data Transmission Procedure

Figure 2.3 shows the LTE protocol stack of the E-UTRAN comprising network elements UE and eNB. LTE-Uu is the radio interface between UE and eNB. L2 of the LTE E-UTRAN comprises three sublayers, PDCP [13], RLC [14] and MAC [15]. L2
relies on the services provided by the Physical (PHY) layer (L1) to receive and transmit data and to assess the signal quality of the radio channel. Above L2 in UE is the IP layer [16] and in eNB an IP relay function between the two protocol stacks shown. The right hand stack in eNB represents the Evolved Packet Core (EPC) using General Packet Radio Service (GPRS) Tunnelling Protocol User Plane (GTP-U) [17].

2.2.1 Packet Data Convergence Protocol (PDCP) Sublayer

The PDCP sublayer offers acknowledged / unacknowledged connection oriented data communication service to the IP layer. PDCP connections across interface LTE-Uu are called Radio Bearers. The PDCP sublayer performs Robust Header Compression (ROHC) [18] before transmission and decompression after reception supporting multiple protocols for compression of IP header and headers of higher layer protocols like Transmission Control Protocol (TCP) [19], User Datagram Protocol (UDP) [20], Real-Time Transport Protocol (RTP) [21], and Encapsulating Security Payload (ESP) [22] specified by the Internet Engineering Task Force (IETF). PDCP maintains
sequence numbers of Protocol Data Units (PDUs) to assure in-order delivery even under handover or connection interrupt. PDCP entities can be configured to either use short or long sequence numbers resulting in 7 or 12 bit Protocol Control Information (PCI). The total PCI length is 8 or 16 bit for short and long sequence numbers, respectively to assure byte aligned PDUs. The PDCP optionally supports a discard timer per PDU in the UL allowing to discard PDUs not scheduled for transmission in a given time duration.

2.2.2 Radio Link Control (RLC) Sublayer

Each PDCP entity communicates with one RLC entity and uses its services for user plane data transport. RLC entities may operate in Unacknowledged Mode (UM) or Acknowledged Mode (AM) depending on the configuration of the Radio Bearer performed at setup time. In-order data delivery is guaranteed in both modes but in UM lost PDUs are neither detected nor retransmitted. Sequence number has 5 or 10 bit for UM and 10 bit for AM. AM guarantees error correction by retransmission of PDUs received in error.

The MAC layer informs a RLC entity about the size of available transmission resources. The RLC layer applies segmentation and concatenation functions allowing to form RLC PDUs of one or multiple RLC Service Data Units (SDUs) and RLC SDU segments to fit the MAC PDU size. A buffer is used at the receiver to reassemble segmented RLC SDUs.

A RLC PDU includes 8 or 16 bit fixed PCI for short and long sequence numbers, respectively. If concatenation is used and more than one RLC SDU or RLC SDU segment is transmitted, additional 12 bit PCI are added per concatenated RLC SDU or RLC SDU segment. Additional 4 padding bit must be added if the PCI length is not byte aligned due to an even number of SDUs in the PDU.
2.2.3 Medium Access Control (MAC) Layer

While each radio bearer is handled by an individual PDCP and RLC entity, there is only one MAC entity per node managing channel access. The DL is directly managed by the MAC layer in eNB. UEs receive resource grants from the MAC layer of their serving eNB through the Physical Downlink Control Channel (PDCCH) [15].

The MAC layer can multiplex RLC-PDUs from multiple radio bearers belonging to the same UE into one MAC-PDU. MAC-PDUs containing only one RLC-PDU have an 8 bit PCI. The length of additional PCI fields depends on whether RLC user- or MAC control data is multiplexed. The MAC-PDU is called Transport Block (TB).

Figure 2.4 shows the layer specific PCI carried in a TB by means of an example. The control information is protocol related overhead. It is assumed the MAC scheduler has granted a TB containing 14 PRBs (PRB_0 to PRB_13) to be transmitted with MCS index 27 in a single antenna system. According to [11] and Figure 2.5, 8760 bit can be transmitted in this TB. This is more than the maximum PHY-PDU code block length of 6144 bit,
so the MAC-PDU is segmented into two blocks 6144 bit and 2616 bit long each coded separately and therefore having its own 24 bit Cyclic Redundancy Check (CRC) field. The MAC scheduler is assigning all resources to a single Radio Bearer and does not multiplex control traffic to it. Therefore the MAC PCI is 8 bit leaving space for 8752 bit of user data to be transmitted. This amount of data is requested from the RLC entity operating the Radio Bearer in question. This RLC entity has a less than 8752 bit long PDCP-PDU in its queue which fits the MAC-PDU. Another complete PDCP-PDU from its queue is concatenated to the RLC-SDU. The remaining space in the RLC-SDU is filled by a segment of a third PDCP-PDU leaving another segment of this PDCP-PDU behind. An 8 bit long fixed header and two 12 bit long extension headers are added to form the RLC-PDU. The total RLC-PCI length is 32 bit, so the header is byte aligned and no padding is necessary.

Each PDCP-PDU is formed from an IP-PDU IP header and
further headers as listed in Section 2.2.1 are compressed by ROHC and a 8 bit PDCP|PCI is added.

2.2.3.1 Link Adaptation (LA)

The LTE standard specifies 29 MCSs [11] the MAC layer may apply for transmission for a TB depending on the current channel quality of a PRB. The number of bits that can be carried in a TB only depends on the MCS index and the number of assigned PRBs as shown in Figure 2.5. MCSs in LTE do not have a fixed but a so called effective code rate. It is the ratio of the TB size in bit and the number of bits transmitted depending on the number of REs available for user data transmission. This ratio depends on the size of the CCH region and number of antennas transmitting RS in the DL see Section 2.1. According to Table 2.2 126 REs are available for DL user data transmission with one antenna and a three symbol wide CCH region. Two, four, or six bit can be transmitted in one RE depending on the modulation order of the selected MCS (QPSK, 16QAM or 64QAM). A TB spanning 10 PRBs transmitted at MCS index 9 using 16QAM modulation fits 1544 bit. It occupies 10 · 126 = 1260 REs, each fitting 4 bit. This results in 5040 transmitted coded bit and 1544/5040 = 0.31 effective code rate. PRBs containing periodic DL control information like the BCH can fit the same amount of bit but the effective code rate is significantly increased. In the UL optional SRS and control signaling multiplexed on the PUSCH region must be considered to calculate the effective code rate.

2.2.3.1.1 DL Link Adaptation (LA) A suitable MCS can be selected for the DL to a UE based on its Channel Quality Indicator (CQI) feedback derived from RS measurements by a UE. For each MCS a minimal SINR threshold is required to meet a requested Block Error Rate (BLER). For DL 15 reference MCSs are specified [11]. A UE reporting a certain reference MCS index thereby signals to eNB it can decode the corresponding
User Plane Data Transmission Procedure – 2.2

MCS with a BLER below 10%. The CQI value is either related to a subset or all PRBs. The channel quality is PRB specific as a result of frequency selective fading and changing interference patterns resulting from scheduling decisions in other cells.

SINR $\gamma_k$ for each PRB $k$ chosen for transmission may be different but only one MCS $m$ can be used for a TB utilizing the set $K$ of PRBs. It is therefore common practice to calculate an average SINR called effective SINR $\bar{\gamma}$, and map it to an appropriate MCS. Link level simulation results [23] show that Mutual Information Effective SINR Metric (MIESM) and Exponential Effective SINR Metric (EESM) [24] are well-suited averaging methods to calculate the effective SINR of a TB. EESM is calculated by Eq. (2.1).

$$I^{-1}(x) = -b \ln(x)$$ is the inverse of function $I(\gamma_k)$. $b$ depends on the MCS and is obtained by link level simulation as described in [25]. Calculating MIESM is more complex as visible from the IEEE 802.16m Evaluation Methodology Document [26].

$$\bar{\gamma} = I^{-1}\left(\frac{1}{|K|} \sum_{\forall k \in K} I(\gamma_k)\right)$$ (2.1)

$$I(\gamma_k) = e^{-\frac{\gamma_k}{b}}$$

2.2.3.1.2 UL Link Adaptation (LA) Single Channel Frequency Division Multiple Access (SC-FDMA) transmission used on UL requires another method to calculate effective SINR $\bar{\gamma}$ of a TB from estimated SINRs $\gamma_k$ of its PRBs. A receiver usually applies frequency domain equalization which influences effective SINR. A method to calculate the effective SINR with frequency domain Minimum Mean Square Error (MMSE) equalization is provided in [27]:

\[ \bar{\gamma} = I^{-1}\left(\frac{1}{|K|} \sum_{\forall k \in K} I(\gamma_k)\right) \]
\[ \tilde{\gamma} = \frac{x^2}{|K|x + x^2}, \]  
\[ x = \sum_{\forall k \in K} \frac{\gamma_k}{\gamma_k + 1} \]

The UL transmit power per PRB \( P_s \) of UE, according to \[11\]
is:

\[ P_s = \max(P_{\text{CMAX}}, \alpha \bar{h}_{c,s} + P_0), \]  
\[ P_0 = P_{0,\text{NOMINAL_PUSCH}} + P_{0,\text{UE_PUSCH}} \]

The maximum transmit power can be set by the UE but may not exceed 23 dBm \[9\]. The cell specific factor \( \alpha \) is multiplied with the estimated path loss \( \bar{h}_{c,s} \) from the serving eNB to UE. It is calculated as the difference between the transmit power of the cell specific RS and the filtered Reference Signal Received Power (RSRP) defined in \[28\]. The RSRP is linearly averaged in frequency domain within one subframe duration and exponential averaging with configurable coefficient is used to average over a longer time span as defined in \[29\]. The factor \( \alpha \) can be set to values between 0.5 and 1 in 0.1 wide steps and to 0. In the last case fixed transmit power \( P_0 \) is used. The power offset \( P_0 \) consists of a cell specific and UE specific part \( P_{0,\text{NOMINAL_PUSCH}} \) and \( P_{0,\text{UE_PUSCH}} \), respectively. The latter allows to compensate for impairments between the estimated path loss from DL measurements and the actual UL path loss and power boosting as a further degree of freedom for LA.

### 2.2.3.2 Hybrid Automatic Repeat Request (HARQ)

Two HARQ entities, one for down- one for UL, exist in eNB MAC layer for each UE. Each entity has exactly 8 Stop-and-Wait HARQ processes together supporting a Selective-Repeat
Automatic Repeat Request (ARQ) protocol with window size 8 [11].

**DL HARQ** is called asynchronous, since retransmissions can be scheduled in any subframe after a Negative Acknowledge-ment (NACK) was received on the Physical Hybrid-ARQ Indicator Channel (PHICH) and is called adaptive because arbitrary **PRBs** can be allocated and modulation can be changed regardless of the parameters of the initial transmission [30]. Only the amount of uncoded bit must remain the same as in the initial transmission, so the code rate cannot be adapted.

**UL HARQ** procedure is synchronous and non-adaptive in general. Unsuccessful decoding of a **TB** received is indicated by a NACK transmitted by eNB four subframes later signaling to the **UE** that retransmission of the data is required using the same **MCS** as before on the same **PRBs** four subframes later (eight subframes later than the previous transmission). The **HARQ** process number is determined by the **UE** from the current subframe number modulo number of **HARQ** processes. Instead of transmitting NACK in the PHICH, eNB may transmit an UL scheduling grant in PDCCH for the **UE** with disabled New Data Indicator (NDI) field. This way adaptive **HARQ** to some extend can be used in the UL. The grant includes information which **PRBs** should be used but the **MCS** must remain the same as for the initial transmissions for all retransmissions [11, 30]. To some extent asynchronous **HARQ** behavior is achieved if eNB sends an Acknowledgement (ACK) as reply but leaves NDI field disabled next time a grant handled by that particular **HARQ** process is transmitted on PDCCH. A **UE** would not perform a retransmission eight subframes later since it assumes the last transmission was successfully received according to the ACK. A resource grant received 16 subframes after the last transmission with disabled NDI field will force the **HARQ** process to perform a retransmission since it is not allowed to transmit new data according to the NDI field. The **UL HARQ** procedure therefore allows some adaptivity by shifting the allocated resources in frequency domain and some asyn-
chronous operation by shifting retransmissions to later subframes handled by the same HARQ process. This is especially required with SPS, see Section 2.3.2, to assure HARQ retransmissions are not scheduled on resources reserved for persistent transmissions of other UEs.

2.3 Scheduling

In LTE DL and UL resource grants for each UE are transmitted on the PDCCH. UL grants are necessary to inform UEs on which resources they should transmit data. DL grants prevent a UE from having to search all resources for data intended for it. Resource grants are described by Downlink Control Information (DCI) [31] e.g. information which PRBs form the TB, which MCS is used, and the HARQ protocol applied, see Section 2.2.3.2. Further, information to support Multiple Input Multiple Output (MIMO) transmission is included, see the DCI formats described in [30]. The DCI is protected by a 16 bit CRC field scrambled with the identifier of the intended UE.

2.3.1 Signaling

Adaptive Modulation and Coding (AMC) is used to dynamically adjust the number of REs required to transmit the DCI in the PDCCH [8]. For signaling Quadrature Phase-Shift Keying (QPSK) is always used. The code rate is adjusted dynamically. A DCI is transmitted on one, two, four or eight Control Channel Elements (CCEs). A CCE comprises 36 REs (each 2 bit in QPSK) in summary 72 bit. The length of DCI in bit depends on the DCI format selected and the system bandwidth. LTE Release-8 at 5 MHz system bandwidth without MIMO needs 43 bit for DCI-Format 0 and 39 bit for DCI-Format 1 for DL and UL grant, respectively. The resulting effective code rate in DL is $\frac{43}{72} \approx \frac{3}{5}$ if one CCE is used for transmission and $\frac{3}{10}, \frac{3}{20}, \frac{3}{40}$ for two, four, and eight CCEs, respectively.
The total number of CCEs that can be transmitted by eNB depends on the number of REs available for PDCCH transmission. According to Section 2.1, CCH region is one, two or three OFDM symbols in each subframe. If three symbols are used, as desirable for VoIP service if there are many UEs in the system, a Resource Block (RB) provides 32 REs in the CCH region if eNB is configured to use one or two antenna ports and 28 REs if four antenna ports are used (see Figure 2.2a). In a 5 MHz system using 25 PRBs, 25 \cdot 32 = 800 REs are available when using one or two transmit antennas. From these 800 REs, 16 are used to transmit PCFICH and 12 are required for each PHICH group. HARQ feedback for up to eight UEs is multiplexed in a PHICH group using orthogonal codes \cite{8}. The number \(N_{\text{group PHICH}}\) of PHICH groups depends on the number \(N_{\text{DL RB}}\) of PRBs corresponding to the selected system bandwidth and the parameter \(N_g\) configured by Radio Resource Control (RRC) \cite{29}. It is calculated from \cite{8} as:

\[
N_{\text{group PHICH}} = \left\lceil N_g \frac{N_{\text{DL RB}}}{8} \right\rceil \tag{2.4}
\]

Figure 2.6 shows the number of available CCEs for PDCCH transmission versus the number of PHICHs for 5 MHz and 10 MHz system bandwidth with one or two antenna ports at eNB. The four different values for each, the 10 MHz and 5 MHz configuration, correspond to the different settings for \(N_g \in \{1/6, 1/2, 1, 2\}\) and the number of PHICHs is obtained as \(8N_{\text{group PHICH}}\) since eight orthogonal codes are available for transmission. No more than 23 and 46\(^1\) transmissions requiring PHICHs can be performed with 5 MHz and 10 MHz, respectively if all TBs only use one PRB. With \(N_g = 1\) 32 and 56 PHICHs are available to provide HARQ feedback for this worst case with 5 MHz and 10 MHz system bandwidth, respectively. The resulting number of 20 CCEs for 5 MHz system bandwidth

\(^1\)Two or four PRBs are reserved for the PUCCH for 5 MHz and 10 MHz, respectively according to baseline assumptions \cite{32}. 

\[\]
can become a capacity limiting bottleneck if many users must be scheduled.

The CCEs are for both, UL and DL scheduling grants, without limit on the distribution among transmission directions. The PHICHs are only used to transmit (N)ACKs for UL transmissions.

### 2.3.2 Semi-Persistent Scheduling

With Semi-Persistent Scheduling (SPS) a resource grant is only signaled once and remains valid for the current subframe and future subframes with a periodicity specified. The period is configured by the RRC layer according to [29]. SPS reservation saves signaling resources since a resource grant is reserved once until it is revoked instead of reserving each PRB individually. A Time Division Multiple Access (TDMA) channel is established thereby, for circuit switched communication instead of packet switched. A UE having a semi-persistent reservation may always also receive dynamic resource grants. This is why the method is called semi-persistent scheduling. If a UE receives a dynamic resource grant in a subframe where it already
has a persistent reservation it will ignore the persistent reservation and comply to the dynamic resource assignment. The distinction between a dynamic and a semi-persistent resource grant is through the identifier of the UE (which is scrambled with the CRC field carried in DCI) in the resource grant. A UE receives a second identifier, the so-called semi-persistent Cell Radio Network Temporary Identity (C-RNTI) if SPS is enabled for it. If the CRC field for a received DCI can be verified when descrambled with the semi-persistent C-RNTI the grant is considered to be periodic with previously configured period.

For both, DL and UL, the reservation can be explicitly released by sending according DCI having its CRC scrambled with the semi-persistent C-RNTI of the addressed UE. For the UL also implicit release is possible if a UE has not used its reservation for a number of times in a row, configured by the RRC layer. A semi-persistent reservation can be changed with regard to the used PRBs or the MCS by sending a new semi-persistent DCI while a persistent reservation is active. The active reservation is then changed [15].

DL HARQ retransmission is performed asynchronously as described in Section 2.2.3.2. In the UL the UE can retransmit lost TBs using the next semi-persistent allocation. This usually causes a backlog at the sender of data originally designated for these reserved resources. Therefore PDCCH signaling can be used to schedule an asynchronous retransmission as described in Section 2.2.3.2.

PDCCH resources are only consumed to setup or change a semi-persistent reservation, to release it in the DL, to explicitly release it in the UL and to schedule asynchronous HARQ retransmissions. Further signaling is required if some data is scheduled dynamically, for example the Silence Indicator (SID)-PDUs of a VoIP call (see Section 4.4.2).
Related Work

Content

3.1 Voice Call Performance Evaluation .................................. 25
   3.1.1 VoIP Capacity State-of-the-art .................................. 28
   3.1.2 Scheduling Algorithms for VoIP ................................. 31
   3.1.3 LA Power- and Admission Control .............................. 32
   3.1.4 Analytical Evaluation of VoIP Capacity ....................... 33

3.2 E-UTRAN VoIP Performance Evaluation ............................. 35
   3.2.1 Activities of the 3GPP Consortium Prior to IMT-A Evaluation ........................................ 35
   3.2.2 LTE-A IMT-A Evaluation Results ............................... 37

3.1 Voice Call Performance Evaluation

Both major standardization bodies for mobile radio networks, 3GPP and 3rd Generation Partnership Project 2 (3GPP2), conducted effort towards an all-IP based network infrastructure to be deployed at the end of the 20th century [33] since following benefit are expected [34]:

- Enabling simultaneous multimedia services like voice, data and audio
- Seamless services, regardless of the actually used access technology
- Synergies with other ongoing and future IP development
- Reduced costs for operation, administration and management
Chapter 3 – Related Work

- Integration of IP terminals in the network
- Reduced cost by using cheap packet transport rather than reserved communication resources

The all-IP effort of 3GPP was based on the existing GPRS specification using GPRS Tunnelling Protocol (GTP) [17] for mobility management. Key effort was therefore put into multi-media applications and resulted in standardization of the IMS [35].

Mobility support for packet switched services on mobile terminals was an open issue in 3GPP2 protocols at that time [33], so reusing IETF standards for that purpose was the first topic towards IP integration. Mobile-IP [36] was selected for mobility management and the concept of Virtual Private Network (VPN) [37] was adopted to increase security [33]. Support for multimedia applications was later added [38] to the 3GPP2 protocol stack based on IMS [35]. A comparison between the two IMS standards is provided in [39]. Further information on the evolution of IMS and VoIP in LTE is found in [40]. Since LTE and LTE-A of 3GPP became the dominant 3.5G and 4G systems, respectively, 3GPP IMS is the dominant system now, worldwide.

A broad overview on VoIP in general is provided in [41]. It includes a description of standard voice codecs and names factors contributing to the end-to-end delay. Service differentiation by the scheduler controlling transmission resource allocation where VoIP traffic gets priority over less delay sensitive services is identified as key enabling technology to achieve high VoIP capacity.

The first proposal of packet voice in cellular is from 1993 [42] where the capacity of GPRS was evaluated. Packets are scheduled OTT and silence descriptors are also transmitted. In [43] for better efficiency of GPRS based packet voice, persistent scheduling of VoIP bearers is introduced, namely to keep the temporary block flow of GPRS alive for voice bearers instead of releasing it during silence periods of the speech modem.
In [44] it is shown that VoIP over 3rd Generation (3G) Universal Mobile Telecommunications System (UMTS) Release 99 [45] is generally feasible. VoIP performance over 3G UMTS using High Speed Packet Access (HSPA)\(^1\) [48] is evaluated in [49]. Key challenges and enablers for the service are described. Low delay required for VoIP is motivated by the 2 ms TTI of HSPA introduced in Release 7 [48] instead of 10 ms used in UMTS. ROHC on the radio interface is used since GPRS to reduce packet header overhead. Release 7 introduced transport block sizes fitting small VoIP packets avoiding padding overhead. Round Robin (RR) vs. Proportional Fair (PF) scheduling for VoIP over HSPA is evaluated in [50] showing that RR outperforms PF with regard to delay resulting in higher VoIP capacity. VoIP capacity is further improved if multiple VoIP packets are transmitted concatenated in one MAC layer transport block and by introducing admission control to avoid overload of the system as shown for HSPA in [51].

LTE and its all-IP concept introduced VoIP as the only option for voice call services. Special considerations were required in protocol design to improve VoIP capacity compared to what was achieved with HSPA. Circuit switched system GSM allowed multiplexing eight users on one 200 kHz wide frequency channel. This accounts for 40 users per MHz bandwidth if guard bands and control channels are discarded. This is doubled if the GSM half-rate codec [52] is used. Accounting for the reuse factor larger than one reduces GSM VoIP capacity to 3 – 13 users / cell / MHz [53]. HSPA could not even reach 20 users per MHz in the UL limiting overall VoIP capacity [54, 55].

To improve VoIP capacity LTE-A was specified as an IMT-A family member. Evaluation of LTE-A VoIP capacity is de-

---

\(^1\)High Speed Downlink Packet Access (HSDPA) was introduced in 3GPP Release 5 [46] and High Speed Uplink Packet Access (HSUPA) in Release 6 [47]. Release 7 [48] combines both and adds additional features, including 2 ms long Transmission Time Intervals (TTIs), and is usually referred to as HSPA but often the term HSDPA is used as synonym in literature to address Release 7 features.
In the following related work from scientific publications, some originating from institutions involved in evaluation of IMT-A compatibility of LTE-A is analyzed.

### 3.1.1 VoIP Capacity State-of-the-art

The main problem with VoIP over radio is that each packet needs individual resource reservation instead of a reservation of resources for a voice bearer to carry a whole session. In [56] a simple analytical model to calculate VoIP capacity in LTE based on mean values is provided. A two state traffic model is assumed with voice packets generated every 20 ms during talk spurts of a user and SID packets send every 160 ms during idle periods. The mean number of resources required for voice packets, SID packets and packet retransmissions is summarized. An upper bound capacity of 141 users / MHz is calculated assuming all packets require just one PRB and a 20 % packet error ratio. Capacity shrinks linearly to 70 and 47 users / MHz if two and three PRB are required per packet, respectively. This is a weak upper bound since VoIP delay constraints are neglected. Simulation results [56] considering 40 ms delay limit of VoIP packets and 2 PRBs per packet show significantly lower capacity. Users are considered not satisfied in [56] if packet loss exceeds 2 % and VoIP capacity is defined by the offered load where more than 5 % users are not satisfied following [57].

Fully dynamic, persistent and semi-persistent scheduling was used to obtain the results. Fully dynamic scheduling requires signaling to reserve resources for each packet every subframe but also offers the highest flexibility at the cost of signaling overhead. Persistent scheduling was proposed for LTE standardization in order to reduce the signaling overhead. There the same resources are reserved every 20 ms for voice packets during a talk spurt and every 160 ms for SID packets during idle periods. The problem is the lack of reserved re-
Voice Call Performance Evaluation – 3.1

resources for retransmissions. This was solved by persistently reserving additional resources for potential HARQ packets. Persistent scheduling was discarded in LTE standardization due to inefficient handling of HARQ packets and replaced by SPS described in Section 2.3.2 only persistently reserving resources for voice packets every 20 ms and scheduling HARQ and SID packets dynamically.

Results for persistent scheduling in [56] show VoIP capacity is reached at around 40 users / cell / MHz. SPS increases the capacity to around 64 users / cell / MHz. Results for fully dynamic scheduling of VoIP and SID packets show the same capacity as SPS. Possible capacity limits of control channels under dynamic scheduling are not considered.

The influence of limited control channel capacity is evaluated in detail in [58]. The mean number of users to be scheduled in each subframe corresponding to the number of required control channels is calculated similar to [56]. VoIP capacity with the same constraint considered in [56] was evaluated by simulation tolerating 50 ms maximum VoIP packet delay. The resulting VoIP capacity is 44 users / cell / MHz with SPS assuming more than six control channels. Capacity is increased to 46 users / cell / MHz. With dynamic scheduling and an infinite number of control channels. Capacity is reduced to 43, 39, 35, and 32 users / cell / MHz if the number of control channels is limited to nine, eight, seven, and six, respectively. Dynamic VoIP packet scheduling instead of SPS enables channel aware scheduling allowing users to transmit on frequency resources with higher current SINR. The channel model used for simulation resembles 3 km/h user velocity. Capacity gains from dynamic scheduling found in [58] are likely to be lower according to [59].

The maximum amount and resolution in time and frequency domain of channel feedback information through CQI standardized in LTE was modeled too optimistically in initial VoIP capacity evaluations [60]. With speed higher than 3 km/h CQI information tends to be quickly outdated by the time it is used
Chapter 3 – Related Work

to support a scheduling decision. Increasing user speed from 3 km/h to 30 km/h reduces VoIP capacity by 42% according to [60]. In [59, 60] VoIP packet bundling, namely transmission of multiple packets of the same user in the same subframe, is proposed. Thereby the number of control channels required is reduced increasing the VoIP capacity under dynamic scheduling.

Bitmaps were proposed to reduce signaling overhead by coding control information [61]. Persistent scheduling of groups of users is proposed in [62] for the Worldwide Interoperability for Microwave Access (WiMAX) system standard IEEE 802.16m [63] to reduce signaling overhead and improve VoIP capacity. There, rather than addressing only one user when initiating, revoking or changing a persistent reservation, a subset of users from a defined group is addressed. Bitmaps are used [61] to indicate which users belonging to a group are being addressed. A similar concept for LTE is evaluated in [64]. Group scheduling is less flexible than SPS and requires more control channels than SPS. Limiting access of users in a group to certain resources reduces trunking gain and VoIP capacity [64].

In [65] a concept is presented for IEEE 802.16 systems to deal with changing channel conditions and fragmentation of resources in a subframe from released persistent reservations. Few bits are used to signal users how to relocate their resources in a subframe and change their MCS relative to the currently used one. The concept is not applicable to LTE that uses a few bits for signaling and scheduling only. In [66] it is shown that channel prediction errors significantly reduce VoIP capacity, especially in the uplink. Relays introduce higher delays due to an additional hop but also reduce interference variance and therefore channel prediction errors increasing VoIP capacity from 25 to 45 users / cell / MHz.

The gain in VoIP capacity discussed so far cannot be achieved in LTE with many VoIP users due to control channel capacity limitations. Packet bundling may increase VoIP capacity and is discussed in [67]. According to [68, 60] VoIP capacity is sub-
stantially increased if delay higher than 50 ms is tolerated. Results for TDD mode are provided in [69, 70] for LTE and [71] for IEEE 802.16m. VoIP capacity of system 3GPP2 Evolution-Data Optimized (EV-DO) Revision A [72] is provided in [73] from a sophisticated analytical model providing the tail delay distribution to enable calculation of the percentage of satisfied users for a given VoIP packet delay bound.

3.1.2 Scheduling Algorithms for VoIP

VoIP capacity is affected by other traffic flows competing for the same radio resources. The work presented in [67, 74] describes a DL scheduling algorithm for SPS of VoIP packets without LA. Two PRBs are assigned per VoIP packet prioritizing those waiting in the transmit buffer close to 50 ms. Other packets and pending retransmissions are scheduled if sufficient resources are available. Preferably VoIP packets are scheduled to the same PRBs in frequency domain used during the preceding talk spurt. Packets exceeding 50 ms delay in the transmit buffer are discarded. [75] evaluates a scheduler where VoIP packets exceeding 40 ms delay are scheduled dynamically using LA.

[76] studies scheduling based on Earliest Deadline First (EDF) algorithm assigning radio resources preferably to packets close to their maximum tolerable delay. Hybrid delay and channel aware scheduling is presented in [76]. Algorithms like the following require dynamic scheduling and are not suitable for a large number of VoIP users due to limited control channel capacity.

The scheduling algorithms in [77] and [78] assign highest priority to the VoIP user with maximal product of queue length and instantaneous SINR. Thereby the buffer is emptied preferably under good channel condition. A limited number of VoIP users is prioritized over non-VoIP users. The limit is dynamically increased with increased number of VoIP packets missing the 50 ms delay constraint and reduced otherwise. The al-
algorithm of [79] is similar but in addition accounts for the packet loss rate.

In [80, 81] time- and frequency domain scheduling of VoIP packets is strictly separated. In time domain users are ranked by a weighted ratio of required data rate and exponential moving average received data rate. The weight function represents delay sensitivity of a traffic class. PF scheduling [82] is used in frequency domain. Elastic traffic has constant weight 1. Using an exponential weight function $2^{d-\alpha}$ for VoIP, where $d$ is the Head of Line (HOL) packet delay, is found there resulting in best throughput performance in a mixed traffic scenario. VoIP packets exceeding delay $\alpha$ have higher priority than best effort traffic. In [83] a similar weight factor is used that increases for packets 10 ms close to their due date and decreases if packets are 5 ms close to their due date accounting for very late packets will likely not make it in time.

3.1.3 LA, Power- and Admission Control

[84] evaluates an LA algorithm suitable for persistent scheduling for IEEE 802.16m systems. The algorithm estimates the statistics of the SINR distribution from a Nakagami-m distribution [85] and related packet error probability distribution taking HARQ into account. Control overhead for signaling MCS change is accounted for.

MIMO transmission increases SINR and may allow more users with higher data rates. Capacity of MIMO systems in general is evaluated in [86] for VoIP in [87]. It is found multiple antennas do not increase VoIP capacity in high SINR scenarios where most VoIP packets need one PRB for transmission.

In [88] a distributed algorithm for calculating optimal transmit power in different regions of the frequency domain is introduced, thereby establishing SFR [89]. The advantage for real time traffic is proven by comparing queuing delay distribution results from simulation for simple and optimal power control.

Transmit power is often the limiting factor for VoIP capacity
Cell-edge users tend to need more PRBs in frequency domain compared to other users owing to a more robust MCS required to meet packet error constraints. Fragmenting VoIP packets to be transmitted in multiple subframes, each fragment using less resources in frequency domain is proposed and evaluated in [90].

Admission control is required in packet switched systems to avoid overload. Admission control in [91] is studied by means of a Markov Model where exclusive resources are dynamically assigned to both VoIP and non-VoIP services such that VoIP users are satisfied.

In [92] channel aware scheduling, such as PF [82] is applied known to provide a higher data rate by preferably transmitting data when channel quality is good. The actual perceived channel quality is calculated using order statistics. Admission control is based on percentiles instead of mean data rate.

### 3.1.4 Analytical Evaluation of VoIP Capacity

In Section [3.1.1] a simple upper bound capacity model based on mean values for the amount of voice-, comfort noise and retransmission packets [56] was discussed. In [93] VoIP capacity is calculated much simpler only taking the total numbers of available resource units and number of resource units required to transmit a VoIP packet into account. This way VoIP capacity is compared of systems HSPA Release 6 [47], IEEE 802.16-2004 [94], IEEE 802.11a [95], and EV-DO [72] Rev. A. explicitly calculating the related overhead. For system IEEE 802.11a mean values for backoff time and data rate obtained from simulation are used as basis for the comparison. The results show 15, 9, 10, and 3 users per MHz bandwidth can be served in systems derived from standards IEEE 802.16-2004, HSPA Release 6, EV-DO Rev. A, and IEEE 802.11a, respectively. Low performance of HSPA and EV-DO results from the large minimal transport block size compared to the size of a VoIP packet.

A more sophisticated analytical model for VoIP capacity
evaluation of OFDMA systems is presented in [96]. Here the actual SINR value at each point in the service area is calculated to derive the according MCS and related resource requirement of VoIP packets. It is assumed that all VoIP users are served in each 20 ms interval. The total transmit power constraint is relaxed allowing to exceed the power limit in certain subframes as long as the average transmit power over 20 ms is below the power limit. VoIP capacity for reuse-1 is reached at 16 and 28 users / cell / MHz for Single Input Single Output (SISO) and 2×2 MIMO transmission, respectively in a scenario with 1000 m and 1732 m Inter-site distance (ISD). Capacity is reduced to 14 (SISO) and 28 (MIMO) users / cell / MHz is the scenario with 500 m ISD.

An Engset model [97] is used in [98] to evaluate VoIP capacity of EV-DO. The mean number of resources required to transmit comfort noise is calculated and reserved not to be available for VoIP traffic by correspondingly reducing the number of servers in the model. Persistent scheduling is used, reserving resources for the initial transmission and possibly following retransmissions. It is assumed any target Packet Error Rate (PER) can be achieved by increasing transmit power with the limits of the total transmit power. An increased interference in a multi-cell scenario is not considered. From the steady state distribution of the number of active users mean queue length, waiting time and probability of delay exceeding 20 ms versus number of users are calculated. A tandem M/M/n queuing network is used in [99] to derive the delay distribution for VoIP traffic in EV-DO. The first queue models establishing a reservation while the second one models user data transmission. Any target packet error rate is assumed achievable by power control with the limit of the total available transmit power. The probability to not achieve the target packet error rate due to total transmit power limit is calculated without considering related interference.

Analytical models for evaluation of VoIP performance for system IEEE 802.16e are provided in [100] and [101] under dy-
dynamic scheduling including modeling of limited control channel resources. The superposition of VoIP traffic from all users is modeled with a Markov-Modulated Poisson Process (MMPP) \[102\]. The distribution of MCSs is input to the model and resources required for user and control channel transmissions are derived together with the number of users that can be scheduled. From the ratio of mean scheduled users and mean arriving VoIP packets per frame system utilization and throughput is calculated.

3.2 E-UTRAN VoIP Performance Evaluation

3.2.1 Activities of the 3GPP Consortium Prior to IMT-A Evaluation

Requirements on the evolved Universal Terrestrial Radio Access Network (UTRAN) \[103\] established the following goals for voice services in LTE and LTE-A systems:

Voice and other real-time services supported in the CS [Circuit Switched] domain in R6 [Release 6] shall be supported by E-UTRAN via the PS [Packet Switched] domain with at least equal quality as supported by UTRAN.

Since high signaling overhead for voice packet scheduling is identified as a significant performance limiting factor multiple UEs addressed in one control message called group scheduling \[104\], and persistent scheduling \[105\] based on periodic reservation of resources are proposed for better throughput performance.

Performance evaluation methodologies for different applications including VoIP are described in \[57\] and were partly adapted by ITU-R for its IMT-A Evaluation Methodology \[7\] described in Section 4.4.2. One difference between \[57\] and \[7\] is that VoIP capacity is defined at 5% instead of 2% non satisfied users. Deployment scenarios \[106\] (called Cases) used for UTRAN are still valid for E-UTRAN Release 8 and beyond. Case 1 scenario resembles the Urban Macro (UMa), Case
Table 3.1: Preliminary DL 3GPP LTE VoIP Capacity Results from May 2007. Case 1 (500 m ISD) and Case 3 (1732 m ISD) Deployment Scenario [106]

<table>
<thead>
<tr>
<th>Company</th>
<th>Scheduling</th>
<th>Cap. [UEs/Cell/MHz]</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Case 1</td>
<td>Case 3</td>
</tr>
<tr>
<td>Ericsson</td>
<td>Dynamic</td>
<td>81</td>
<td>62</td>
</tr>
<tr>
<td>Motorola</td>
<td>Semi-pers.</td>
<td>57</td>
<td>44</td>
</tr>
<tr>
<td>NEC</td>
<td>Dynamic</td>
<td>&gt;70²</td>
<td>&gt;70</td>
</tr>
<tr>
<td>Nokia/NSN</td>
<td>Dynamic</td>
<td>56</td>
<td>55</td>
</tr>
<tr>
<td>Nokia/NSN</td>
<td>Semi-pers.</td>
<td>71</td>
<td>64</td>
</tr>
<tr>
<td>Qualcomm</td>
<td>Semi-pers.</td>
<td>47</td>
<td>45</td>
</tr>
<tr>
<td>Samsung</td>
<td>Semi-pers.</td>
<td>54</td>
<td>46</td>
</tr>
</tbody>
</table>

The Rural Macro (RMA) scenario from the IMT-A Evaluation Methodology [7] discussed in more detail in Section 4.1. All scenarios comprise 19 base station sites with 3 sectors each, respectively 57 cells in total.

Quite different LTE VoIP performance results for Case 1 and Case 3 scenario under dynamic, group and (semi-)persistent scheduling were found by manufacturers according to Tables 3.1 and 3.2 [107]. Reference with details on the respective simulation model used is provided in the last column. Persistent scheduling in UL means HARQ packets are not scheduled dynamically.

Capacity for the urban scenario (Case 1) is higher than of the rural scenario (Case 3) owing to the properties of the respective channel models specified in [106]. The capacity difference is much larger in the UL. Some companies have found more than 5% outage in Case 3 scenario for any number of users shown.

²The ratio of unsatisfied users did not exceed 5% with 70 users / cell / MHz. Scenarios with more users were not evaluated.
Table 3.2: Preliminary UL 3GPP LTE VoIP Capacity Results from May 2007. Case 1 (500 m ISD) and Case 3 (1732 m ISD) Deployment Scenario [106]

<table>
<thead>
<tr>
<th>Company</th>
<th>Scheduling</th>
<th>Cap. [UEs/Cell/MHz]</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Case 1</td>
<td>Case 3</td>
</tr>
<tr>
<td>Alcatel-Lucent</td>
<td>Persistent</td>
<td>30</td>
<td>-</td>
</tr>
<tr>
<td>Ericsson</td>
<td>Dynamic</td>
<td>68</td>
<td>31</td>
</tr>
<tr>
<td>Motorola</td>
<td>Group</td>
<td>45</td>
<td>-</td>
</tr>
<tr>
<td>Motorola</td>
<td>Semi-pers.</td>
<td>36</td>
<td>-</td>
</tr>
<tr>
<td>NEC</td>
<td>Dynamic</td>
<td>64</td>
<td>26</td>
</tr>
<tr>
<td>Nokia/NSN</td>
<td>Dynamic</td>
<td>47</td>
<td>25</td>
</tr>
<tr>
<td>Nokia/NSN</td>
<td>Semi-pers.</td>
<td>44</td>
<td>24</td>
</tr>
<tr>
<td>Qualcomm</td>
<td>Dynamic</td>
<td>40</td>
<td>24</td>
</tr>
<tr>
<td>Qualcomm</td>
<td>Persistent</td>
<td>200</td>
<td>16</td>
</tr>
<tr>
<td>Samsung</td>
<td>Persistent</td>
<td>44</td>
<td>16</td>
</tr>
</tbody>
</table>

in Table 3.2 as zero VoIP capacity.

3.2.2 LTE-A IMT-A Evaluation Results

The LTE-A IMT-A evaluation result by 3GPP consortium [115] submitted to ITU-R contains VoIP capacity of five companies obtained by simulation according to IMT-A evaluation guidelines [7] and 3GPP modeling assumptions for LTE-A evaluation [32], see Chapter 4. The results are shown in Table 3.3.

Two antennas at the mobile terminal and four antennas at the base station are assumed. Two base station antenna configurations are evaluated. Antenna elements at the base station are $4\lambda$ and $0.5\lambda$ apart in configuration $A$ and $C$, respectively. Two companies contributed results for each antenna configuration.
Table 3.3: VoIP capacity of IMT-A scenarios UMa, UMi, and RMa as reported to ITU by 3GPP member companies.

<table>
<thead>
<tr>
<th>Antenna Configuration</th>
<th>Company</th>
<th>Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>UMa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DL</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>61</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>80</td>
</tr>
</tbody>
</table>

Company 1 provided results for both configurations. Table 3.3 provides VoIP capacity results for DL and UL in scenarios UMa, UMi, and RMa. VoIP capacity is reached at the smaller of the values shown for UL and DL in bold, Table 3.3. Most companies agree that the DL limits capacity in UMa scenario. For the other two scenarios there is no clear picture since some companies say UL, some DL to be the bottleneck.

The higher mean VoIP capacity shown in Figure 3.1a and 3.1b was reported to ITU-R as VoIP capacity of LTE-A, namely 69, 80, and 94 users / cell / MHz for the UMa, UMi, and RMa scenario, respectively. Most results in this thesis are obtained with single receive and transmit antenna due to computational complexity of simulation and cannot be directly compared to 3GPP findings. Section 7.2 shows for InH scenario that having two receive antennas and small scale channel fading at 3 km/h results in almost same VoIP capacity as with one receive antenna and no small scale channel fading. Further Section 7.3.9 shows that adding another receive antenna and assuming theoretically maximum SINR gain 3 dB [116] results in 68 users / cell /
**Figure 3.1:** VoIP capacity reported to ITU by 3GPP consortium and independent evaluators.

MHz in the UMa scenario which is almost identical to 69 users / cell / MHz found by 3GPP.

VoIP capacity has significantly increased compared to intermediate results presented in Section 3.2.1. The bottleneck for the intermediate results was always the UL. The VoIP capacity for the UMa-like Case 1 scenario with SPS is 36 users /
cell / MHz (Motorola), 44 users / cell / MHz (Nokia/NSN), 40 users / cell / MHz (Qualcomm), and 44 users / cell / MHz (Samsung) resulting in 41 users / cell / MHz mean VoIP capacity. Final results for UMa VoIP capacity 69 users / cell / MHz is 69% higher than the intermediate one. Results cannot be directly compared due to slightly different modeling assumptions between the 3GPP and IMT-A scenarios. A comparison of 3GPP and IMT-A large-scale channel models in [117] shows slight differences in the capacity achieved resulting from different assumptions made for base station antenna height.

Besides the five companies discussed above five organizations have submitted VoIP capacity results for LTE-A to ITU-R:

- Alliance for Telecommunications Industry Solutions (ATIS) Wireless Technologies and Systems Committee (WTSC) [118]
- Canadian evaluation group [119]
- Chinese evaluation group [120]
- Telecommunications Industry Association (TIA) committee TR-45 [121]
- Telecommunications Technology Association of Korea (TTA) Project Group 707 [122]
- European Union project Wireless World Initiative New Radio Plus (WINNER+) [123]

The results together with scenario specific means are shown in Figure 3.1c together with mean VoIP capacity provided by 3GPP taken from Figure 3.1a and 3.1b. Most of VoIP capacity results of evaluation groups confirm the ones from 3GPP. Results from TTA PG 707 [122] derivate most, especially in UMi and RMa scenarios.

Besides TTA the Canada group obtained lower capacity for the UMi scenario than 3GPP. The respective document [119]
does not explain reasons for that. The largest derivation from the mean for UMa and UMi scenario is 2.0 % and 7.7 %, respectively.

All groups except for TTA obtained similar VoIP capacity for the RMa scenario. The highest deviation from the 94 users / cell / MHz VoIP capacity reported by 3GPP are 99 users / cell / MHz (+5.3 %) obtained by the Chinese evaluation group. The largest derivation from the mean is 8.4 %.

Minimum VoIP capacity requirement for IMT-A 40 users / cell / MHz for scenario UMa and UMi and 30 users / cell / MHz for scenario RMa indicated by horizontal lines in Figure 3.1 is significantly exceeded in any case.

Apparently LTE-A VoIP capacity results as shown in Figure 3.1c vary although the same scenarios and channel models were assumed. The reason for this in my view are different radio resource control algorithms and schedulers assumed in the respective studies. It is therefore a main goal of this thesis to make transparent the influence of these algorithms and strategies.
The IMT-A Evaluation Methodology document \[7\] defines four test environments in which candidate radio interface technologies (RITs) have to prove they meet the performance specified by the ITU-R \[124\]. System and link level simulation, mathematical analysis, and inspection are specified in \[7\] to be used for performance evaluation, each method for a specific task. The cell spectral efficiency (CSE), cell edge user spectral efficiency (CEUSE) and VoIP capacity are to be evaluated by system level simulation.

4.1 Scenarios

IMT-A candidate system performance is evaluated in the test environments Indoor, Microcellular, Base coverage urban, and
High speed. Each environment has a specific geometric deployment scenario namely Indoor Hotspot (InH), Urban Micro (UMi), Urban Macro (UMa), and Rural Macro (RMa), defined by the cell size and ISD for the last three as shown in Figure 4.1a. The InH scenario is formed by a rectangular floor spanning 120 m x 50 m with two Base Station (BS) sites as shown in Figure 4.2. The BSs in the cellular scenarios serves three 120° sectors with identical antenna patterns. All UEs and the BSs in the InH scenario use omnidirectional antennas. Antenna patterns are modeled by azimuth and elevation angle dependent attenuation functions, respectively. For each scenario carrier frequency, transmission bandwidth, maximum transmission power, transceiver height, and number of antenna elements is specified as shown in Table 4.1. UEs are deployed in the scenario area at uniformly distributed random positions but have to keep minimum scenario dependent distance from BSs. Position of a UE is fixed throughout a simulation run. Changing channel conditions due to mobility are modeled by the small scale channel model described in Section 4.2.2. UEs can be either located outdoor, indoor or in vehicles. UEs are assumed located in vehicles in UMa and RMa scenarios. For the UMi scenario 50% of UEs are deployed outdoors and 50% indoors. All UEs are located indoors for the InH scenario.

4.2 Channel Model

The channel model specified in [7] comprises a large- and small scale fading component with individual parameters for each scenario. It is worth noting that this model is also used by 3GPP, see [32].

4.2.1 Large Scale Channel Model

The large scale model defines deterministic, distance dependent path loss plus log-normally distributed shadowing loss. The channel conditions on each link between UE and BS can
be either line-of-sight (LoS) or non line-of-sight (NLoS). Channel conditions on a link are assumed not to change during one simulation run, since UEs do not move. The channel condition of the link between BS and UE is chosen randomly and the probability for LoS channel conditions $P(C_{c,s} = \text{LoS} | d_{c,s})$ depends on distance $d$. Distance $d$ between BS at spatial position $[x_c, y_c, z_c]$ and UE at position $[x_s, y_s, z_s]$ is defined in [7] as the two dimensional Euclidean distance

$$d_{c,s} = \sqrt{(x_c - x_s)^2 + (y_c - y_s)^2}. \quad (4.1)$$
Table 4.1: Selected IMT-Advanced Evaluation Scenario Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>InH</th>
<th>UMi</th>
<th>UMa</th>
<th>RMa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier Frequency [GHz]</td>
<td>3.4</td>
<td>2.5</td>
<td>2.0</td>
<td>0.8</td>
</tr>
<tr>
<td>Transmission Bandwidth [MHz]</td>
<td>40</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Max. TX Power BS TDD [dBm]</td>
<td>24</td>
<td>44</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>Max. TX Power BS FDD [dBm]</td>
<td>21</td>
<td>41</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Max. TX Power UE [dBm]</td>
<td>21</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Transceiver Height BS [m]</td>
<td>6</td>
<td>10</td>
<td>25</td>
<td>35</td>
</tr>
<tr>
<td>Transceiver Height UE [m]</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Max. Antenna Elements BS</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Max. Antenna Elements UE</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

All large scale channel model parameters and random value realizations are reciprocal, so channel conditions are identical for the UL and DL \( (C_{c,s} = C_{s,c}) \). Channel conditions (LoS/NLoS) are assumed identical between a UE and all three sector antennas of a BS site. Table 4.2 summarizes expressions for calculating LoS probability in all scenarios and Figure 4.3 shows scenario specific LoS probability over distance.

For each scenario and channel condition the mean path loss is

\[
\bar{h}_{PL,C,c,s} = \beta_C + \Lambda_C \log_{10} d_{c,s} \tag{4.2}
\]

except for scenario RMa. In the RMa scenario path loss also linearly depends on distance accounting to

\[
\bar{h}_{PL,C,c,s} = \beta_C + \Theta d_{c,s} + \Lambda_C \log_{10} d_{c,s}. \tag{4.3}
\]

In LoS channel condition parameters \( \beta, \Lambda \) differ for link distances below and above a break point distance \( d_{BP} \), see
### Table 4.2: Distance Dependent LoS Probability

<table>
<thead>
<tr>
<th>Scenario</th>
<th>LoS Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>InH</td>
<td>$\frac{1}{27}e^{-(d-18)}$, $d \leq 18$ m&lt;br&gt;0.5, $d &gt; 18$ m</td>
</tr>
<tr>
<td>UMi</td>
<td>$\min\left(\frac{18}{d}, 1\right) \left(1 - e^{(d/36)}\right) + e^{(d/36)}$</td>
</tr>
<tr>
<td>UMa</td>
<td>$\min\left(\frac{18}{d}, 1\right) \left(1 - e^{(d/63)}\right) + e^{(d/63)}$</td>
</tr>
<tr>
<td>RMa</td>
<td>$e^{-\frac{d-10}{1000}}$, $d &gt; 10$ m</td>
</tr>
</tbody>
</table>

![Figure 4.3: LoS channel condition probability over distance.](image)

Table 4.3. Path loss for indoor users in the UMi scenario is calculated from a uniformly distributed indoor path length $d_{in,c,s} \sim U(0, \min(25, d_{c,s})[m]$ and an outdoor path length $d_{out,c,s} = d_{c,s} - d_{in,c,s}$, namely the remainder of the whole path length minus indoor path length. Outdoor component of the path loss of indoor users is calculated the same way as for outdoor users only considering outdoor distance $d_{out,c,s}$. 20 dB wall penetration loss and 0.5 dB loss per meter indoor distance are added to this. UEs in vehicles in UMa and RMa scenarios experience a normally distributed penetration loss.
Table 4.3: Path Loss Model Parameters

<table>
<thead>
<tr>
<th>Scenario</th>
<th>( \beta_C ) [dB]</th>
<th>( \Lambda_C )</th>
<th>( \sigma_C ) [dB]</th>
<th>( \Theta ) [dB/m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>InH LoS</td>
<td>43.43</td>
<td>16.90</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>InH NLoS</td>
<td>22.13</td>
<td>43.4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>UMi LoS</td>
<td>35.96</td>
<td>22.00</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>UMi NLoS</td>
<td>19.46</td>
<td>43.30</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>UMa LoS</td>
<td>34.02</td>
<td>22.00</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>UMa NLoS</td>
<td>19.57</td>
<td>39.09</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>RMa LoS</td>
<td>29.80</td>
<td>20.48</td>
<td>4</td>
<td>1.4 \times 10^{-3}</td>
</tr>
<tr>
<td>RMa NLoS</td>
<td>1.70</td>
<td>38.63</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>

\( \bar{h}_{\text{Car},s} \sim N(\mu_{\text{Car}}, \mu_{\text{Car}}) \) with mean \( \mu_{\text{Car}} = 9 \) dB and standard deviation \( \sigma_{\text{Car}} = 5 \) dB. Car penetration loss of a specific UE is identical for serving and interfering links.

Mean signal shadowing \( \bar{h}_{\text{Sh},C,c,s} \sim N(0, \sigma_C) \) is assumed normally distributed in logarithmic scale (dB) with zero mean and standard deviation \( \sigma_C \) which is scenario and channel condition specific, see Table 4.3. Finally, overall large scale channel loss \( \bar{h}_{C,c,s} \) of a link is modeled as a normally distributed random variable

\[
\bar{h}_{C,c,s} \sim N \left( \bar{h}_{\text{PL},C,c,s} + \mu_{\text{Car}}, \sqrt{\sigma_C^2 + \sigma_{\text{Car}}^2} \right). \quad (4.4)
\]

The mean car penetration loss \( \mu_{\text{Car}} \) and standard deviation \( \sigma_{\text{Car}} \) are set to zero for indoor- and outdoor users not in vehicles.
4.2.2 Small Scale Channel Model

Small scale fading creates channel quality fluctuation in time and frequency domains resulting in channel quality prediction errors causing choice of inappropriate MCS either wasting resources or causing BLER. Performance gain may be achieved when scheduling UEs on resources with best quality at current time instant.

4.2.2.1 IMT-A Small Scale Channel Model

The IMT-A small scale channel model [7] describes the fading on all links between all receive and transmit antenna pairs. It captures the correlation of the transmitted signal in space, time and frequency domain. The correlation in time and frequency domain depends on UE speed, which is scenario specific. Higher speed results in shorter correlation. The antenna geometry describes the number of antennas, their alignment (linear or circular) and polarization. It is used as input to the model to determine the correlation among antennas. The scattering parameters for each scenario differ according to the fact that the propagation environment is different in rural-, urban- and indoor settings. A more detailed description of the IMT-A small scale fading model is provided in [125].

The model implementation must be capable to calculate \( n_{TX} \times n_{RX} \times |K| \times |C| \times |S| \) complex channel attenuation values for a scenario with \(|S|\) UEs served by \(|C|\) cells, transmitting on \(|K|\) subchannels in frequency domain with \(n_{TX}\) and \(n_{RX}\) transmit and receive antennas, respectively. In a cellular IMT-A evaluation scenario with 57 BS sectors and 10 UEs per sector transmitting at 10 MHz bandwidth on 50 PRBs with two receive and two transmit antennas approximately \(6.5 \cdot 10^6\) values have to be calculated every millisecond according to the LTE subframe duration, separately for UL and DL. In a simplified scenario (Figure 4.1b) with three sectors in the center and 18 surrounding sectors at one ring of interfering sites \(0.8 \cdot 10^6\)
channel realizations must be calculated. For VoIP capacity evaluation with at least 200 UEs per sector, one UE transmit and two BS receive antennas at 5 MHz channel bandwidth 32.5 \cdot 10^6 and 4.4 \cdot 10^6 channel realizations must be calculated for 57 and 21 sectors, respectively, of the cellular scenarios shown in Figure 4.1.

4.3 System Model Baseline Assumptions

The IMT-A Evaluation Guidelines define requirements for simulation models used to evaluate candidate system performance described in this section. 3GPP provides information on meeting these requirement in a reference simulation model described in [32]. For calibration purposes and in order to establish a common baseline values and assumptions are provided proposing simplifications to certain protocols.

4.3.1 Cell Assignment

Due to the channel model introduced in Section 4.2 BS closest to a UE is not alway the best one with respect to channel quality. According to [32] a UE should determine all BS links with large scale fading channel loss $\bar{h}_{\text{Min}}$ and with loss $\geq \bar{h}_{\text{Min}} - 1$ dB and from there select randomly one and the corresponding BS as serving. This models the measurement impairments of UEs during cell selection.

4.3.2 Protocol Overhead

According to [7] user- and control plane overheads should be modeled realistically.

4.3.2.1 User Plane Overhead

For the user plane, all Layer 1 (L1) and Layer 2 (L2) protocol overheads are included. The user plane comprises all layers
of the E-UTRAN protocol stack (Figure 2.3) devoted to user data transmission. The PDCP and RLC sublayer described in Section 2.2.1 and 2.2.2 use short sequence numbers and Unacknowledged Mode (UM). Long sequence numbers and Acknowledged Mode (AM) would add more overhead. The 3GPP reference simulator model [32] does not define a header compression technique in PDCP sublayer since the overhead saving is neglectable for large PDU size used for CSE evaluation. IMT-A Guidelines [7] for evaluation of VoIP capacity foresee header compression as described in Section 4.4.2.

4.3.2.2 Control Plane Overhead

The control plane comprises all layers of the E-UTRAN protocol stack serving for transmission of control data. Control channel overhead includes channel quality feedback, ARQ feedback, reference signals, and signaling of resource grants. ARQ feedback and reference signals are bidirectional and therefore consume UL and DL resources. Channel quality feedback is only transmitted in the UL and resource grants are only signaled in the DL. One to three symbols within the entire frequency band in each subframe can be dedicated as the control region transmitting DL control channels as described in Section 2.1. If less resources are dedicated for the control channel, less users can be scheduled and less HARQ feedback can be transmitted. The base line for LTE simulation assumes a three symbol wide control region. Since AMC (see Section 2.3) is applied for the PDCCH, the actual number of users that can be scheduled depends on the DL channel quality. A simplified assumption is that eight UEs can be scheduled in both UL and DL [32] confirmed by simulation results [126]. According to baseline assumptions [32], four (at 10 or 20 MHz) or two PRBs (at 5 MHz) are reserved for the PUCCH, mainly occupied by CQI and Precoding Matrix Index (PMI) reports, see Section 2.3. The amount of available PUCCH resources affects the feedback resolution.
Multiple REs in the DL and UL are reserved for transmission of reference signals, see Figure 2.2 reducing the gross data rate available for control and user data. The number of cell-specific RSs in the DL depends on the number of transmit antennas. DM RS are always transmitted in the UL while transmission of SRSs is optional. SRSs are disabled in UL if full-buffer traffic is simulated according to LTE base-line assumptions [32]. SRS transmission may be enabled at appropriate intervals to estimate channel parameters of UEs that have not transmitted any data recently.

4.3.3 Error Modeling

Packet error of user- and control plane transmission must be modeled accurately [7]. The LTE simulation model [32] assumes an error free control channel. The assumption is reasonable due to the very robust MCSs used for control channel transmission. Although AMC is applied to the PDCCH no packet error is assumed to occur.

The LTE reference simulation model [32] does not say how errors in user data should be modeled. An effective SINR model is specified to derive present channel quality over the entire transmission bandwidth as described in Section 2.2.3. This is appropriate since error probability is evaluated on entire TBs transmitted over multiple PRBs with different fading levels. Mapping of the effective SINR and related MCS and number and length of code blocks to BLER is based on link level simulation results.

4.4 Traffic Source Models

Two traffic types are defined in [7]: Full Buffer and VoIP.
4.4.1 Full Buffer Traffic Model

The full buffer model serves to evaluate UL and DL capacity and therefore always provides user data waiting for transmission to/from UEs. For that purpose IP-PDUs with maximum size of 1500 byte [127] are transmitted. MAC TB size depends on MCS used and can be larger then requiring concatenation of multiple IP-PDUs into one TB (see Figure 2.4). In a 20 MHz wide LTE configuration providing 100 PRBs a maximum TB size of 9422 byte can be transmitted with one spatial layer (single antenna) and 37482 byte with four spatial layers [11].

The total available bandwidth for full buffer simulations is 40 MHz for the InH and 20 MHz for the other scenarios. The bandwidth is divided equally between UL and DL for Frequency Division Duplex (FDD) mode.

4.4.2 VoIP Traffic Model

A discrete time two state Markov Chain is defined in [7] to model the VoIP traffic source at both sides BS and UE Figure 4.4. Accordingly a VoIP source is either in active state (talking) or inactive state (listening). UL VoIP sources are defined to be independent of the DL sources. Call establishment and release is not modeled, instead a fixed number of calls last for the entire simulation duration.

Time slot duration in the discrete time Markov Model is $T = 20$ ms according to the speech frame rate of common VoIP speech codecs [128, 129]. Transition probabilities $P(x|y) = 0.01$ for active- and inactive states are identical and $P(x|x) = 0.99$ is the probability to remain in a given state. The state sojourn time is therefore geometrically distributed with mean value $T/P(x|y) = 2$ s. In active state 244 bit long voice-PDU’s are generated every $T = 20$ ms containing output of the speech codec. In inactive state 40 bit long SID-PDU’s are generated every 160 ms containing information for generating comfort noise at the receiver. The related protocol overhead added by lower lay-
Chapter 4 – IMT-Advanced Evaluation Guidelines

Figure 4.4: Two-State Discrete Time Markov Chain (DTMC) VoIP Source

ers to voice- and SID PDU s is specified in [7] for compressed RTP [21], UDP [20], and IP [16] PCI to be 74 bit long resulting in 320 bit and 120 bit long compressed byte aligned voice- and SID SDU s of the Data Link Layer (DLL), respectively.

VoIP PDUs with transmission error or delayed more than 50 ms within E-UTRAN are considered lost. A voice user must receive 98 % of the PDUs successfully to be satisfied [7]. System capacity is defined by the number of VoIP calls served such that 98 % of the users are satisfied.

VoIP capacity evaluation is performed with reduced system bandwidth, namely 10 MHz for all reference scenarios, 5 MHz for UL and DL respectively.
5.1 Related Work

The idea of evaluating the performance of communication systems using Markov processes dates back to 1917 when Agner Krarup Erlang developed the Erlang Formula \[ \text{[130]} \] capable to evaluate the number of phone lines (trunks) to reach a defined call blocking probability under a given call arrival intensity. The assumption is that an arriving call will occupy a trunk for an exponentially distributed service time according to the call duration with mean value \( \mu^{-1} \) and no more calls can be served if all trunks are held by calls. The calls originate from a sufficiently large population so that they arrive with exponentially distributed inter arrival time (IAT) characterized by the mean arrival rate \( \lambda \).

The model was extended by Tore Olaus Engset \[ \text{[97]} \] to allow to evaluate the blocking probability if calls originate from a finite population with \( |S| \) users. In this case \( \lambda \) is the call arrival rate for one user. If \( x \leq |S| \) users are active performing a call using one of the trunks only \( |S| - x \) users can become active with aggregated call arrival rate \( (|S| - x)\lambda \). The Engset Model
enables evaluation of the number $|S|$ of users in a communication system with $k$ lines in the trunk that are served with a blocking probability $l(|S|, k)$. This model is used to calculate VoIP capacity for 4G systems described by the IMT-A Evaluation Guidelines [7].

SPS described in Section 2.3 to reserve PRBs for the duration of a VoIP talk spurt is well covered by Erlang or Engset models. VoIP packets of a talk spurt may be queued if all PRBs of a subframe are busy at time of arrival and are delayed until served by PRBs freed by VoIP calls switching to inactive state. The Engset loss model was enriched to become a queuing model in [131] described in more detail in [132]. Numerical calculation of the delay distribution is infeasible since the factorial of a large number of users and servers for values larger than about 170 leads to overflows even with modern computers. LTE operating at 5 MHz bandwidth has 25 PRBs per subframe. From 20 subframes per VoIP-IAT $25 \cdot 20 = 500$ servers per trunk are resulting. Due to 50% voice activity ratio up to $500/0.5 = 1000$ users would communicate at 100% load of the system.

In [133] a recursive numerically stable solution is provided for calculating the Engset queuing system delay distribution.

The VoIP source model was described in Section 4.4.2 as a DTMC. To calculate VoIP capacity with SPS a DTMC model is introduced as the discrete time version of the Engset queuing model studied in [133]. To the best of my knowledge only the solution for the delay distribution of the discrete time single server queue M/M/1/k/k with finite source population is known [134]. In [135] a closed form solution for the discrete time multi server queue with geometrically distributed service times M/Geo/n/$\infty$/k is presented using Probability Generating Function (PGF) with infinite source population. As said
before in a 5 MHz LTE system VoIP capacity is in order of the number of servers. Therefore a finite source population is derived using matrix analytic methods following especially [136] and [137].

5.2 Model Description

Figure 5.1 shows the system model as a queuing network with two multi server nodes. |S| jobs in the system represent the fixed number of VoIP calls accepted by a LTE system. A VoIP call in inactive state (Figure 4.4) is served in the upper node, which has as many servers as users in the system, so blocking or queuing never occurs and no queue is required. The lower node has $k$ servers representing the $k$ resources which each can serve a call\(^2\). More than $k$ calls in active state are queued and served by the First in First Out (FIFO) strategy. The system operates at discrete time steps of $T = 20$ ms known from the two state VoIP model introduced in Section 4.4.2.

In the following the steady state probability for the lower

\(^2\)One PRB per VoIP packet is assumed necessary.
node is derived. It is assumed that at each time step first calls that enter inactive state leave the system before new calls become active in the upper node and enter the lower node. Let $x$ be the number of active calls at current-, and $y$ at the previous time slot.

The probability for $x$ out of $X$ calls to change the state is binomially distributed with Probability Mass Function (PMF) described by Eq. (5.4). The transition probability $P(x|y)$ is given by Eq. (5.1). The term $\delta(x - y)$ is a Dirac impulse at the number of active calls $y$ in previous time slot. The number of arriving calls $a$ added to the system with distribution $a(x|y)$ is accounted for by convolution. It is binomially distributed according to Eq. (5.2) reflecting the fact that between zero and $|S| - y$ calls can become active. The number of departing jobs $d$ in Eq. (5.1) is distributed binomially according to Eq. (5.3). It is subtracted by convolution with Eq. (5.3) mirrored on the y-axis, indicated by the $-x$ in Eq. (5.1). If the total number of active calls $y$ in the system is larger than the number of servers $k$, only up to $k$ jobs in service can leave the system.

$$P(x|y) = \delta(x - y) * a(x|y) * d(-x|y), \quad (5.1)$$
$$a(x|y) = B(x||S| - y, p), \quad (5.2)$$
$$d(x|y) = B(x|\min(y, k), p), \quad (5.3)$$

$$B(x|X, p) = \binom{X}{x} p^x (1 - p)^{X-x} \quad (5.4)$$

The number $k$ of resources is equivalent to the PRBs of a LTE system, e.g. $25 \cdot 20 = 500$ for 5 MHz bandwidth.

$$\pi(x) P = \pi(x), \quad (5.5)$$
$$P = (P(x|y))$$

The steady state distribution of the number of active calls $\pi(x)$ is obtained by solving the equation system in Eq. (5.5), where $P$ is a $|S| \times |S|$ matrix containing the entries of $P(x|y)$. 
Figure 5.2 presents the steady state distribution $\pi(x)$ for a system with $|S| = 950$ calls and $k = 500, 950$ resources obtained from Eq. (5.5). If the number of active calls equals the number of resources active calls must never wait in queue. The steady state distribution is the binomial distribution $\pi(x) = B(x||S|, p = 0.5)$. Apparently the probability for $k > 500$ calls are active is low which is the reason for the small deviation for $\pi(x)$ under $k = 500, 950$, respectively. The dashed graph $P(x|475)$ in Figure 5.2 is the conditional state probability in the current time slot given the number of active calls in the previous one was $y = 475$ (Eq. (5.1)). Both, steady state and the conditional probability decrease rapidly to both sides of their maximum.

To be able to derive the waiting time distribution, the queue length distribution as seen by a call becoming active must be known. The transition matrix $P$ cannot directly be used for that, because its transitions are caused by the sum total of calls becoming active and remaining inactive. For example if there are $y = 10$ active calls in the previous slot and $a = 5$ calls become active and $d = 7$ calls become inactive, the resulting number of active calls in the current time slot is $x = 8$. The same
transition is executed if \( d = 2 \) calls become inactive and no call becomes active. To solve this matrix \( Q \) is introduced (see Eq. (5.6)). Its left side represents transitions with arrivals \((a > 0)\), and its right side with no arrivals \((a = 0)\). The respective transition probabilities \( P(x, a = 0|y) \) and \( P(x, a > 0|y) \) are provided in Eq. (5.7), see [138].

\[
Q = \begin{pmatrix}
P_{a>0} & P_{a=0} \\
P_{a>0} & P_{a=0}
\end{pmatrix}
\]

\( P_{a>0} = (P(x, a > 0|y)), \)
\( P_{a=0} = (P(x, a = 0|y)), \)
\( P(x, a = 0|y) = \delta(x - y) \cdot a(x = 0|y) \cdot d(x|y), \)
\( P(x, a > 0|y) = \delta(x - y) \cdot (a(x|y) - a(x = 0|y) \delta(x)) \cdot d(x|y) \) (5.7)

The system is now described by a two dimensional state variable \((x, \alpha)\), \( x \in 0...|S|, \alpha \in 0, 1 \) where \( x \) still represents the number of active calls and \( \alpha = 0, \alpha = 1 \) describes if the state was reached through a transition without or with arrival, respectively. The steady state probability of the system \( \pi_Q(x, \alpha) \) is found by solving

\[
\pi_Q(x, \alpha)Q = \pi_Q(x, \alpha),
\]
\[
\alpha = \begin{cases} 
0 : a = 0 \\
1 : a > 0
\end{cases}
\]

(5.8)

The states reached after at least one arrival occurred describe the system state after all arrivals. To calculate the waiting time distribution, the system state probabilities before an arrival happens must be known, since these are the ones encountered by calls switching to active state. The following is based on [138]. The state probability distribution conditioned on at least one arrival \( \pi_{Q,a>0}(x) = \pi_Q(x|\alpha = 1) \) is calculated by Eq. (5.9).
Analogously the conditional transition matrix $Q_{a>0}$ only considering transitions with at least one call becoming active is calculated by Eq. (5.10). By multiplying the current state probability vector $\pi_{Q,a>0}(x)$ with the matrix inverse $Q_{a>0}^{-1}$, the state probabilities in the previous time slot $\pi_{Q,a>0}^{-1}(y)$ are calculated (Eq. (5.11)). Those probabilities describe the system as it is seen by arrivals to the system.

$$\pi_{Q,a>0}(x) = \frac{\pi_Q(x, \alpha = 1)}{\sum_{i=0}^{\lfloor s \rfloor} \pi_Q(x = i, \alpha = 1)}$$  \hspace{1cm} (5.9)$$

$$Q_{a>0} = (P(x|y, a > 0)),$$  \hspace{1cm} (5.10)

$$P(x|y, a > 0) = \frac{P(x, a > 0|y)}{1 - a(0|y)}$$

$$\pi_{Q,a>0}^{-1}(y) = \pi_{Q,a>0}(x)Q_{a>0}^{-1}$$  \hspace{1cm} (5.11)

Calls becoming active arrive at the start of a slot at same time in batches with binomially distributed batch size. From the point of view of a tagged job the queuing position within the batch is uniformly distributed. In the following an extension to [138] for state dependent batch sizes is provided. The result is the distribution $q(x|y)$ of active calls in the system in front of a tagged job entering the system at slot start with $y$ jobs present

$$q(x|y) = \delta(x - y) * e(x|y) * d(-x|y)$$  \hspace{1cm} (5.12)

Before the tagged job, $e$ other jobs with distribution $e(x|y)$ have entered the system. The number of calls becoming inactive is binomially distributed with PMF $d(x|y)$, see Eq. (5.3). According to [138] the distribution of the number $x$ of active calls in front of a tagged call $e(x|y)$ is calculated as the cumulated probability to have more than $x$ arrivals, normalized to its expected value $E[a>0(x|y)]$ as shown in Eq. (5.13). The distribution $a>0(x|y)$ accounts for the fact that at least one call must become active.
Chapter 5 – LTE VoIP Queueing Model

\[ e(x | y) = \frac{\sum_{i=x+1}^{\vert S \vert} a_{>0}(x = i | y)}{E[a_{>0}(x | y)]}, \]

\[ a_{>0}(x | y) = \frac{a(x | y) - a(0 | y) \delta(x)}{1 - a(0 | y)}, \]

\[ E[a_{>0}(x | y)] = \frac{(\vert S \vert - y)p}{1 - a(0 | y)} \] (5.13)

\[ e(x | y), \text{ the probability for a tagged call becoming active when } y \text{ calls were active in the time slot before can be explained as follows:} \]

\[ a_{>0}(z | y) \] is the probability to have \( z \) arrivals in any slot having at least one arrival. The probability to have \( z \) arrivals in the slot where the tagged call becomes active is different as given by Eq. (5.14). This is due to the fact that the probability for the tagged call to be part of the batch of arriving calls increases linearly with the batch size \( z \). The expected value \( E[a_{>0}(z | y)] \) is normalizing the expression.

\[ a_{>0, tag}(z | y) = \frac{za_{>0}(z | y)}{E[a_{>0}(z | y)]} \]

(5.14)

\[ g(x + 1 | z) = \begin{cases} 
\frac{1}{z}, & x < z \\
0, & x \geq z 
\end{cases} \]

(5.15)

\[ e(x | y) = \sum_{i=x+1}^{\vert S \vert} a_{>0, tag}(z = i | y)g(x + 1 | z) \]

(5.16)

The probability to be the \( x \)th call becoming active conditioned that \( z \) calls become active in total is uniformly distributed (Eq. (5.15)) since all calls are treated equally. The unconditioned probability is obtained by summing up (Eq. (5.16)) only including the non-zero terms according to the fact that the number of calls becoming active before the tagged call cannot exceed the total number of calls that became active. Inserting
Eq. (5.14) and (5.15) into (5.16) gives the result presented in Eq. (5.13).

The arrival distribution $a_{>0}(x|y)$ depends on the number of active calls in the previous time slot. In states where a lower number of calls becoming active is more likely, the likelihood for the tagged call to be among those is also lower. Analogous to the findings in [138] for a state independent arrival process resulting in Eq. (5.13)-(5.16), the probability for the tagged call to be part of an arriving batch increases linearly with the expected number of calls becoming active in state $y$ as described in Eq. (5.17). The denominator describes the normalization and is the expected value of activated calls $x$ of the expected values over all states $y$. The state probabilities $\pi_{Q,a>0}^{-}(y)$ from Eq. (5.11) describe the system as seen by any activated call.

$$w(y) = \frac{\mathbb{E}[a_{>0}(x|y)] \pi_{Q,a>0}^{-}(y)}{\sum_{i=0}^{\abs{S}-1} \left( \mathbb{E}[a_{>0}(x|y = i)] \pi_{Q,a>0}^{-}(y = i) \right)}$$  \hspace{1cm} (5.17)

The unconditioned PMF $q(x)$ as given in Eq. (5.18) describes the probability of an arbitrary call becoming active to enter the system at position $x+1$. Entering at position $1 \leq x + 1 \leq k$ results in immediate service of the call. Calls entering at positions $x + 1 > k$ are queued until they can be served. The state $y = \abs{S}$ is not possible since no calls can become active if all calls are already in active state in the previous time slot.

$$q(x) = \sum_{i=0}^{\abs{S}-1} q(x|y = i) w(y = i)$$  \hspace{1cm} (5.18)

$$D = (D(x|y)),$$

$$D(x|y) = \delta(x - y) \ast d(-x|y)$$  \hspace{1cm} (5.19)
The departure matrix $D$, given in Eq. (5.19), contains the probabilities of calls to become active before the tagged call became inactive (and released its occupied resources). The distribution of the position $q_\tau(x)$ of the tagged call after $\tau$ time slots is the product of the initial position in the system $q(x)$ multiplied by the matrix $D^\tau$ (Eq. (5.20)).

$$q_\tau(x) = q(x)D^\tau \quad (5.20)$$

$$l(|S|, k) = 1 - \sum_{i=0}^{k} q_2(x = i) \quad (5.21)$$

Calls must be serviced within two slot durations of 20 ms to meet the maximum VoIP delay requirement of 50 ms. The fraction of calls not meeting the delay requirement $l(|S|, k)$ is described by the complement to the sum over all state probabilities representing calls that are served up to state $k$ (Eq. (5.21)).

### 5.3 Numerical and Simulation Results

#### 5.3.1 Interference Free Scenario

In the following the results of the analytic model are compared to system level simulation results following IMT-A Evaluation Guidelines [7] as described in Chapter 4. There the same traffic model is assumed as for this analytic model. More details on the model implemented in the open source Wireless Network Simulator (openWNS) framework are provided in Chapter 7. A single cell is simulated and UEs are placed close to the eNB transmitting voice packets in one PRB only. No SID packets are transmitted. The DL is simulated for different system bandwidth resulting in 6, 15, 25, and 50 PRBs per subframe for bandwidths 1.4 MHz, 3 MHz, 5 MHz, and 10 MHz, respectively, see Table 2.1. The number $k$ of available PRBs during 20 ms is then 120, 300, 500, and 1000. Half of the $|S|$ users are active on average so the mean offered load $\rho$ is:
To determine VoIP capacity loss rate $l(|S| = 2k\rho, k)$ is evaluated. VoIP capacity is reached at highest offered load $\rho$ where packet loss is below 2%. As said before packet loss in the model is purely caused by packet queuing delay without packet retransmission caused by packet error.

Figure 5.3a shows the ratio of successfully received packets versus offered load $\rho$. Results of simulation and analysis are compared for varying system bandwidths. Higher system bandwidth results in higher VoIP capacity due to multiplexing gain. VoIP capacity counted in users per cell and users per cell per MHz is show in Figure 5.3b and 5.3c, respectively. Simulated VoIP capacity is $4(+1.9\%)$, $6(+1.1\%)$, $7(+0.7\%)$, and $5(+0.3\%)$ users higher than found by the analytic model. The reasons for that are:

**Time Granularity of Scheduling:** The analytic model considers 20 ms resolution in time and treats all requests for persistent reservations not served within 40 ms as too late, since they will experience $60\,\text{ms} > 50\,\text{ms}$ delay in the next time slot. The system level simulator has 1 ms resolution counting packets having $0 - 46\,\text{ms}$ queuing delay as successful since processing delay according to HARQ timing is 3 ms and transmission delay is 1 ms.

**Packet Versus Talk Spurt Scheduling:** The analytic model treats talk spurts as jobs requiring a server. Talk spurts finding all servers occupied are queued. In reality and in the simulator packets are generated every 20 ms and are queued if the preceding packet has not been served within 20 ms. Packet delay of more than 40 ms means three voice packets of a talk spurt are queued. After a server becomes free a persistent reservation is established and the HOL packet is transmitted in the simulator. The packets remaining in the queue may be transmitted dynamically between two reservations provided that there
(a) Received Packet Ratio vs. Offered Load

(b) Capacity

(c) Capacity per MHz Bandwidth

Figure 5.3: Simulated and analytic ratio of received packets rate versus offered load and resulting VoIP capacity.

are free PRBs. This way delay of voice packets of a talk spurt that without dynamic scheduling would wait more than 50 ms may be reduced below 50 ms. Figure 5.4a shows the mean of VoIP packets per subframe scheduled versus number of users per cell in a system with bandwidth 1.4 MHz. Up to six packets could be scheduled dynamically in the six PRBs available per subframe of which 5% to 7% are scheduled, (see Figure
Numerical and Simulation Results – 5.3

Figure 5.4: Mean of VoIP packets per subframe scheduled dynamically and ratio of failed dynamic scheduling attempts found by simulation. 1.4 MHz system bandwidth.

5.4a) since the other PRBs are reserved persistently. Almost no queued VoIP packets are dynamically scheduled successfully since 90% to 99% find no free PRB, see Figure 5.4b.

Results gained from the analytical model deviate by about 1% from system level simulation results and can be considered to provide an upper bound on VoIP capacity. The small deviations found result from modeling on talk spurt rather than packet level. The simulator can be considered validated by the analytic model for this special case of assumptions (no interference, no LA).

5.3.2 Scenario with Interference

A Queueing Network Simulator (QNS) following the assumptions of the analytic model based on Figure 5.1 is used to obtain VoIP capacity with users transmitting with different TB sizes depending on their channel condition (SINR), see Section 2.2.3.1. Results are compared against VoIP capacity results by system level simulation using openWNS simulator, see Section 7.1. The QNS and analytic model operate on talk spurt level.
openWNS determines the appropriate TB size for each user in DL based on channel estimation from RS and CQI feedback, see Section 2.2.3.1.1. The QNS and analytic model need the TB size distribution as input parameter. The analytic model in Chapter 6 allows to obtain the SINR and by this TB size distribution for a two cell scenario. A discussion on how to extend the model to more cells is provided, but this is computationally complex and has not been implemented. Therefore the TB size distribution from DL openWNS simulation for scenario UMa is used for the following evaluation, see Section 7.3.4 and Figure 5.5. There maximum TB size is limited to four PRBs. Users demanding more PRBs according to their reported CQI value transmit on four PRBs causing high PER. HARQ recovers the lost packets at the expense of increased delay. This is not accounted for in the QNS and analytic model and therefore PER is set to zero for openWNS results presented in the following. Transmission of SID-PDUs is also disabled in openWNS because not present in the analytic and QNS model.

The analytic model calculates the ratio of talk spurts with
queueing delay beyond 50 ms equaling the ratio of not satisfied users since all users are identical in that model. Due to nature of random simulation the QNS result is a user packet loss rate distribution since each user has a slightly different packet loss rate. The variance of the user packet loss rate is reduced with increased simulation runtime. If different TB sizes are introduced all users with same TB size have same packet loss rate according to the ratio of talk spurts queued for more than 50 ms.

QNS and openWNS were configured for fixed TB size distribution 17.5 %, 27.5 %, 17.5 %, and 37.5 % for TB size 1, 2, 3, and 4 PRBs, respectively, see Figure 5.5. Users with same TB size belong to the same class.

Figure 5.6a and 5.6b show the Complementary Cumulative Distribution Function (CCDF) of the delay for QNS and openWNS with load 65, respectively. The load is defined as the number of users per MHz bandwidth. Packets exceeding 50 ms delay are considered lost. Results for each class and the mean delay distribution over all classes is shown together with the analytic result, Eq. (5.20). The number of servers was reduced from $k = 20 \cdot 25 = 500$ for 5 MHz to $k_{\bar{K}} = 500/2.75 \approx 182$ for the analytic model according to the mean TB size 2.75 PRBs. This mean value based approach shows significant difference in the tail distribution of the delay where the analytic model has 0.5 % probability to exceed 50 ms while simulation results show 1 % probability.

Figure 5.7a and 5.7b show the complementary loss rate for QNS and openWNS, respectively and compare it to the result from the analytic model. The loss rates for load 65 match the probabilities to exceed 50 ms delay for each class, Figure 5.6a and 5.6b. In the same way loss probabilities for other loads in steps of 0.2 are derived. The mean loss from simulation and the analytic one both show highest integer load with loss rate below 2 % is 66. The loss rates for the classes representing users with different TB sizes slightly differ for QNS but show significant difference for openWNS. There four PRB users experience
more than 2 % packet loss for load 64. Users with TB size one have loss rate zero for load 67 where TB size two users experience very low loss 0.15 %.

The different loss rates for the TB size classes for QNS and openWNS is explained as follows: For QNS users with TB size below four PRBs get persistent reservations if one to three servers are free giving them a slight advantage over users with TB size four PRBs. In reality as implemented in openWNS this is true for each of the 20 subframes i.e. if all subframes have three free PRBs many users with smaller TB sizes can be scheduled while users with TB size four will have to wait in the queue. Users with TB size one are the only ones that can transmit on
Figure 5.7: Satisfied user ratio and loss rate distribution for openWNS and QNS for scenario UMa DL.

The last free PRB in each subframe and therefore have significantly lower loss rate.

VoIP capacity is defined as the load where 2% of the users have loss rate 2%. Since the four PRB users are the first ones to be unsatisfied and 37.5% of the users belong to this class VoIP capacity is the point where 5.3% of the users with TB size four PRBs are unsatisfied while all users with lower TB size are still satisfied.

Figures 5.7a and 5.7b show the Satisfied User Ratio (SUR) for QNS and openWNS, respectively. Simulation runtime for QNS was set to 100000 s resulting in very low variance of the loss rate distribution for each class, Figure 5.7c. Therefore increasing load from 65.8 to 66 causes all 37.5% users with TB size four
to be unsatisfied. VoIP capacity is 65 according to QNS result. openWNS simulation runtime was 1250 s \(^3\) and the determined loss rate for each class has higher variance, Figure 5.7d. Increasing load by 0.2 does not immediately result in all users with TB size four to be unsatisfied, but typical step size \([7]\) load one does. VoIP capacity determined by openWNS simulation is load 64.

Figure 5.7d shows the loss rate for users with TB size less than four PRBs is almost zero as TB size four PRBs users exceed loss rate 2 \%. This observation is used to further refine the analytic model: The loss rate of users with TB size less than four PRBs is assumed to be zero. Then the mean loss rate at the point where the TB size four PRB users have 2 \% loss rate is \((1 - 0.375) \cdot 0 + 0.375 \cdot 0.02 = 0.75\) %. Loss rate 0.75 \% is reached at load 65.4 in the analytic model, Figure 5.6d. Resulting VoIP capacity is load 65 which is very close to load 64 VoIP capacity determined by openWNS system level simulation.

It was shown that the analytic model is not only capable of deriving the upper bound VoIP capacity assuming all VoIP PDUs need one PRB for transmission but also for scenarios with interference. Required parameters are the mean TB size for VoIP PDU transmission and the ratio of users transmitting with highest TB size, as was shown exemplary for the UMa scenario. The number of servers in the analytic model is scaled by the mean TB size and the tolerated loss 2 \% is scaled by the ratio of users requiring largest TB size.

\(^3\)openWNS packet level simulation is more than 100 times slower than QNS talk spurt level simulation. Using multiple random number seeds only increases the confidence of the determined mean loss rate for each class but does not reduce the variance.
In this Section results obtained from openWNS [139] simulation of the InH scenario are compared to results gained from an analytical model introduced in the following.

The aim is to support confidence in the results gained by the openWNS simulator in a co-channel interference scenario.

### 6.1 DL Throughput Capacity in InH Scenario

3GPP consortium partners contributing system level simulation results for IMT-A evaluation assured their simulator implementations achieve similar results for SINR Cumulative Distribution Functions (CDFs) under common modeling assumptions [7] for LTE Release 8 reference parameters [32]. The reference scheduler serves DL resource requests RR. Small scale fading is the only time varying influence in serving requests from a full buffer traffic model.
6.1.1 SINR Distribution

All random values influencing CSE and CEUSE are drawn once at simulation start. Thereby, positions of nodes (eNBs, UEs) and large-scale channel model parameters are defined for all links. According to Section 4.2.1 path loss of a link between eNB \(c\) and UE \(s\) for channel condition \(C \in \{\text{LoS, NLoS}\}\) is normally distributed

\[
\bar{h}_{C,c,s} \sim N \left( \bar{h}_{PL,C,c,s} + \mu_{\text{Car}}, \sqrt{\sigma_C^2 + \sigma_{\text{Car}}^2} \right)
\]  

(6.1)

In the following, CSE and CEUSE are derived for the InH scenario [140]. There, all UEs are located indoors, so that mean car penetration loss \(\mu_{\text{Car}}\) and its standard deviation \(\sigma_{\text{Car}}\) equal zero in Eq. (6.1). According to Table 4.3 standard deviation is \(\sigma_{\text{LoS}} = 3\) and \(\sigma_{\text{NLoS}} = 4\) for links experiencing LoS and NLoS channel conditions, respectively. The InH scenario has two eNBs, see Figure 4.2. The position of the left eNB is chosen to be the origin of the coordinate system \([x_{c=1}, y_{c=1}] = [0\,\text{m}, 0\,\text{m}]\). Link distance of a UE located at position \([x_s, y_s]\) from both eNBs are \(d_{s,1} = \sqrt{x_s^2 + y_s^2}\), \(d_{s,2} = \sqrt{(x_s - 60\,\text{m})^2 + y_s^2}\). If both links experience LoS channel conditions the mean large scale channel loss following Eq. (4.2) is

\[
\bar{h}_{\text{PL},\text{LoS},s,c} = \beta_{\text{LoS}} + \Lambda_{\text{LoS}} \log_{10} (d_{s,c}).
\]  

(6.2)

Parameters \(\beta_C\) and \(\Lambda_C\) are provided in Table 4.3. UE \(s\) will be served by eNB \(c\) if large scale loss \(\bar{h}_{\text{LoS},1,s}\) on that link is lower than the loss on the other link \(\bar{h}_{\text{LoS},2,s}\). The probability of UE \(s\) to be served by eNB \(c\) is

\[
P(s \in S_1) = P(\bar{h}_{\text{LoS},1,s} < \bar{h}_{\text{LoS},2,s})
\]

1The measuring impairment described in Section 4.3.1 is neglected, due to its marginal influence. It is only applicable in the unlikely event, that the large scale loss difference on the links is below 1 dB.
\[ P(s \in S_1) = P(\bar{h}_{\text{LoS},2,s} - \bar{h}_{\text{LoS},1,s} > 0) = 1 - \frac{1}{2} \text{erf} \left( \frac{\bar{h}_{\text{PL},\text{LoS},1,s} - \bar{h}_{\text{PL},\text{LoS},2,s}}{\sqrt{2} \sigma_{\text{LoS}}} \right) \]  

proved in Eq. (A.4). The SINR distribution \( p(\gamma|s \in S_1) \) of UE, having LoS channel conditions on both links and being served by eNB1 is the probability density function (PDF) of the difference between the large scale losses \( \bar{h}_{\text{LoS},2,s} - \bar{h}_{\text{LoS},1,s} \) and is always positive:

\[ p(\gamma|s \in S_1) = \frac{\phi(\gamma|\bar{h}_{\text{LoS},2,s} - \bar{h}_{\text{LoS},1,s}, \sqrt{2} \sigma_{\text{LoS}}) \mathbb{1}_{\{\gamma \geq 0\}}}{P(s \in S_1)} \]  

(6.4)

The proof is provided in Eq. (A.5)-(A.7). \( \phi(x|\mu, \sigma) \) is the PDF of the normal distribution \( N(\mu, \sigma) \). The indicator function \( \mathbb{1}_{\{\gamma \geq 0\}} \) assures that only SINR values greater than zero are possible. Values below zero are not possible, since in this case eNB2 would be serving UE. Due to symmetry, the SINR distribution on the link connecting UE to eNB2 is analogous after swapping mean values \( \bar{h}_{\text{LoS},c,s} \). The SINR CDF \( P(\gamma|s \in S_c) \) of UE served by eNB with LoS channel conditions on links to both eNBs is

\[ P(\gamma|s \in S_c) = \frac{\Phi(\gamma|\bar{h}_{\text{LoS},\bar{c},s} - \bar{h}_{\text{LoS},c,s}, \sqrt{2} \sigma_{\text{LoS}}) + p(0|s \in S_c) \mathbb{1}_{\{\gamma \geq 0\}}}{P(s \in S_c)} \]  

(6.5)

\[ p(0|s \in S_c) = 1 - P(s \in S_c) \]

\( \Phi(x|\mu, \sigma) \) is the CDF of the normal distribution \( N(\mu, \sigma) \) and \( \bar{c} \) is the index of the interfering eNB.
6.1.2 Throughput Distribution and Cell Spectral Efficiency

For each SINR value, a MCS $0 \leq m \leq 28$ maximizing the net data rate $(1 - \text{BLER}(m, \gamma))r_m$ can be found. The SINR versus net data rate function is shown in Figure 6.1. BLERs are very small if SINR $\gamma$ is greater than a MCS dependent threshold $\gamma_{m,min}$ indicated by vertical lines. Assuming zero BLER and rate $(1 - \text{BLER})r_m = r_m$ as done in the following calculations only introduces a small error. The probability of UE, to transmit with data rate $r_m$, served by eNB, at position $[x_s, y_s]$ and LoS channel conditions on links to both eNBs is

$$P(r_m| x_s, y_s, C = \text{LoS}, s \in S_c) =$$

$$P(\gamma_{m+1,\text{min}}| s \in S_c) - P(\gamma_{m,\text{min}}| s \in S_c), \quad 0 \leq m \leq 27$$

$$P(\gamma_{28,\text{min}}), \quad m = 28. \quad (6.6)$$

Data rate $r$ is always larger than zero since SINR is always greater 0 dB in this scenario and the lowest MCS $m = 0$ requires less than 0 dB SINR to provide a positive throughput.

To obtain the data rate distribution for all possible combinations of LoS and NLoS channel conditions with regard to the two eNBs, the unconditional probability must be calculated:
\[ P(r_m|x_s, y_s) = \sum_{\forall C_{1,s}} \sum_{\forall C_{2,s}} \sum_{\forall c} P(C_{1,s}, C_{2,s}, s \in S_c) f(r_m), \]

\[ f(r_m) = P(r_m|x_s, y_s, C_{1,s}, C_{2,s}, s \in S_c), \]

\[ P(C_{1,s}, C_{2,s}, s \in S_c) = P(C_{1,s}|d_{s,1}) P(C_{2,s}|d_{s,2}) P(s \in S_c) \quad (6.7) \]

\( C_{c,s} \) is the channel condition on the link between UE and eNB, with \( C_{c,s} = \text{LoS} \) and \( C_{c,s} = \text{NLoS} \) meaning there are LoS and NLoS channel conditions on the link between eNB and UE, respectively. It depends on respective distance \( d_{s,c} \) in the InH scenario, see Table 4.2. The channel conditions on the links and the probability to be served by eNB are mutually independent, so the joint probability can be expressed as the product of the individual probabilities.

The MAC layer throughput of a station depends on the SINR dependent data rate of the MCS selected and the share of resources the station gets according to the resource assignment algorithm of the scheduler. LTE-A base line [32] specifies RR scheduling strategy to be used in the MAC layer. UEs transmit on all PRBs in frequency domain during their turn. Perfect channel knowledge and no packet loss is assumed. A UE served by an eNB serving \(|S_c|\) UEs in total has the DLL throughput \( B = r_m/|S_c| \). UEs randomly placed in the InH scenario area have a 50% chance to be served by either of the two eNBs. Repeating the experiment of placing UEs in the scenario results in a Binomial distribution of the number of UEs being served by a given eNB. For a total number of 20 UEs \(|S|\) in the InH scenario this distribution is

\[ P(|S_c|) = \binom{20}{|S_c|} 0.5^{|S_c|} 0.5^{20 - |S_c|} \quad (6.8) \]
The throughput distribution\(^2\) \(P(B|x_s,y_s)\) of UE\(_s\) at position \([x_s,y_s]\) is obtained by unconditioning Eq. (6.7) thereby taking into account that data rate distribution and distribution of the number of served UE\(_s\) are independent.

\[
P(B|x_s,y_s) = \sum_{i=1}^{\left|S\right|} P(|S_c| = i)P(r_m|x_s,y_s) \quad (6.9)
\]

Each UE\([x_s,y_s]\) occurs with equal probability since UE positions are assumed uniformly distributed in the area \(A \in \{x| -30 \text{ m} \leq x \leq 90 \text{ m}\} \times \{y| -25 \text{ m} \leq y \leq 25 \text{ m}\}\), see Figure 4.2. The UE throughput distribution is achieved by integrating area \(A\) and normalizing to \(A = 120 \cdot 50 \text{ m}^2\):

\[
P(B) = \frac{1}{A} \iint_\mathbb{A} P(B|x_s,y_s)dx_sdy_s = \\
\frac{1}{A} \int_{-25\text{ m}}^{25\text{ m}} \int_{-30\text{ m}}^{90\text{ m}} P(B|x_s,y_s)dx_sdy_s \\
A = 120 \cdot 50 \text{ m}^2 \quad (6.10)
\]

\(\text{CSE}\eta\) is obtained by integrating the expected data rate

\[
E(r_m) = \sum_{m=0}^{28} P(r_m|x_s,y_s)r_m \quad (6.11)
\]

and normalizing to area \(A\), bandwidth \(\omega\) and number of eNBs \(|\mathbb{C}|\):

---

\(\text{The extreme case } |S_c| = 0 \text{ (all UE\(_s\) served by one eNB) is omitted in the summation since there would be no interference in that case and the thermal noise power (so far neglected in the calculations because it is much lower than the interference power) would have to be accounted for. The introduced error is neglectable since } P(|S_c| = 0) \sim 10^{-6}\)
\[ \eta = \frac{1}{A \omega |C|} \int_{\mathcal{A}} \sum_{m=0}^{28} P(r_m|x_s, y_s) r_m \, dx \, dy = \]

\[ \frac{1}{A \omega |C|} \int_{-25 \text{ m}}^{25 \text{ m}} \int_{-30 \text{ m}}^{90 \text{ m}} \sum_{m=0}^{28} P(r_m|x_s, y_s) r_m \, dx \, dy \quad (6.12) \]

### 6.1.3 Numerical Evaluation

It is sufficient to evaluate just one quarter of the scenario area \( \mathcal{A} \in \{ x | -30 \text{ m} \leq x \leq 30 \text{ m} \} \times \{ y | 0 \text{ m} \leq y \leq 25 \text{ m} \} \) due to symmetry. Since no closed form solutions for both integrals are known the integrals are evaluated using Riemann Sums with subinterval lengths \( \Delta x_s \) and \( \Delta y_s \) with \( n_x = \lfloor 60 \text{ m} / \Delta x_s \rfloor \), \( n_y = \lfloor 25 \text{ m} / \Delta y_s \rfloor \) subintervals in \( x \) and \( y \) direction, respectively, according to Eq. (6.13) and (6.14).

\[ P(B) = \frac{1}{n_x n_y} \sum_{i=0}^{n_x-1} \sum_{j=0}^{n_y-1} P(B|x_s = x, y_s = y) \Delta x_s \Delta y_s, \]

\[ x = -30 \text{ m} + i \Delta x_s, y = 0 \text{ m} + j \Delta y_s \quad (6.13) \]

\[ \eta = \frac{1}{n_x n_y \omega |C|} \sum_{i=0}^{n_x-1} \sum_{j=0}^{n_y-1} P(r_m|x_s = x, y_s = y) \Delta x_s \Delta y_s \quad (6.14) \]

Figure 6.2a compares UE throughput distribution obtained from Eq. (6.13) with that gained from openWNS simulation of the InH scenario. The results match very well giving confidence that the simulation model for the InH scenario is correctly implemented.

Analytical results for CSE are shown in Figure 6.2b for different interval lengths \( \Delta x \) and \( \Delta y \) and compared to simulation
Chapter 6 – *Mathematical Analysis of InH Scenario*

Figure 6.2: Analytic and simulation results for the UE DL throughput distribution and CSE in the InH scenario.

Results supplemented with 95% confidence interval. Mean CSE is \(2.237 \pm 0.022\) bit/s/Hz according to the simulation result. Analytical CSE is almost constant at 2.236 bit/s/Hz for interval lengths below 0.1 m. For higher interval lengths, CSE shows a jitter with decreasing mean value, which is a modeling effect.

6.2 UL SINR Distribution in InH Scenario

An analytical evaluation of static UL system performance is more complicated to achieve than for the DL. Interference to an eNB can originate from any position in the scenario. Additionally, the contribution of UL power control, as described in Section 2.2.3.1.2, must be modeled.

In the following SINR distribution on UL for the InH scenario is analyzed. The method in principle can be extended to scenarios with more cells but a numeric evaluation is practically non-feasible owing to enumeration of all combinations of transmitting and interfering UEs.
6.2.1 SINR Distribution

UEs served by eNB1 transmits with power \( P = \alpha \bar{h}_{C,c,s} + P_0 \) where \( \bar{h}_{C,c,s} \) is the path loss, \( \alpha \) a factor for path loss compensation and \( P_0 \) the power control offset. The UL power control factor \( \alpha \) is set to 0.8 and the offset \( P_0 \) is assumed as \(-83\) dBm in the following. The received signal power at eNB1 is \( P_{S,c} = \bar{h}_{C,c,s}(\alpha - 1) + P_0 \). At the other eNB1 eNB2, interference power \( P_{I,c} = \alpha \bar{h}_{C,c,s} + P_0 - \bar{h}_{C,c,s} \) is received. UEs is served by eNB1 if the condition \( \bar{h}_{C,c,s} < \bar{h}_{C,c,s} \) holds, which means the path loss on the link to eNB2 is lower than the one on the link to eNB1. 

Figure 6.3 shows the received signal and interference power at eNB1 caused by UEs versus the path losses on both links. Each point in the graph represents a UE position experiencing a path loss to eNB1 and eNB1 according to the value on the x-axis and y-axis, respectively. The lower right part under the diagonal stands for the situation when the path loss to eNB2 is
higher than to eNB₁ and UE, is served by eNB₁. In this case the color represents the received signal power $P_{S,1}$ for UE at eNB₁. It is independent of the path loss to eNB₁. It is for example $-89$ dBm if the path loss to eNB₁ equals $30$ dB, $\alpha = 0.8$ and $P_0 = -83$ dBm.

The upper left above the diagonal stands for the situation when the path loss to eNB₁ is higher than to eNB₂. UE is then served by eNB₂ and becomes an interferer to eNB₁. Hence the color represents the received interference power from UE to eNB₁. The power also depends on the path loss $\bar{h}_{C,2,s}$ to eNB₂ because it influences the transmission power. The linear equation $\bar{h}_{C,2,s} = \alpha \bar{h}_{C,2,s} + (P_0 - P_{I,1})$ describes the dashed line for which the received interference power at eNB₁ is constant, e.g. $P_{I,1} = -95$ dBm. In Figure 6.3 an example is given with $P_0 - P_{I,1} = 12$ dBm resulting in $P_{I,1} = -95$ dBm received interference power.

The path losses $\bar{h}_{C,c,s}, c \in \{1, 2\}$ on both links are assumed to be independent and normally distributed random variables. The probability to experience a certain received signal or interference power at eNB₁ can be determined by integrating the joint probability density function along the line of constant received power. As in Section 6.1.1, $\phi(x|\mu, \sigma)$ and $\Phi(x|\mu, \sigma)$ describe the Normal PDF and CDF respectively. PDF of signal power received from UE at eNB located at $[x_s, y_s]$ conditioned on UE is served by that eNB $p(P_{S,1}|s \in S_c)$ is

$$p(P_{S,1}|x_s, y_s, C_1, C_2, s \in S_1) =$$

$$\frac{\phi(x|\bar{h}_{PL,C_1,1,s}, \sigma_{C_1}) \int_{x}^{\infty} \phi(y|\bar{h}_{PL,C_2,2,s}, \sigma_{C_2}) dy}{\Phi(0|\bar{h}_{PL,C_1,1,s} - \bar{h}_{PL,C_2,2,s}, \sqrt{\sigma_{C_1} + \sigma_{C_2}})} =$$

$$\frac{\phi(x|\bar{h}_{PL,C_1,1,s}, \sigma_{C_1})(1 - \Phi(x|\bar{h}_{PL,C_2,2,s}, \sigma_{C_2}))}{\Phi(0|\bar{h}_{PL,C_1,1,s} - \bar{h}_{PL,C_2,2,s}, \sqrt{\sigma_{C_1} + \sigma_{C_2}})},$$

$$x = \frac{P_{S,1} - P_0}{\alpha - 1}.$$  \hspace{1cm} (6.15)

The interference power received from UE at eNB conditioned on
tioned eNB is the serving eNB is

\[ p(P_{I,1}|x_s, y_s, C_1, C_2, s \in S_2) = \]

\[ \int_{-\infty}^{x_{max}} \phi(y|\bar{h}_{PL,C_2,2,s}, \sigma_{C_2})\phi(\alpha y + P_0 - P_{I,1}|\bar{h}_{PL,C_1,1,s}, \sigma_{C_1})dy = \]

\[ 1 - \Phi(0|\bar{h}_{PL,C_1,1,s} - \bar{h}_{PL,C_1,2,s}, \sqrt{\sigma_{C_1} + \sigma_{C_2}}) \]

\[ \phi(P_{I,1}|\bar{h}_{PL,C_1,1,s} - P_0 - \alpha \bar{h}_{PL,C_2,2,s}, \frac{\sigma_{C_2}^2}{\sigma_{C_1}^2} \Phi(x'|\mu', \sigma') \]

\[ 1 - \Phi(0|\bar{h}_{PL,C_1,1,s} - \bar{h}_{PL,C_2,2,s}, \sqrt{\sigma_{C_1} + \sigma_{C_2}}) \]

\[ x_{max} = \frac{P_0 - P_{I,1}}{1 - \alpha}, \]

\[ x' = (\alpha \sigma_{C_1}^2 + \sigma_{C_2}^2)P_{I,1}, \]

\[ \mu' = \alpha \sigma_{C_1}^2 (\bar{h}_{PL,C_1,1,s}(\alpha - 1) + P_0) + \sigma_{C_2}^2 (\bar{h}_{PL,C_2,2,s}(\alpha - 1) + P_0), \]

\[ \sigma' = \frac{\sigma_{C_1}^2 \sigma_{C_2}^2 \sqrt{\alpha^2 \bar{h}_{PL,C_1,1,s}^2 + \bar{h}_{PL,C_2,2,s}^2}}. \] (6.16)

The distributions of both, received signal and interference power, are valid for any combination of the channel conditions \( C_1 \in \{\text{LoS}, \text{NLoS}\} \) and \( C_2 \in \{\text{LoS}, \text{NLoS}\} \). The respective distributions \( p(P_{S,2}|s \in S_2) \) and \( p(P_{I,2}|s \in S_1) \) at eNB are derived analogously. The overall received power distribution function \( p(P_S) \) is derived by unconditioning Eq. (6.15) as weighted sum of the received power at a given eNB for each possible combination of channel conditions \( C_1 \) and \( C_2 \), see Eq. (6.17). The distribution of the received interference power \( p(P_I|x_s, y_s) \) is derived analogously.

\[ p(P_S|x_s, y_s) = \]

\[ \sum_{\forall C_1} \sum_{\forall C_2} \sum_{\forall C \in \{\text{LoS}, \text{NLoS}\}} \sum_{\{1,2\}} P(C_1, C_2, s \in S_c)p(P_{S,c}|x_s, y_s, C_1, C_2, s \in S_c), \]

\[ P(C_1, C_2, s \in S_c) = P(C_1|d_{s,1})P(C_2|d_{s,2})P(s \in S_c) \] (6.17)

The received thermal noise power \( P_{\text{Noise}} \) in the UL is approximately \(-116.4\) dBm on one 180 kHz wide PRB according to
The received UL interference power can be in the order of or even below the thermal noise power. This is caused by potentially very low transmit powers depending on the power control parameters. The received interference and noise power in dBm is

\[ P_{I+N} = 10 \log_{10}(10^{0.1P_I} + 10^{0.1P_{Noise}}). \]  
(6.18)

The distribution of \( P_{I+N} \) is derived through PDF transformation according to Equation (6.19) resulting in Equation (6.20).

\[
p(y) = \frac{10}{y \log 10} p(x(y))
\]

\[
y(x) = 10^{0.1x}, x(y) = 10 \log_{10}(y),
\]

\[
dx = \frac{10}{y \log 10}
\]

\[
dy = \frac{10^{0.1z} \log(10)}{10} p(y(z) + 10^{0.1N})
\]

\[
z(y) = 10 \log_{10}(y + 10^{0.1N}), y(z) = 10^{0.1y} - 10^{0.1N},
\]

\[
dy = \frac{10^{0.1z} \log(10)}{10}
\]

\[
p(z) = p(x(z)) \frac{10^{0.1z}}{10^{0.1z} - 10^{0.1N}}
\]

\[
x(z) = 10 \log_{10}(10^{0.1z} - 10^{0.1N})
\]

\( p(P_{I+N}) =\)

\[
p(P_I(10 \log_{10}(10^{0.1z} - 10^{0.1N})) \| x_s, y_s) \frac{10^{0.1z}}{10^{0.1z} - 10^{0.1N}}
\]

The distributions of received signal to interference plus noise power for the entire scenario is gained by integration over the area in the same way as done for the DL (see Section 6.1). To
(a) Received signal and noise power and estimated effective distribution.

(b) SINR and estimated effective SINR distribution.

**Figure 6.4: Comparison of analytical and simulated CDFs.** \( \alpha = 0.8, P_0 = -83 \, \text{dBm} \)

my best knowledge no closed form solution for integrating Eq. (6.15) and (6.16) over the scenario area exist. Therefore numerical evaluation to achieve \( p(P_S) \) and \( p(P_{I+N}) \) is done using Riemann Sums. SINR distribution is obtained by convolution:

\[
p(\gamma) = p(P_S) \ast p(-P_{N+I}) \quad (6.21)
\]

Figure 6.4 compares results of system level simulation using openWNS with analytic results gained from Eqs. (6.17), (6.20), and (6.21). The power control parameters were set to \( \alpha = 0.8 \) and \( P_0 = -83 \, \text{dBm} \) reflecting values later used for performance evaluation. The analytical results for SINR components and for overall SINR match the simulation results very well.

Unlike for the DL (Eq. 6.6), the UL throughput distribution cannot be easily obtained from the SINR distribution. The UL reference scheduler [32] schedules all UEs in each subframe giving each UE an almost equal number of PRBs. Since the number of served UEs is generally not equal for the two eNBs, a UE receives interference from multiple UEs served by other eNBs on a TB. Since all PRBs forming a TB use the same MCS, the effective SINR is calculated for LA according to Eq. (2.2).
Calculating the throughput distribution as done for the DL would require PDF transformation of effective SINR Eq. (2.2) for all possible combinations of served users in each cell. In [141] a method assuming linear SINR averaging in logarithmic (decibel) domain is presented to determine the throughput distribution. It proves the throughput distribution only depends on the SINR distribution calculated by Eq. (6.21). Numerically solving the transformed PDF of Eq. (2.2) is more difficult since computationally complex transformation from logarithmic to linear domain must be performed.

Extending the model to more than two cells is difficult for a similar reason: It would require calculating the sum of the random interference powers. The according PDFs need to be transformed to linear domain, numerically convoluted and transformed back to logarithmic domain. This even applies for normally distributed interference power, since a closed form solution for the sum of two not identical log-normally distributed random variables is not known [142].
## System Level Simulation Model and Results

### Content

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>VoIP Scheduler for Multi-Cell Scenarios</td>
<td>88</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Queues and Scheduling Strategies</td>
<td>88</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Call State Estimation</td>
<td>89</td>
</tr>
<tr>
<td>7.1.3</td>
<td>Resource Assignment and Link Adaptation (LA)</td>
<td>91</td>
</tr>
<tr>
<td>7.1.4</td>
<td>Queue Management</td>
<td>100</td>
</tr>
<tr>
<td>7.1.5</td>
<td>Preempting Persistent Reservations</td>
<td>101</td>
</tr>
<tr>
<td>7.2</td>
<td>Scenario</td>
<td>103</td>
</tr>
<tr>
<td>7.3</td>
<td>VoIP Capacity Results</td>
<td>107</td>
</tr>
<tr>
<td>7.3.1</td>
<td>Optimal Queue Sizes and Discard Timers</td>
<td>107</td>
</tr>
<tr>
<td>7.3.2</td>
<td>Relocation of Persistent Reservations</td>
<td>110</td>
</tr>
<tr>
<td>7.3.3</td>
<td>DL with OTT Scheduling of VoIP Packets</td>
<td>112</td>
</tr>
<tr>
<td>7.3.4</td>
<td>DL with SPS of VoIP Packets</td>
<td>117</td>
</tr>
<tr>
<td>7.3.5</td>
<td>Uplink</td>
<td>119</td>
</tr>
<tr>
<td>7.3.6</td>
<td>Comparison of the Impact of Resource Assignment Algorithms on VoIP Capacity</td>
<td>123</td>
</tr>
<tr>
<td>7.3.7</td>
<td>Preemption of Persistent Reservations</td>
<td>129</td>
</tr>
<tr>
<td>7.3.8</td>
<td>Soft Frequency Reuse (SFR) and VoIP Capacity in the UL</td>
<td>132</td>
</tr>
<tr>
<td>7.3.9</td>
<td>SIMO</td>
<td>135</td>
</tr>
<tr>
<td>7.3.10</td>
<td>Individual Contribution of the Algorithms and Parameters</td>
<td>137</td>
</tr>
</tbody>
</table>
7.1 VoIP Scheduler for Multi-Cell Scenarios

7.1.1 Queues and Scheduling Strategies

SPS is preferred for VoIP traffic due to limited control channel capacity as explained in Section 2.3. The scheduler has to distinguish among three packet types: voice-PDUs, SID-PDUs, and HARQ-retransmissions. It is not necessary to distinguish HARQ-retransmissions for voice- and SID-PDUs since both are scheduled dynamically. SID-PDUs are also scheduled dynamically while voice-PDUs are scheduled persistently meaning they receive a TB in a subframe and this reservation is repeated in every twentieth subframe.

Since voice traffic is delay sensitive, the EDF scheduling strategy \cite{76, 143} is applied to all packet types. Users are sorted in descending order with regard to the time their head of line PDU is waiting\(^1\). In case of identical waiting time the user with lower ID is scheduled first. In the DL, the scheduler can drop PDUs from the queues if they cannot be received within delay bounds anymore. The same is applied to initial packet transmissions in the UL using the PDU discard timer of the PDCP-sublayer described in Section 2.2.1. UL HARQ-retransmissions are under control of the scheduler by setting the NDI bit for a HARQ-process in the resource grant indicating to the UE that it should transmit new data instead of HARQ-retransmissions.

Figure 7.1 shows the scheduler structure. Four different queue classes per user to be scheduled, namely Persistent, HARQ, Activated, and Dynamic are shown. Within each queue class EDF scheduling is used. In a first step PRBs for all users having persistent reservations granted earlier are assigned.

\(^1\)This results in EDF scheduling since all packets have the same 50 ms deadline.
Then all pending HARQ retransmissions are scheduled dynamically. In the next step voice calls that have just switched to active state are granted persistent resources. Also calls that had persistent resources granted but need relocation of resources due to bad channel conditions are treated in this step. Unlike all other classes, newly activated calls may receive reservations in this or one of the next 19 subframes. Finally all remaining traffic is scheduled dynamically, namely SID-PDUs and voice-PDUs not recognized by the scheduler as such.

### 7.1.2 Call State Estimation

The state diagram presented in the following serves to distinguish between PDUs belonging to active and inactive voice calls. A voice call is recognized by its bearer type. Voice- and SID-PDUs may be distinguished by PDU size. The DL MAC buffer status provides the total amount of queued data and the number of queued PDUs. UL Buffer Status Reports (BSRs) provide information on the total amount of queued data for a group of bearers with limited resolution only. It is therefore difficult to distinguish between one 320 bit long queued voice-PDU and two SID-PDUs, each 144 bit long. Also different voice
codecs and ROHC configurations can result in other PDU sizes.

One method to characterize queued data is by considering the IAT. A call in active state should create voice-PDUs every 20 ms while inactive calls only create SID-PDUs every 160 ms. Even with delay jitter, especially in the DL, the inter arrival times of the two PDU types differ significantly and can be used to classify queued data.

Figure 7.2 presents a state diagram useful to detect the state of a call just from presence of data in the associated queue. This information is directly available for the DL and indirectly by BSRs for the UL. For simplicity no delay jitter is assumed, but it could be included by regarding subframe ranges rather than single subframes. It is assumed that upon bearer establishment it is known during which VoIP IAT period (20 subframes) according voice-PDU s will arrive. This subframe is referred to as the arrivalFrame. Initially a user is considered inactive. If there is data in the queue for the user (bufferedData == True) in the ar-
rivalFrame, a transition to PotentiallyActive state is performed. This intermediate state serves to distinguish between voice- and SID PDUs. If another PDU has arrived in the queue after 20 ms in the next arrival frame, this is interpreted as a voice-PDU and Active state is entered. Resources recurring every 20 ms are reserved for the call. In case no free resources can be found for a persistent reservation the deactivate() signal is sent forcing to treat the call as if it was inactive and therefore try to dynamically assign resources until the queued data is transmitted. If there is no queued data for a call during its active frame a transition to Inactive state is performed.

If the scheduler cannot find free resources for a new active call in its arrival frame it must relocate the call to a later subframe causing delay. The state machine must be informed about this relocation by the relocate() signal with the newFrame parameter providing the frame in which resources are reserved persistently. VoIP packets arriving to the queue at the arrival frame must not be transmitted dynamically but remain buffered until the newFrame is reached. In all other states except for the Relocate state (* Relocate) any data entering the queues at other than the arrival frame is transmitted dynamically if resources are available. The main reason for such data are subframes with all resources occupied causing delay to transmissions of the user queue considered.

7.1.3 Resource Assignment and Link Adaptation (LA)

The quality of a PRB used to transmit data depends on the current path loss of both serving node and interferers, transmit power and whether a PRB is reused in interfering cells.

Figure 7.3a gives an example for the DL interference from eNB2 experienced by UE1,1 served by eNB1. Interference power is the same regardless which UE is served by eNB2 assuming eNB2 applies no DL power control and serves all users with same transmit power. In multi cell scenarios a UE receives interference from multiple surrounding eNBs, although only
few dominantly contribute to the total interference power. Due to the statistical channel propagation properties it is still likely that short term fading caused by interfering channels with high and low attenuation present is averaged out. The quality of a PRB therefore approximately depends only on the current attenuation of the serving link and whether surrounding eNBs, especially the ones close to the UE transmit on it. For the schedule of the surrounding eNBs a worst case estimation can be made, namely that all of them transmit on the considered PRB, resulting in a given SINR value. This way LA under given estimated SINR for persistent resource allocations to a talk spurt lasting 2 s on average is significantly simplified.

The UL situation is shown in Figure 7.3b and 7.3c. The received interference power here depends on the distribution of the interfering UEs. In this example UE2,1 is closer to eNB1 serving UE1,1 and therefore causes significantly more interfer-
Figure 7.4: ACF of estimated SINR for a UE measured on one PRB for 250 subframes in the UMa scenario with VoIP traffic only.

ence (Figure 7.3b) than UE2, which is farther away (Figure 7.3c). The effect is magnified by closed loop UL power control (see Section 2.2.3.1.2) letting cell edge users transmit at higher powers than UEs closer to their serving eNB. Calculating a worst case received interference power as done for the DL is hardly possible for the UL. As long as eNBs do not coordinate their schedule among each other they should at least perform a predictable resource assignment by changing their schedule as rarely as possible. An eNB detecting certain interference on a PRB will then assume the same interference to exist every 20 subframes.

Persistent resource allocation to transmit voice-PDUs will create a regular interference pattern every 20th subframe due to the recurrence period of the voice codec. This can be exploited for IMT-A VoIP capacity evaluation since only VoIP users are present in the system according to the evaluation guidelines [7]. In reality all cells would have to schedule VoIP calls in the same regions of the resource grids and the ACF of the interference power for each subframe would have to be evaluated in order to detect interfering VoIP connections having peaks in the ACF every 20 ms. The ACF $C_{\gamma_i,\gamma}(n) = E[(\gamma_i - \bar{\gamma})(\gamma_{i+n} - \bar{\gamma})]$
of the SINR estimation over 250 subframes on a given PRB for a user in a scenario with 21 cells is shown in Figure 7.4. There are 4200 UEs in total in the scenario and all UEs perform VoIP calls. The peaks in the ACF every 20 subframes result from the 20 ms voice-PDU IAT. This proves the ACF over a relatively short period can be applied to detect VoIP traffic in interfering cells and estimate an individual SINR for each subframe within a 20 subframe period.

Figure 7.5 shows the TB size in PRBs required to carry a voice- and SID-PDU versus the effective SINR of the TB. UL and DL results for voice- and SID-PDUs are provided. According to Section 2.2.3, 15 MCSs represented by the 15 distinct CQI values can be used in the DL and 29 MCSs in the UL. As long as for a given SINR selecting a higher order MCS does not result in a smaller TB size in PRBs, the LA process selects the most robust MCS. This method was proposed in [144] and results in long SINR value intervals for each TB size, especially if the TB size is one or two PRBs, making LA more robust to varying channel conditions. The intervals are very narrow for low SINR values, especially for UL voice-PDUs. It is therefore favorable in terms of reduced transmission errors to omit some MCSs for UL voice-PDUs.

Each PRB in the subframe has a different SINR due to fre-
quency selective fading and interference. If a single PRB does not provide sufficient quality to use a MCS allowing TB size of one PRB, the TB must be extended to span further PRBs. In the UL, due to the SC-FDMA constraint, adjacent PRBs must be used. In the DL, theoretically arbitrary PRBs may be used to form the TB. To limit the set of potential PRB combinations establishing a TB, further constraints have to be defined to keep scheduler complexity within limits. The DL scheduler used in this thesis extends the TB to the next free PRB when needed until either the PDU fits the TB or the end of the resource grid in frequency domain is reached.

Figure 7.6 depicts how a set of candidate TBs, TB₀, is formed. For illustration the arithmetic mean of the SINRs of the PRBs forming the TB is used to calculate the effective SINR of the TB. A TB is described by the set of PRBs forming it. In this example it is assumed 5 dB SINR is required to fit the PDU into a single PRB and 3 dB for a two PRB wide TB. PRBs 0 and 4 are occupied and cannot be used to form a TB. The SINR of PRB 1 is too low for TB size one, so the TB is extended to PRB 2 resulting in an effective SINR of 4 dB, which is sufficient to transmit the PDU. In a next step the scheduler tries to form a TB starting at PRB 2. The SINR on PRB 2 is sufficient to form a TB spanning one PRB. Next, PRB 3 is selected as the first PRB of a potential TB. It does not provide sufficient quality for TB size...
one. In the **UL** no **TB** including **PRB** 3 can be created for this **PDU** since the next **PRB** is occupied and **TBs** must be formed from **PRBs** adjacent in frequency domain. In the **DL** the next free **PRB** **PRB** 5, is included in the **TB**. The resulting effective **SINR** of 3 **dB** is sufficient to transmit the **PDU**.

For **HARQ**-retransmissions and dynamic transmissions the **TB** candidate set only considers resources in the current sub-frame to form **TBs**. For calls that became active and require a persistent reservation, **PRBs** in the current and the next 19 sub-frames are evaluated to form **TBs** with the constrain that all **PRBs** forming a **TB** must belong to the same subframe. Another degree of freedom when forming the **TB** candidate set is the transmit power which influences the estimated **SINR**. Increased power may result in smaller **TB** size due to increased **SINR**. The implemented scheduler supports two power levels, **normal** and **boosted**. The normal power level is calculated according to the closed loop power control algorithm (see Section 2.2.3.1.2) in the **UL** and the boosted power level is a fixed offset relative to it.

The resource assignment process must select one **TB** K ∈ **TB**₀ from the **TB** candidate set. To allow more sophisticated decisions, further information is attached to each **TB** as listed in Table 7.1.

The last two attributes are used to implement fragmentation aware resource assignment algorithms in the **UL**. A **space** is defined as the set of consecutive, free **PRBs** between occupied **PRBs** or the edge of the resource grid in frequency domain. For example the first space in Figure 7.6 starts at **PRB** 1 and ends at **PRB** 3.

Table 7.2 lists the implemented resource assignment algorithms called **filters** since they reduce the **TB** candidate set to a single or no **TB**. Returning no **TB** allows to use the filter chain for admission control by deciding it is better not to serve the user at all than allowing to use one of the **TBs** from the **TB**

---

²Κ₁⁻ is the set of **PRBs** previously occupied by the respective **UE**
candidate set. The filters are divided into four groups, namely Frequency domain, Time domain, Channel quality aware, and, for UL only, Fragmentation aware. The Random filter does not belong to any group and returns one randomly chosen TB from the TB candidate set.

The filters can be concatenated as shown in Figure 7.7. In each step only TBs complying to the condition described in Table 7.2 are passed to the next filter. According to the last column in the table, filters either return 0, 1, or \( n \) TBs. Most filters operate on integer attributes like the size of the TB in PRBs so it is very likely that more than one TB in the candidate set will fulfill the filter condition. The Random filter can be obviously used as a final filter selecting a single TB from a set. This can also be achieved by first applying the Min. Delay filter returning only TBs from the subframe closest to the arrival frame of the call and then applying the First filter to select the TB with lowest PRB start index. Some filters may return an empty set depending on their input set. The RBList(\( \hat{K} \)) filter will return an empty set if none of the TBs in the input set consists solely of PRBs contained in the set \( \hat{K} \). The Limit Delay(\( d \)) filter will not
Table 7.2: **TB** Candidate Set Filters

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>out</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Frequency domain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>TB with lowest start PRB index</td>
<td>N</td>
</tr>
<tr>
<td>RBLList($\hat{K}$)</td>
<td>$K \subseteq \hat{K}$</td>
<td>N$_0$</td>
</tr>
<tr>
<td>Previous</td>
<td>$\max(</td>
<td>K \subseteq K_{t-}</td>
</tr>
<tr>
<td><strong>Time domain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min. Delay</td>
<td>$\min((sfIndex - afIndex) \mod 20)$</td>
<td>N</td>
</tr>
<tr>
<td>Limit Delay($d$)</td>
<td>$(sfIndex - afIndex) \mod 20 \leq d$</td>
<td>N$_0$</td>
</tr>
<tr>
<td><strong>Both domains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>Return one random <strong>TB</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Channel quality aware</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest SINR</td>
<td>$\max(estSINR)$</td>
<td>N</td>
</tr>
<tr>
<td>Shortest</td>
<td>$\min(</td>
<td>K</td>
</tr>
<tr>
<td>Boosted</td>
<td>Boosted <strong>TBs</strong></td>
<td>N$_0$</td>
</tr>
<tr>
<td>Not Boosted</td>
<td>Not boosted <strong>TBs</strong></td>
<td>N</td>
</tr>
<tr>
<td>Min. unboosted size($l$)</td>
<td>$tbSizeNB \geq l$</td>
<td>N</td>
</tr>
<tr>
<td>Weakest($p$)</td>
<td>Return fraction $p$ of <strong>UEs</strong> with highest path loss</td>
<td>N</td>
</tr>
<tr>
<td><strong>Fragmentation aware</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Largest Space</td>
<td>$\max(endSpace - startSpace + 1)$</td>
<td>N</td>
</tr>
<tr>
<td>Smallest Space</td>
<td>$\min(endSpace - startSpace + 1)$</td>
<td>N</td>
</tr>
<tr>
<td>Least Fit</td>
<td>$\max(endSpace - startSpace + 1 -</td>
<td>K</td>
</tr>
<tr>
<td>Best Fit</td>
<td>$\min(endSpace - startSpace + 1 -</td>
<td>K</td>
</tr>
</tbody>
</table>

return any **TBs** if all **TBs** of its input set result in higher delay between arrival and scheduling subframe than $d$. Boosting is only applied if the **TB size with normal transmit power exceeds one **PRB** and if the **TB size with boosting is smaller than the
Figure 7.7: Resource Assignment Filter Chain

Figure 7.8: NonEmpty Alternative filter selecting a non empty set if the preferred one (TB_{(i+1),a}) has no elements.

one without (tbSizeNB). It can therefore happen that an input set does not contain boosted TBs and consequently the Boosted filter will return an empty set.

Scheduling algorithms may define preferred resources but allow the use of non preferred ones otherwise. The NonEmpty
Alternative filter shown in Figure 7.8 is used to configure such behavior. It takes two input sets labeled $a$ and $b$ returning set $a$ if it includes at least one element, set $b$ otherwise.

### 7.1.4 Queue Management

Transmitting packets that cannot reach their destination within 50 ms is a waste of resources [145]. Multiple mechanism to drop such packets are available as illustrated in Figure 7.9. In literature [146] the maximum number of HARQ retransmissions is limited to three. A packet being transmitted more than five times cannot reach its destination in time, since each retransmission introduces 8 ms of delay at minimum.

Packets retransmitted less than five times can still experience too high delays due to waiting in the queue. Limiting the number of retransmissions does not account for the time a packet waited for its initial transmission (lower queue in Figure 7.9). Therefore packets are dropped from the queue if their waiting time exceeds a defined value [145]. This is called Reneging in Queuing Theory [147]. The scheduler in the eNB knows the exact age of DL data and the time passed since the initial UL transmission of a packet. Reneging is implemented in the UL.
by setting the NDI flag in the corresponding resource grant. The UE must then transmit new data and drop the pending retransmission from the HARQ buffer of the corresponding HARQ process. The eNB may drop the packet from the HARQ buffer in the DL. Same is being done in the DL for voice data waiting too long. In the UL the PDCP timer can serve this purpose as described in Section 2.2.1.

Another straight forward and easy to implement solution is limiting the buffer size. This will introduce a drop tail behavior not admitting new packets to a full buffer. In this case packets with higher delay remain in the buffer while newer ones are dropped. Limiting queue size can improve system performance by allowing to faster recover from a short period of high load [148].

### 7.1.5 Preempting Persistent Reservations

Data with a persistent reservation has highest priority according to Figure 7.1. This could prevent scheduling data that urgently needs to be transmitted due to the high delay experienced. This is often true for HARQ retransmissions.
scheduler therefore implements preemption of persistent reservations in order to free resources for HARQ retransmissions.

According to Section 7.1.1 data from the HARQ queue is scheduled first. After that the scheduler knows which retransmissions could not be scheduled due to insufficient free PRBs. For each retransmission the number of required PRBs is also known, since it is the same as for the initial transmission. A DL retransmission cannot be scheduled if it requires more free PRBs than available in the subframe. It is enough to release any persistent reservation occupying the required number of PRBs to fit the retransmission. In the UL the SC-FDMA constraint requires consecutive PRBs in frequency domain to fit the HARQ retransmission. Figure 7.10 shows an example how the set of all possible preemptions that will allow to schedule a retransmission requiring three PRBs is constructed. Persistent reservations are released, starting at lowest frequency index, until the resulting number of consecutive free PRBs is greater or equal to the TB size of the retransmission. The first possibility is to release the reservation for UE3 in this example and would result in exactly three adjacent free PRBs. Another option would be to unschedule UE1 and UE9 resulting in four free PRBs. Four free PRBs also result from releasing the reservation of UE7. Four different strategies exist to select one candidate from the set of possible preemptions:

**First:** Select the one with lowest index in frequency domain

**Random:** Select one according to uniform random distribution

**MaxPRBs:** Select the preemption resulting in the largest number of free adjacent PRBs

**MinUsers:** Select the one preempting the least amount of persistent reservations from different UEs

One PDCCH is required to signal the explicit release of a persistent reservation to a UE. Another one is used to schedule the HARQ retransmission on the now available PRBs. Preemption is therefore only possible as long as at least two PDCCHs are available. The MinUsers preemption strategy tries to minimize the amount of required signaling in order to maximize the
number of HARQ retransmissions benefiting from freed PRBs.

Preemption is performed for one HARQ retransmission at a time, starting with the longest waiting one. The preemption strategy selects which reservations should be released and then schedules the retransmission. Resources just reserved for the previously processed HARQ retransmission may not be released again in this round. As a result the probability of successfully preempting persistent reservation in favor of retransmission decreases with every iteration.

The data of UEs that have lost their persistent reservations is considered for scheduling again together with the UEs that have just switched to active state as described in Section 7.1. Finally further data is scheduled dynamically. Both is only possible if there are still PDCCHs available.

It is possible to configure a delay threshold to select which HARQ retransmissions are considered critical. Only those are then considered within the preemption process in order to limit the amount of required PDCCHs. Distinguishing between critical and noncritical HARQ retransmissions was motivated by the work published in [67].

### 7.2 Scenario

In the following VoIP capacity results for the IMT-A evaluation scenarios InH, UMi, UMa, and RMa described in Section 4.1 are presented. In order to limit simulation runtime only 21 eNBs each operating one of three sectors at seven different sites are deployed rather than 57 at 19 sites as specified in [7]. SUR is the key performance indicator specifying when a VoIP system reaches its capacity limit. The load is defined as the number of users per cell and MHz system bandwidth.

Figure 7.11 shows the SUR versus load in a 21 and 57 cell scenario. Transmission bandwidth is limited to 3 MHz resulting in 15 PRBs in order to be able to execute the simulations within the memory limits of the simulation platform. DL res-
ults for the UMa, UMi, and RMa scenario are shown. Horizontal lines in Figure 7.11d and all following VoIP capacity graphs indicate the minimum VoIP capacity requirement for IMT-A systems: 40 for scenario UMa and UMi and 30 for scenario RMa [124]. Figures 7.11a, 7.11b, and 7.11c, plot the SUR versus load for the three scenarios. Figure 7.11d summarizes the VoIP capacity of the scenarios. The VoIP capacity increases from load 32 to 35 (+9.4 %), 36 to 43 (+19.4 %), and 39 to 43 (+10.3 %) for UMa, UMi, and RMa scenarios, respectively. Simulating 21 cells results in higher capacity due to less interference as shown in Figure 7.12. The CCDF of the received signal power $S$ is almost the same for 21 and 57 while very low received interference power $I$ is less likely with 57 than 21 cells. This results in reduced probability of very high SINR values explaining reduced capacity.
The focus of this thesis lies in the capacity improvement for voice traffic in LTE with regard to resource assignment algorithms in time and frequency domain as well as power control. The small scale channel fading model described in Section 4.2.2 is disabled for the following evaluation in order to reduce the computational effort of simulations. Obtained capacity in this case is lower, since gains from spatial multiplexing cannot be achieved. At the same time channel estimation errors resulting from small scale channel fading are avoided leading to higher capacity.

Figure 7.13 shows the SUR in the InH scenario versus load. Results for one, two, and four receive antennas with small scale channel fading enabled and one receive antenna without small scale channel fading are provided. All UEs move at 3 m/s speed. Even at such low velocity and therefore high channel coherence time for a system with a single antenna and small scale channel fading less than 20 % of users are satisfied at load 90 i.e. experience a packet error rate below 2 %. Voice-PDU's are schedule persistently and the MCS is chosen when resources are allocated. Due to high channel quality variance the
channel quality measurement on talk spurt establishment does not reflect the later channel quality and packet errors and frequent reallocations of the persistent allocation highly degrade the performance.

With four receive antennas system capacity is reached at load 126. Maximum Ratio Combining (MRC) [125] is applied in the four receive antenna system significantly increasing SINR compared to less antennas allowing almost all users to transmit their VoIP packet in only one PRB.

With two receive antennas and small scale channel fading enabled as well as with one receive antenna and small scale channel fading disabled, system capacity is reached at approximately load 100. It is found for the InH scenario that the capacity of a single antenna system without considering small scale channel fading is about the same as that of a $1 \times 2$ SIMO system with small scale channel fading considered.

A fair comparison of VoIP capacity of OTT versus VoLTE calls with focus on the radio link is the objective of this thesis, see Section 1.2. VoLTE calls benefit from SPS enabled at the MAC layer to overcome control channel capacity limitations. Packets of each class (Persistent, HARQ, Activated, Dynamic) are scheduled according to EDF algorithm, see Figure 7.1. It is assumed OTT packets are also scheduled with EDF algorithm with classes HARQ and Dynamic to assure a fair comparison. RR results in lower delays than PF scheduling [50] and is therefore preferred for VoIP traffic. RR sorts users according to the order they were last scheduled and treats the user that has not been scheduled for the longest time first. EDF shows almost identical behavior since all packets have the same deadline 50 ms. Systems using different scheduling algorithms than EDF for OTT VoIP PDU would achieve even lower VoIP capacity.
7.3 **VoIP** Capacity Results

In the following, results for the **VoIP** capacity of **LTE** are presented. Results for **DL** and **UL** with **OTT** scheduling and with **SPS** with and without control channel constraint are presented. System bandwidth is set to 5 MHz (25 **PRBs**) if not stated different.

### 7.3.1 Optimal Queue Sizes and Discard Timers

Figure 7.14 shows the different components of the delay experienced by a **PDU**. At first a queuing delay applies waiting for the initial transmission (TX) in one of the non-HARQ queues, see Figure 7.1. Minimum **UL** scheduling delay is 3 ms because **UL** scheduling grants refer to resources 3 subframes later. Transmission delay is 1 ms (one subframe duration). At the receiver 3 ms processing delay apply. In case of packet er-
ror a HARQ retransmission is initiated starting with receiving a NACK after 4 ms processing and transmission delay. Further a HARQ retransmission queueing delay applies until the respective PDU is scheduled. The following HARQ retransmission experiences the same transmission and processing delay as the initial one. Minimum delay is 4 ms and 7 ms for DL and UL respectively allowing 46 ms and 43 ms delay budget for queuing and HARQ retransmissions.

Figure 7.15 shows the ratio of satisfied users versus queue size in bit for different PDU queuing delay limits. A voice PDU has 320 bit and a SID PDU has 144 bit. The impact of limiting the delay in the non-HARQ queue for initial transmissions (PDU Age) and the HARQ queue for retransmissions (HARQ PDU Age) is shown for the UMi scenarios with load 54 in UL. PDUs exceeding the limits shown as legends to the respective curves are dropped from the respective queue. The ratio of satisfied users is below 90 % if no delay limits apply. The queue size does not have a significant impact on the result. The ratio of satisfied users is increased with queue size 440 bit to 90 % if queuing delay for initial transmission is limited to 43 ms. PDUs older than 43 ms cannot be delivered on time since the scheduled grant refers to a subframe 3 ms later and 3 ms processing plus 1 ms transmission delay apply.
Limiting queuing delay of PDU$s in the HARQ retransmission queue to 43 ms significantly improves the ratio of satisfied users to around 95%. The size of the queue again has little impact. The 43 ms delay limit is measured from when the PDU entered the initial transmission queue. VoIP PDU$s requiring retransmission are likely to have suffered poor channel condition and are candidates for violating delay constraint under high system load. Therefore limiting the queuing delay of HARQ retransmissions rather than limiting the number of retransmissions increases SUR by 5%.

All following simulation results are valid for 43 ms max delay of initial and HARQ retransmission PDU$s and 1050 bit long queues per UE fitting no more than three voice-PDU$s.
Chapter 7 – System Level Simulation Model and Results

7.3.2 Relocation of Persistent Reservations

The effective SINR of each persistent reservation is evaluated before granting PRBs. If the effective SINR is too low relocation of PRBs granted could be considered. Persistent reservations are relocated only if estimated SINR is too low multiple times in a row as defined by parameter relocation threshold.

Figure 7.16a shows the ratio of satisfied users versus relocation threshold for the UMa scenario at load 48. Zero means relocation is performed whenever the estimated effective SINR is too low for successful transmission. Setting the threshold to infinity disables relocation. The SUR is above 98% for relocation threshold one to nine.

Figure 7.16b shows the number of persistent reservations relocated per subframe. Absolute and relative number of relocation

Figure 7.16 shows performance versus relocation threshold in UL UMa scenario at load 48.

(a) Satisfied User Ratio  (b) Number of Relocation

(c) HARQ Retransmissions  (d) Failed Relocations
cated reservations per subframe are shown. Less than 2% of the reservations are relocated regardless of the threshold equivalent to a mean of one relocation every 10 subframes.

As expected packet error rate increases with relocation threshold, see Figure 7.16c. Without relocation error rate is 27% and setting the threshold to zero decreases it to 2.5%. Figure 7.16d shows the ratio of failed relocations due to insufficient free resources in the corresponding subframe relative to the total number of reservations scheduled for relocation. Relocated users are treated equally as users that just became active and scheduled according to the EDF strategy. Around half of the relocations fail for threshold two or higher.

The bad performance of relocation with threshold zero results from dynamically scheduled HARQ- and SID-PDU transmissions causing interference in a particular subframe but not in the one 20 ms later. Unexpected low SINR during multiple consecutive 20 ms distant subframes indicate a persistent reservation in another cell causing increased interference. Then relocation is recommended because interference will likely be present for a long time.

Relocation threshold is set to one for all SPS scenarios evaluated in the following sections.

Figure 7.17a shows the ratio of unsatisfied users (1 - SUR)
and the mean $\text{PER}$ versus load. VoIP capacity is load 45, 47, and 48 for relocation threshold 0, $\infty$, and 1, respectively, marked by circles. The ratio of unsatisfied users exceeds the 2% limit at a point where the mean $\text{PER}$ of all users is around 0.2% because 98% of the users are still satisfied with very low $\text{PER}$ while few users with bad channel conditions determine VoIP capacity.

Figure 7.17b shows the mean TB size in PRBs versus load. It shows differences of around 20% where never relocating results in smallest TB size, always relocating in largest and relocation threshold one is in between. Also the mean number of HARQ retransmissions significantly changes depending on the relocation threshold, see Figure 7.16c. Highest VoIP capacity is reached at an optimal configuration with regard to TB size through more aggressive MCS choice and PER causing HARQ retransmissions. Despite the large changes in TB size and PER VoIP capacity only changes in the range of load 3.

### 7.3.3 DL with OTT Scheduling of VoIP Packets

The following results are valid for Over the Top (OTT) scheduling each PDU individually without persistent resource reservation for talk spurts. The scheduler is configured to apply a very simple resource assignment strategy using the First and Minimal Delay filters, Table 7.2, where PRBs are filled in ascending order in frequency and time domain. Power boosting is not permitted. The resulting filter chain is shown in Figure 7.18.

Figure 7.19 presents VoIP capacity of UMa, UMi, and RMa scenarios. Figure 7.19a shows the SUR versus load. Results for unlimited number of control channels ($\infty$ PDCCHs) and eight
Figure 7.19: DL VoIP Performance with and without control channel limitation and OTT scheduling.

PDCCHs [32] are shown with 95% confidence intervals.

A user is unsatisfied if its packet loss exceeds 2% and system capacity is reached when more than 2% of the users are unsatisfied, as described in Section 4.4.2. Figure 7.19b shows the scenario specific VoIP capacity defined as the load where...
the SUR is still greater than 98%. A confidence bound below 98% means the result cannot be confirmed with 95% probability. This inaccuracy is tolerated since the worst case error is overestimating the VoIP capacity by one user / cell / MHz.

With unlimited control channel resources capacity is reached at load 43, 49, and 51 for the UMa, UMi, and RMa scenario, respectively. With eight PDCCHs, capacity changes to load 41 (−4.7%), 46 (−6.1%), and 47 (−7.8%) for the UMa, UMi, and RMa scenario, respectively.

Evaluated simulation time is 10 seconds, although the IMT-A Evaluation Guidelines demand 20 seconds. Figure 7.19e shows the impact of simulation duration on the capacity value estimation via SUR for the UMa scenario which has the smallest capacity. It can be seen that the results remain almost unchanged if simulation time is increased. The number of satisfied users is slightly higher for 41 than for 40 users as a result of the random processes simulated.

Figure 7.19c shows the mean number of occupied PDCCHs versus load for the three scenarios with and without limited number of control channels. The number of used PDCCHs scenario independently grows linearly with increased load for unlimited number of control channels but nonlinearly for eight PDCCHs. The linear growth matches the analytic results from [58] also shown in Figure 7.19c. Although the mean stays below seven PDCCHs there seems to be a high variance causing the difference in capacity visible in Figure 7.19a. Control channel resources appear to be a bottleneck for overall VoIP capacity. Control channel congestion in a given subframe may vanish in

---

3Confidence bounds are calculated using normal distribution since each experiment was repeated at least 20 times. If SUR− is the lower confidence bound then the probability for SUR to be below 98 % is \( \Phi(0.98|\text{SUR}, s) - \Phi(\text{SUR}−|\text{SUR}, s) \) where SUR is the mean value and \( s \) the estimated variance. Worst case is if SUR = 0.98 (mean value exactly at 98 % bound) and probability for overestimating VoIP capacity by load one is 47.5 % then.

4The first five seconds of the simulation form the warm-up period and are not evaluated resulting in 15 seconds total simulation time.
Figure 7.20: DL TB size distribution and SUR when forbidding low MCSs.

(a) TB Size Distribution; 8(b) Mean TB Size and Number of Retransmissions versus MCS_{\text{Min}}; UMa Load 45

(c) Capacity versus MCS_{\text{Min}}; UMa Load 45

Figure 7.19d shows the mean number of non-persistently (OTT) occupied PRBs per cell versus load for the three scenarios. Due to scenario size and channel model parameters, the UMa scenario offers the lowest mean SINR and the highest is provided in scenario RMa [149]. This results in a higher number of usable and therefore occupied PRBs for the UMa compared to the RMa scenario. Owing to low SINR probability congestion is highest in the UMa scenario explaining its lowest VoIP capacity compared to the other scenarios.

Figure 7.20a shows the CCDF of TB size in PRBs for the DL in three scenarios when loaded to system capacity. SINR on DL
is estimated (see Section 7.1.3) assuming all interfering eNBs transmit concurrently resulting in a worst-case SINR estimation. The largest possible TB size of 13 PRBs is applied rarely (1 to 2 %) depending on scenario. More than 5 PRBs per TB are used with less than 20 % probability. Similar to [67] where the TB size is limited to one to three PRBs, the impact of limiting the MCSs applicable down to MCS$_{\text{Min}}$, thereby eliminating MCSs consuming excessive number of PRBs, is studied.

Figure 7.20b shows mean TB size and mean number of HARQ retransmissions versus MCS$_{\text{Min}}$ for a system in UMa scenario at load 45 operated beyond capacity limit\(^5\). Results with and without control channel limit are shown. The mean number of retransmissions is zero if all available MCSs are permitted, but this results in largest mean TB size. Increasing MCS$_{\text{Min}}$ leads to an increased number of retransmissions but a decreased TB size. The number of retransmissions decreases for MCS$_{\text{Min}} \geq 12$ where TBs require one or two PRBs. The reason is not a decreased error rate but an increased number of late PDUs in the retransmissions buffer being dropped. With unlimited number of PDCCHs mean of retransmissions is one for MCS$_{\text{Min}} \geq 12$.

Figure 7.20c shows SUR versus MCS$_{\text{Min}}$. It slightly increases from MCS$_{\text{Min}} = 0$ to MCS$_{\text{Min}} = 4$ and MCS$_{\text{Min}} = 6$ for a limited and unlimited number of PDCCHs, respectively. Apparently, a decreased number of PRBs per TB outweighs the increased packet error rate. Increasing MCS$_{\text{Min}}$ beyond 7 and 10, respectively reduces the SUR. With eight PDCCHs no more than $3 \cdot 8 = 24$, $2 \cdot 8 = 16$, and $1 \cdot 8 = 8$ out of 25 PRBs are used if the maximum TB size is decreased to 3, 2, an 1 PRB(s), respectively.

Figures 7.21a - 7.21c show SUR versus load for UMa, UMi, and RMa scenarios for MCS$_{\text{Min}} = 0$ and 6. MCS$_{\text{Min}} = 6$ was

---

\(^5\)The capacity limit is exceeded at load 45 with MCS$_{\text{Min}} = 0$ with and without limited control channel resources. Increasing MCS$_{\text{Min}}$ increases VoIP capacity beyond load 45.
Figure 7.21: Theoretical and actual DL VoIP capacity improvement if low MCSs are prohibited.

found optimal without control channel limit. Clearly, as summarized in Figure 7.21d, VoIP capacity is increased with unlimited control channel resources under $\text{MCS}_{\text{Min}} = 6$ in all scenarios. With limited number of PDCCHs VoIP capacity is only increased in the UMa scenario.

7.3.4 DL with SPS of VoIP Packets

Figure 7.22 shows the DL VoIP capacity with SPS. Results for $\text{MCS}_{\text{Min}} = 0$ and 6 are presented for number of PDCCHs limited to eight. With $\text{MCS}_{\text{Min}} = 6$ a maximum TB size of four PRBs results increasing VoIP capacity in all scenarios beyond the capacity (at eight PDCCHs) found for OTT scheduling in Figure 7.21d. Clearly SPS is superior in capacity compared to OTT scheduling.
Figure 7.22: **DL VoIP** capacity with **SPS** and prohibiting low **MCS**s.

### 7.3.4.1 Comparison of **DL VoIP** Capacity for **OTT** Scheduling and **SPS**

**OTT** scheduling without control channel capacity limitation shows slightly higher capacity than **SPS** because it strictly obeys the **EDF** strategy for all packets. This is not possible with **SPS** scheduling because persistent reservations have highest priority, see Figure 7.1. **SPS** outperforms **OTT** scheduling if control channel capacity is limited to eight **PDCCH** in the **UMi** and **RMa** scenario and shows identical capacity in the **UMa** scenario. Capacity is increased by more than 10% if **MCS**Min = 6 is set preventing **TB** sizes larger than four **PRBs**. With **OTT** scheduling this gain is only achieved with unlimited number of **PDCCHs**. **DL VoIP** capacity is 47, 53, and 54 users / cell / MHz for the **UMa**, **UMi**, and **RMa** scenario, respectively.
7.3.5 Uplink

7.3.5.1 Optimal Transmit Power

The $\text{IoTN}$ may not exceed 10 dB according to the IMT-A Evaluation Guidelines [7] described in Chapter 4. Transmit power is the main factor influencing $\text{IoTN}$, see Eq. (2.3) with the parameters $\alpha$ and $P_0$. Usually [123, 32] $\alpha$ is set to 0.8 adjusting transmit power of a UE to equal 80% of its path loss plus offset $P_0$.

Figure 7.23a shows the ratio of satisfied VoIP users in UMa, UMi, and RMa scenarios versus power offset $P_0$ for load 47, 49, and 54, respectively. The load is chosen such that some users are unsatisfied regardless of the value of $P_0$. In order to maximize the ratio of satisfied VoIP users $P_0$ should be set to $-80$ dBm, $-82$ dBm, and $-78$ dBm for the UMa, UMi, and RMa scenario, respectively. According to Figure 7.23b the 10 dB limit specified [7] for mean IoTN is violated for this configuration. Therefore $P_0$ is set to $-84$ dBm for the UMi and RMa scenario and to

with SPS and $\text{MCS}_{\text{Min}} = 6$.

---

63GPP specification [32] and IMT-A Evaluation Guidelines [7] use IoT as acronym for interference over thermal noise. Throughout this thesis IoTN is used since IoT commonly refers to the Internet of Things.
Figure 7.24: **SINR** versus **UL** power offset $P_0$ for all and the 2 % weakest users of each cell.

$-83$ dBm for the **UMa** scenario which appears to be suboptimal in terms of **VoIP** capacity.

Figure 7.24 shows the mean **SINR** versus transmit power offset $P_0$ in the **UMa**, **UMi**, and **RMa** scenario for load 47, 49, and 54, respectively. It is increasing for all scenarios as $P_0$ increases contradicting the results for the **VoIP** capacity in Figure 7.23 where maximums are present. **VoIP** capacity is reached when more than two percent of the users experience more than 2 % packet loss. An increased mean **SINR** over all users does therefore not translate into an increased **VoIP** capacity. A better insight would be possible if the results for the unsatisfied users could be evaluated separately. Whether a user is satisfied or not remains unknown until the end of the simulation. Collecting and keeping all results per user until the end of the simulation requires a huge amount of memory. The Wide Band Loss (WBL)$^7$ known at simulation start is used to collect packet loss results for users with high WBL that will likely be unsatisfied.

Figure 7.25a and 7.25c show packet loss ratio of user versus its WBL for load 51 and 55 in the **UMi** scenario. There is obvi-

$^7$Wide band loss includes all effects reducing the received signal power except for small scale fading. This includes path loss, shadowing and the impact of antenna directivity [31][28].
VoIP Capacity Results – 7.3

Figure 7.25: Correlation between WBL of a UE and its UL packet error rate.

VolP Capacity Results – 7.3

(4) WBL of a UE versus its packet loss rate; Load 51

(6) Relative WBL of a UE versus its packet loss rate; Load 51

(8) WBL of a UE versus its packet loss rate; Load 55

(9) Relative WBL of a UE versus its packet loss rate; Load 55

Figure 7.24 shows SINR versus transmit power offset $P_0$ for the weakest users. This now resembles the results for VoIP ca-
capacity from Figure 7.23 showing that SINR of the weakest users has a maximum and decreases as transmit power is increased beyond that point.

7.3.5.2 UL without LA

Figure 7.26 shows the UL VoIP capacity for the UMa, UMi, and RMa scenario. LA is disabled; instead a fixed MCS is used for all transmissions resulting in two and one PRB for VoIP- and SID-PDUs, respectively. Results for OTT and SPS are shown with eight and infinite number of PDCCHs. VoIP capacity with OTT scheduling and eight PDCCHs is reached at load 43, 44 and 47 in the UMa, UMi and RMa scenario, respectively.

Capacity is increased to around 48 (+11%) for UMa and UMi scenario and around 55 (+17%) for RMa scenario, if either SPS is used or an infinite number of PDCCHs is assumed. OTT VoIP
7.3.6 Comparison of the Impact of Resource Assignment Algorithms on VoIP Capacity

7.3.6.1 UL Frequency Domain SPS

For the following simulation studies LA is applied on UL based on interference measurement on each PRB per subframe every 20 ms, see Section 7.1.3.

Figure 7.27 shows filter chains for the four resource assignment algorithms First, Random, LeastFit, and BestSINR. Filter Min. Delay limits the set of candidate TBs to the ones close to the arrival time of a PDU. First, Random, LeastFit, and BestSINR filters are chained to the Min. Delay filter. LeastFit and BestSINR filters may return a set with more than one candidate TB which the First filter reduces to a single TB.

Figures 7.28a-7.28c show the SUR versus load for scenario UMa, UMi, and RMa comparable to the results in Figure 7.26.
Chapter 7 – System Level Simulation Model and Results

Figure 7.28: UL VoIP capacity for different resource assignment algorithms.

The difference is that now PRB assignment is further tried to be optimized resulting in higher VoIP capacity. In Figure 7.28d VoIP capacities are summarized. The lowest capacity is achieved with algorithm First where all cells tend to occupy PRBs with low index causing heavy mutual interference leaving PRBs with high index free.

Algorithm Random distributes TB requests equally among all PRBs reducing mutual interference across cells slightly increasing VoIP capacity.

Algorithm BestSINR performs best because the TBs with highest SINR require the smallest amount of PRBs. This results in reduced number of PRBs occupied permitting more TBs to be carried by the system.

The LeastFit algorithm shows similar results as algorithm BestSINR because it tends to select TBs with small size because
Figure 7.29: UL IoTN for First and BestSINR resource assignment algorithm with default and increased transmit power offset $P_0$.

Those are the ones filling the spaces least.

Figure 7.29 shows the IoTN for algorithm First and BestSINR. IoTN for algorithm First is 9.4 dBm, 10 dBm, and 9.8 dBm at VoIP capacity in UMa, UMi, and RMa scenario, respectively. The 10 dBm limit [7] for the IoTN is not exceeded. A slight transmit power increase by higher offset $P_0$ is recommended for the UMa scenario and might increase VoIP capacity according to Figure 7.23a.

The IoTN for the BestSINR algorithm is 8 dBm, 7.6 dBm, and 7.7 dBm when VoIP capacity is reached in the UMa, UMi, and RMa scenario, respectively. The increase in IoTN when the number of users is increased is significantly larger than for the First algorithm. In the following VoIP capacity for algorithm BestSINR with transmit power offset $P_0$ increased by 2 dBm from $-83$ dBm to $-81$ dBm for scenario UMa and $-84$ dBm to
−82 dBm for scenarios UMi and RMa is evaluated.

The SUR versus load for algorithm BestSINR with $P_0 = -81$ dBm for scenario UMa and $-82$ dBm for the other two scenarios is plotted in Figures 7.30a-7.30c. For comparison results for algorithm First are shown, see Figure 7.28. Increasing $P_0$ by 2 dBm results in an increased VoIP capacity according to Figure 7.30d. Relative to the results for algorithm First it is +11 %, +12 %, and +9 % for UMa, UMi, and RMa scenario, respectively. According to Figure 7.29 (circle marker) the corresponding mean IoTN values are 8.6 dBm, 9.4 dBm, and 9.1 dBm for the UMa, UMi, and RMa scenario, respectively. The 10 dBm limit [7] for the mean IoTN is therefore not violated and the power offset could even be increased by 1 dBm for the UMa scenario.

![UL VoIP capacity of the First and BestSINR resource assignment algorithms with and without increased transmit power.](image)
7.3.6.2 **UL Frequency Domain OTT Scheduling**

Figure 7.32a shows the SUR versus load for OTT scheduling for resource assignment algorithm *First* and *Previous* in scenario UMa and infinite number of PDCCHs. Algorithm *Previous* tries to assign as many PRBs as possible which were used during last transmission [67, 74]. If no previously used PRBs are available algorithm *First* is used, see Figure 7.31. This way a more regular channel occupation pattern to support interference estimation for LA and reduce the mean number of transmissions is created. Without LA the mean number of retransmissions is around 1.4 due to high BLER see 7.32b. VoIP capacity is load 46 and 47 for algorithm *First* and *Previous*, respectively. Both are below load 49 VoIP capacity achieved without LA. Mean number of transmissions is significantly reduced with LA (Figure 7.32b) but at the cost of increased mean TB size (Figure 7.32c). Without LA mean TB size is 1.89 PRBs because two and one PRBs are used per VoIP and SID PDU respectively and the ratio between the two PDU kinds is 8 : 1. The benefit of enabling TB size one for VoIP PDU is repealed by having transmissions
Figure 7.32: Impact of frequency domain scheduling and LA on UL VoIP capacity with OTT scheduling. Scenario UMa, ∞ PDCCHs.

with TB size beyond two PRBs. According to Figures 7.32b and 7.32c mean TB size for SPS is even larger but the BLER is significantly lower. Resource assignment algorithm Previous is unable to create an interference pattern for OTT scheduling as predictable as SPS does for VoLTE.

7.3.6.3 Comparison of UL VoIP Capacity for OTT Scheduling and SPS

Same as for DL VoIP capacity for OTT scheduling is lower than for SPS if control channel capacity is limited to eight PDCCHs.
VoIP Capacity Results

LA exploiting the 20 ms periodicity to estimate interference only increases VoIP capacity in the UMi scenario compared to using fixed TB size two and one PRB for VoIP and SID PDUs, respectively, if PRBs are filled from low to high index (First strategy). SPS and the BestSINR resource assignment strategy enabled by LA and higher transmit power increase VoIP capacity by around 10% for UMa and RMa and 15% for UMi scenario compared to the case without LA. VoIP capacity with OTT scheduling is not increased by LA due to high BLER caused by frequent channel estimation errors from the irregular channel occupation pattern.

7.3.7 Preemption of Persistent Reservations

In the following the influence of preempting persistent reservations in order to free resources for pending HARQ retransmissions is evaluated. The details of the mechanism are described in Section 7.1.5.

There four strategies for deciding which persistent reservation should be preempted in favor of a pending HARQ-PDU are introduced. Figure 7.33 compares VoIP capacity for strategies First, Random, MaxPRBs, and MaxUsers and preemption disabled (see Figure 7.28d) in UMa, UMi, and RMa scenarios. MaxPRBs strategy performs best followed by MaxUsers and Random strategies. The First strategy performs worst and results in no or only slight (UMi scenario) capacity gain compared to the capacity without preemption.

Figure 7.34 shows the VoIP capacity for preemption with MaxPRBs strategy and LA with BestSINR algorithm and transmission power increased by 2 dBm, see Figure 7.30. VoIP capacity is increased by load one for all scenarios compared to applying only LA but no preemption. The very irregular PRB assignment by scheduling algorithm BestSINR results in lower gains from preemption than for algorithm First.

Figure 7.35 shows the ratio of failed attempts to activate calls and to schedule HARQ PDUs versus load. Activation
fails if there are insufficient free PRBs for a call in active state not yet having a persistent reservation. Preemption improves VoIP performance by significantly increasing the probability to schedule HARQ-PDUs while leaving the probability for successful activation almost unchanged. The small capacity increase by load two is because only a small fraction of the PDUs are HARQ-PDUs benefiting from preemption.

Figure 7.36 shows the mean number of used PDCCHs per subframe versus load for the four preemption strategies and preemption disabled. All preemption strategies result in almost identical average number of used PDCCHs per subframe. Strategies MaxPRBs and MinUsers increase the mean number of
Figure 7.34: UL VoIP capacity with preemption and link adaptation.

Figure 7.35: Ratio of failed attempts to activate calls and to schedule HARQ PDUs in scenario UMa UL and Max-PRBs strategy.

OTT can achieve slightly higher VoIP capacity than SPS if sufficient PDCCHs are available because it strictly follows the EDF scheduling algorithm, see Figure 7.26. Preemption enables the same for SPS at the cost of around 25% more occupied PDCCHs. This is not critical due to the low number of used PDCCHs compared to OTT scheduling. Therefore the MaxPRB
Chapter 7 – System Level Simulation Model and Results

Figure 7.36: Mean number of used UL PDCCHs per subframe versus load for the four preemption strategies and with preemption disabled in scenario UMa.

Figure 7.37: Cell edge/center area and corresponding outer/inner PRB sets.

should be used in to improve VoIP capacity by up to 4 %.

7.3.8 Soft Frequency Reuse (SFR) and VoIP Capacity in the UL

SFR partitions the PRBs of a cell into two sets: outer set used to serve edge users with boosted transmit power and inner set for serving all other users with normal transmit power, see Figure 7.37. Cell edge users may be served on inner set PRBs with normal transmit power and cell center users may be served on outer set PRBs with normal transmit power.
VolP Capacity Results – 7.3

Studies show that SFR can increase the throughput of users located at the cell edge [150, 89, 151, 152]. In [153] LTE with SFR is evaluated assuming a traffic mix including VoIP traffic. Results show improved SINR and improved throughput under SFR but the ratio of satisfied VoIP users is not evaluated.

Figure 7.38 shows the structure of the filter chain used to realize SFR allowing boosted transmit power on PRB 1 to 8 for sector 1.

If TB size is greater than one PRB power boosting is applied and the respective PRBs assigned. This way a criterion to distinguish between cell edge and cell center UEs is established. Boosted transmit power is 4.77 dB higher than normal power (factor three). The filter chain assures cell edge UEs are preferably scheduled on outer set PRBs and use boosted transmit power. If no TB containing solely outer set PRBs is available a TB is
chosen containing solely inner set PRBs and normal transmit power is used. If this is not possible a TB containing PRBs from both sets is chosen and normal transmit power is used. Cell center UEs are preferably scheduled on inner set PRBs and use normal transmit power.

Figure 7.39b shows the UL VoIP capacity versus load for the UMa, UMi, and RMa scenario with and without SFR. Transmit power with SFR is reduced by 1 dBm in order to not violate maximum uplink IoTN 10 dB [7]. The capacity remains almost the same for all scenarios. At higher loads the SUR is higher for UMi and RMa scenario, but this does not affect VoIP capacity.
7.3.9 **SIMO**

SIMO with MRC is modeled assuming theoretically maximum gain [116] scaling the SINR by the number of receive antennas resulting in 3 dB SINR gain for two receive antennas (1 × 2 SIMO). Figure 7.40a shows the SUR versus load for OTT scheduling with and without control channel limit in the UMa scenario for one and two receive antennas. Without control channel limitations (∞ PDCCHs) VoIP capacity is increased by 41.7% from load 48 to load 68. Limited control channel capacity decreases VoIP capacity to load 45 and 49 for one and two receive antennas, respectively. Insufficient control channel resources prevent the system to exploit the huge receive antenna diversity gain.

Increasing the number of PDCCHs to 13 would allow to exploit multi antenna gains with OTT. Having more than eight PDCCHs is feasible, for example if total system bandwidth is 10 MHz and the remaining 5 MHz are used for other traffic like video streaming requiring less PDCCHs. Research indicates VoIP will only account for a small portion of total mobile Internet traffic [154] and other traffic classes with lower control channel demands due to larger packet sizes dominate.

Figure 7.40c shows the SUR versus load for the three scenarios with SPS and Figure 7.40d the respective VoIP capacity for SPS and OTT scheduling. VoIP capacity for scenario UMa and SPS with two receive antennas is load 66 and therefore slightly lower than load 68 for OTT with infinite PDCCHs since EDF scheduling is strictly applied for OTT, see Section 7.3.4.1. VoIP capacity is load 75 and 76 for scenario UMi and RMa respectively.

Increasing the number of receive antennas from one to two significantly increases VoIP capacity by around 40% with SPS. VoIP capacity is hardly increased for OTT scheduling with limited number of PDCCHs due to insufficient control channel capacity for so many users. Results are in line with findings from literature [58] and intermediate 3GPP VoIP capacity evaluation.
Figure 7.40: Comparison of DL VoIP capacity for SPS and OTT scheduling with eight and infinite PDCCHs.

[111, 114] showing that VoIP capacities beyond load around 45 are not possible with eight PDCCHs and OTT scheduling. VoIP capacity load 66 for UMa is very close to load 69 limited by DL assessed by 3GPP, see Table 3.3 and Figure 3.1c. Load 75 VoIP capacity for UMi is within the range of load 80 reported by 3GPP but the results for scenario RMa show significant difference of load 18, see Table 3.3 and Figure 3.1c.
7.3.10 Individual Contribution of the Algorithms and Parameters

Figure 7.41a shows the impact on DL VoIP capacity of limiting the lowest applied MCS to \( \text{MCS}_{\text{Min}} \) and adding a second receive antenna for OTT and SPS VoIP capacity for OTT scheduling with eight PDCCH is hardly improved by those means. VoIP capacity for SPS is increased by almost 15\% if largest TB size is limited to four PRBs. This is reached by finding an optimal tradeoff between increased PER and decreased TB size reducing resource occupation. Similar results can theoretically
be reached with OTT and more PDCCHs.

OTT scheduling shows slightly higher VoIP capacity than SPS because long waiting HARQ retransmissions have highest scheduling priority. Preempting persistent reservations in favor of long waiting HARQ retransmissions achieves the same VoIP capacity with SPS. This is true for UL and DL. Further in the UL optimized transmit power and using resources providing highest SINR increases VoIP capacity by more than 10% while SPS surprisingly reduces it, see Figure 7.41b. SPS creates a very regular interference patterns allowing to apply LA and the BestSINR scheduling algorithm. OTT cannot benefit from LA due to high SINR variance introducing high PER.

VoIP capacity is determined by the DL supporting less calls than the UL. Figure 7.42 shows the VoIP capacity for scenarios UMa, UMi, and RMa for OTT and SPS equal to DL VoIP capacity with $\text{MCS}_{\text{min}} = 6$, Figures 7.20 and 7.22. OTT and SPS show almost identical performance with OTT supporting one more user / cell / MHz for all scenarios. Frequency selective scheduling as for the UL with LA is not feasible for DL due to limited control channel resources for CQI feedback [155]. Preempting persistent reservation in the DL would show same gain as for UL allowing to support one or two more users / cell / MHz. Control channel capacity limitations for OTT are not
considered here since only few more PDCCHs are required to achieve presented VoIP capacity (Figure 7.40b) and those are available since the system will not be fully loaded with VoIP traffic in reality [154].

VoIP capacity can hardly be improved by means of radio resource management compared to gains achieved by multi antenna systems.
The main focus of this work is on evaluation of VoIP capacity of LTE according to the guidelines specified by ITU-R for IMT-A systems [7] operated in various scenarios. It is established that VoIP capacity of LTE systems in all scenarios is higher for both, UL and DL transmission direction, than demanded by the evaluation guidelines [7]. The UMa scenario representing urban macro cells appears to be a challenge for LTE systems since capacity on DL 41 users / cell / MHz is just beyond IMT-A demand 40 users / cell / MHz. These results are gained from a system level event driven simulator partly developed to be able to perform this study. Additionally an analytic model based on Engset multi-server queues was developed allowing to determine VoIP capacity in the absence of packet errors.

It is found that VoIP users with low SINR condition (roaming at the cell edge) that require multiple PRBs to carry their VoIP packets limit VoIP system capacity, since close to capacity limit the number of PRBs requested appear not to be available in the current LTE subframe. It is not the packet error ratio experienced by VoIP users but the packet delay limit specified in the evaluation guidelines [7] that limits system capacity.

It appears that the number of control channels available in LTE is a capacity limiting factor in systems purely loaded by VoIP traffic, if VoIP packets are scheduled Over the Top (OTT) of the LTE radio interface protocol stack. This observation, however, appears to be of minor importance since a real world LTE system will mostly carry non-VoIP data like mobile video, Web browsing etc. These services result in much less load to LTE control channels compared to VoIP service so that no control channel bottleneck is expected in real systems. According
the current traffic load to mobile broadband systems like LTE-A is about 1% VoIP, 20% Web browsing and 67% mobile video, besides others.

One main focus in this study is comparison of VoIP capacity achieved under Semi-Persistent Scheduling (SPS) of VoIP packets (Voice over LTE (VoLTE)) specified as an option in standard LTE and of OTT scheduling. All results clearly indicate that OTT scheduling should be preferred to SPS, an observation that is really surprising given that much efforts have been put into standard LTE to “improve” system performance by providing SPS - which is a very complex technique as found when implementing it in the simulator. Allowing certified OTT VoIP application, e.g. Rich Communication Services (RCS) compliant ones, access to prioritized bearers managed by IMS would provide identical end-to-end VoIP capacity with less complexity on the radio link.

A number of techniques and VoIP packet scheduling strategies applied in LTE systems purely loaded by VoIP packets have been studied as part of this thesis and the main findings are summarized as follows:

Neither the order of choice of free PRBs to serve VoIP packets, nor the order of service of queued VoIP packets have much impact on the capacity. The reason is that close to the capacity limit of a system there is not much choices possible and the number of alternate orders of service also shrinks to a minimum.

As mentioned earlier edge users with low SINR threshold limit LTE VoIP capacity. Therefore it is investigated whether or not Soft Frequency Reuse (SFR) enhanced LTE systems would achieve higher capacity. The result is disappointing showing that SFR based systems achieve lower capacity than systems without this technology.

A next step is evaluating VoIP capacity with OTT RCS voice and presence of traffic from other applications in IMS enabled end-to-end scenarios.
Equation Proofs

Proof of Eq. (6.3):

\[ X \sim N(\mu_x, \sigma) \] (A.1)
\[ Y \sim N(\mu_y, \sigma) \] (A.2)
\[ Z = X - Y \sim N(\mu_x - \mu_y, \sqrt{\sigma_x^2 + \sigma_y^2}) \] (A.3)

\[ P(z > Z) = \int_{-\infty}^{Z} p(z)dz \]

\[ = 1 - \frac{1}{2} \text{erf} \left( \frac{Z - (\mu_x - \mu_y)}{\sqrt{\sigma_x^2 + \sigma_y^2}} \right) \]

\[ = 1 - \frac{1}{2} \text{erf} \left( \frac{\mu_y - \mu_x}{\sqrt{2}\sigma} \right), \text{for } Z = 0, \sigma_x = \sigma_y = \sigma \] (A.4)

Proof of Eq. (6.4) with x, y, z defined in Eq. (A.1)-(A.3), \( \mathbb{1} \) the indicator function and \( \phi(x|\mu, \sigma) \) the PDF of the normal distribution \( N(\mu, \sigma) \):

\[ p(z = x - y|x \geq y) = \frac{p(z = x - y, x \geq y)}{P(x \geq y)} \] (A.5)

\[ p(z, x \geq y) = \begin{cases} \phi(z|\mu_x - \mu_y, \sqrt{2}\sigma) : z \geq 0 \\ 0 : z < 0 \end{cases} \]

\[ = \phi(z|\mu_x - \mu_y, \sqrt{2}\sigma) \mathbb{1}_{\{z \geq 0\}} \] (A.6)

\[ p(z, x \geq y) = \frac{\phi(z|\mu_x - \mu_y, \sqrt{2}\sigma) \mathbb{1}_{\{z \geq 0\}}}{P(z \geq 0)} \] (A.7)
### LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a(x</td>
<td>y)$</td>
</tr>
<tr>
<td>$A$</td>
<td>Area of the scenario [$m^2$]</td>
</tr>
<tr>
<td>$A$</td>
<td>Continuous set of all positions in the scenario</td>
</tr>
<tr>
<td>$A$</td>
<td>UE to PRB assignment matrix</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Factor for uplink power control</td>
</tr>
<tr>
<td>$b$</td>
<td>EESM mapping parameter</td>
</tr>
<tr>
<td>$B$</td>
<td>Throughput</td>
</tr>
<tr>
<td>$\beta_C$</td>
<td>Scenario dependent path loss offset under channel condition $C$</td>
</tr>
<tr>
<td>$BW_{\text{Channel}}$</td>
<td>Channel bandwidth of the LTE cell</td>
</tr>
<tr>
<td>$c$</td>
<td>Index of eNB</td>
</tr>
<tr>
<td>$C_{c,s}$</td>
<td>Channel condition (LoS or NLoS) on link between UE $s$ and eNB $c$</td>
</tr>
<tr>
<td>$C$</td>
<td>Set of all eNBs</td>
</tr>
<tr>
<td>$d_{c,s}$</td>
<td>Distance between eNB $c$ and UE $s$</td>
</tr>
<tr>
<td>$d(x</td>
<td>y)$</td>
</tr>
<tr>
<td>$D$</td>
<td>State transition matrix considering only departures</td>
</tr>
<tr>
<td>$\Delta_{s,c}$</td>
<td>Regret of UE $s_c$ served by eNB $c$</td>
</tr>
<tr>
<td>$e(x</td>
<td>y)$</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Cell spectral efficiency</td>
</tr>
</tbody>
</table>
\( \Theta \) Distance dependent path loss offset for RMa scenario
\( \gamma \) SINR
\( \bar{\gamma} \) Effective SINR
\( \gamma_c \) Effective SINR hypermatrix of all UEs served by eNB \( c \) considering all possible combinations of interfering UEs served by other eNBs
\( h_{c,s,k} \) Channel gain from eNB \( c \) to UE \( s \) on PRB \( k \)
\( \bar{h}_{c,s} \) Mean channel gain from eNB \( c \) to UE \( s \)
\( \bar{h}_{PL,C,c,s} \) Path loss on link between UE \( s \) and eNB \( c \) under channel condition \( C \) [dB]
\( \bar{h}_{Car,s} \) Random car penetration loss of UE \( s \) [dB]
\( \bar{h}_{Sh,C,c,s} \) Random shadowing on link between UE \( s \) and eNB \( c \) under channel condition \( C \) [dB]
\( \bar{h}_{C,c,s} \) Total large scale channel loss on link between UE \( s \) and eNB \( c \) under channel condition \( C \) [dB]
\( k \) Index of PRB
\( k \) (Chapter 5) Number of trunks
\( K \) Set of PRBs
\( \Lambda_C \) Scenario dependent path loss coefficient under channel condition \( C \)
\( \Lambda \) Common matrix describing which PRBs form which uplink TBs in all cells
\( m \) Index of MCS
\( \mu_{Car} \) Mean car penetration loss [dB]
\( n_{TX} \) Number of transmit antennas
\( n_{RX} \) Number of receive antennas
$N_g$  Parameter to configure number of PHICH groups

$N_{RB}$  Number of PRBs

$N_{group}$  Number of PHICH groups

$\nu$  Common vector describing which PRBs form which uplink TBs in all cells

$\omega$  Spectrum allocated to the system

$p$ (Chapter 5) Transition probability from active to inactive and from inactive to active state

$P$  Transmit power

$P_0$  Offset for uplink power control

$P_a$  UE specific offset for downlink power control

$P_{\text{Noise}}$  Noise power

$P_{S,c}$  Received uplink signal power at eNB $c$

$P_{I,c}$  Received uplink interference power at eNB $c$

$P_{I+N,c}$  Received uplink interference plus noise power at eNB $c$

$P$ (Chapter 5) State transition matrix

$P_{c}$  PRB to TB assignment matrix of eNB $c$

$\pi(x)$  Steady state distribution vector

$Q$  State transition matrix for two dimensional Markov process conditioned on the event that at least one arrival occurs

$q(x|y)$  Probability for $x$ calls in the system in front of a tagged call if $y$ calls were active in the previous time slot

$q(x)$  Probability for a call to arrive in the system with $x$ other calls in front of it (waiting or being served)
Chapter A – Equation Proofs

\( r \)
Data rate

\( r_m \)
Data rate of MCS \( m \)

\( R^\Sigma \)
Cumulated throughput of all UEs

\( \mathbf{R}_c \)
Data rate hypermatrix of all UEs served by eNB \( c \) considering all possible combinations of interfering UEs served by other eNBs

\( s \)
Index of UE

\( s_c \)
UE \( s \) served by eNB \( c \)

\( S \)
Set of UEs

\( S_c \)
Set of UEs served by eNB \( c \)

\( \sigma_C \)
Scenario dependent standard deviation of shadowing under channel condition \( C \)

\( \sigma_{\text{Car}} \)
Standard deviation of car penetration loss [dB]

\( \varsigma \)
Assignment vector of UEs transmitting on the same resources

\( \mathbf{T}_c \)
UE to PRB assignment matrix of eNB \( c \)

\( t \)
Index of TB

\( \tau \)
Assignment vector of UEs to TBs in all cells

\( w(y) \)
Probability for the tagged call to arrive within the current arriving batch if \( y \) calls are active in the previous time slot

\( x \) (Chapter 5)
System state in current time slot

\( y \) (Chapter 5)
System state in previous time slot
# List of Figures

2.1 LTE Radio Frame ............................................ 8
2.2 PRB Structure ............................................. 10
2.3 LTE Protocol Stack .......................................... 12
2.4 Example TB transmitting three IP-PDUs and one IP-PDU-Segment in 14 PRBs using MCS with index 27. ................................. 14
2.5 Size of TB in bit over MCS index and number of PRBs. ........................................ 15
2.6 Number of CCEs for PDCCH transmission vs. number of PHICH ........................................ 22
3.1 VoIP capacity reported to ITU by 3GPP consortium and independent evaluators. ..................... 39
4.1 Cellular IMT-Advanced scenarios (UMa, UMi, RMa) with 21 and 57 sectors. ............................ 45
4.2 IMT-Advanced Indoor Hotspot (InH) Scenario ........................................ 45
4.3 LoS channel condition probability over distance ........................................ 47
4.4 Two-State Discrete Time Markov Chain (DTMC) VoIP Source ........................................ 54
5.1 VoIP Queuing Network Model ................................ 57
5.2 Steady state- and conditional distribution of number of active calls in VoIP queuing model, Figure 5.1 ........................................ 59
5.3 Simulated and analytic ratio of received packets rate versus offered load and resulting VoIP capacity. ........................................ 66
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>Mean of VoIP packets per subframe scheduled dynamically and ratio of failed dynamic scheduling attempts found by simulation. 1.4 MHz system bandwidth.</td>
</tr>
<tr>
<td>5.5</td>
<td>TB Size Distribution</td>
</tr>
<tr>
<td>5.6</td>
<td>Delay and loss rate for openWNS, QNS and analytic model for scenario UMa DL</td>
</tr>
<tr>
<td>5.7</td>
<td>Satisfied user ratio and loss rate distribution for openWNS and QNS for scenario UMa DL</td>
</tr>
<tr>
<td>6.1</td>
<td>Throughput versus SINR</td>
</tr>
<tr>
<td>6.2</td>
<td>Analytic and simulation results for the UE DL throughput distribution and CSE in the InH scenario.</td>
</tr>
<tr>
<td>6.3</td>
<td>Received signal and interference power at eNB versus path loss on both links from a UE: $\alpha = 0.8, P_0 = -83$ dBm</td>
</tr>
<tr>
<td>6.4</td>
<td>Comparison of analytical and simulated CDFs. $\alpha = 0.8, P_0 = -83$ dBm</td>
</tr>
<tr>
<td>7.1</td>
<td>Semi-persistent VoIP scheduler structure.</td>
</tr>
<tr>
<td>7.2</td>
<td>State diagram to evaluate the VoIP traffic source state in the MAC layer.</td>
</tr>
<tr>
<td>7.3</td>
<td>Origin of UL and DL Interference</td>
</tr>
<tr>
<td>7.4</td>
<td>ACF of estimated SINR for a UE measured on one PRB for 250 subframes in the UMa scenario with VoIP traffic only.</td>
</tr>
<tr>
<td>7.5</td>
<td>TB Size in PRBs versus SINR</td>
</tr>
<tr>
<td>7.6</td>
<td>Example TB candidate sets $\text{TB}<em>{0,UL}$ and $\text{TB}</em>{0,DL}$ for UL and DL within a single subframe.</td>
</tr>
<tr>
<td>7.7</td>
<td>Resource Assignment Filter Chain</td>
</tr>
<tr>
<td>7.8</td>
<td>NonEmpty Alternative filter selecting a non empty set if the preferred one ($\text{TB}_{(i+1),a}$) has no elements.</td>
</tr>
<tr>
<td>7.9</td>
<td>Mechanisms to remove packets that cannot be delivered within defined delay bounds.</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>7.10</td>
<td>Example set of preemptions for HARQ retransmission requiring three PRBs</td>
</tr>
<tr>
<td>7.11</td>
<td>DL VoIP capacity at 3 MHz bandwidth in a 21 and 57 cell simulation scenario</td>
</tr>
<tr>
<td>7.12</td>
<td>DL S, I, and SINR for Load 39 in Scenario UMi</td>
</tr>
<tr>
<td>7.13</td>
<td>SUR vs. number of users per cell in the InH scenario UL with small scale channel fading disabled and enabled and different number of receive antennas</td>
</tr>
<tr>
<td>7.14</td>
<td>Components of LTE Radio Link Delay</td>
</tr>
<tr>
<td>7.15</td>
<td>VoIP Capacity versus Queue Size; UMi Scenario; Load 54; UL</td>
</tr>
<tr>
<td>7.16</td>
<td>Performance versus relocation threshold in UL UMa scenario at load 48</td>
</tr>
<tr>
<td>7.17</td>
<td>Ratio of unsatisfied UL users (1 - SUR), PER and mean TB size in PRBs versus load</td>
</tr>
<tr>
<td>7.18</td>
<td>Basic Scheduler Filter Chain</td>
</tr>
<tr>
<td>7.19</td>
<td>DL VoIP Performance with and without control channel limitation and OTT scheduling</td>
</tr>
<tr>
<td>7.20</td>
<td>DL TB size distribution and SUR when forbidding low MCSs</td>
</tr>
<tr>
<td>7.21</td>
<td>Theoretical and actual DL VoIP capacity improvement if low MCSs are prohibited</td>
</tr>
<tr>
<td>7.22</td>
<td>DL VoIP capacity with SPS and prohibiting low MCSs</td>
</tr>
<tr>
<td>7.23</td>
<td>UL SUR versus power offset $P_0$ under 10 dB IoTN limit $\text{[7]}$</td>
</tr>
<tr>
<td>7.24</td>
<td>SINR versus UL power offset $P_0$ for all and the 2% weakest users of each cell</td>
</tr>
<tr>
<td>7.25</td>
<td>Correlation between WBL of a UE and its UL packet error rate</td>
</tr>
<tr>
<td>7.26</td>
<td>UL VoIP capacity with OTT and SPS</td>
</tr>
<tr>
<td>7.27</td>
<td>Filter chains for different resource assignment algorithms. From left to right: First, Random, Least-Fit, BestSINR</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Number of <strong>PRBs</strong> for different system bandwidths.</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Number of available <strong>REs</strong> out of 168 <strong>REs / PRB</strong> for user data if <strong>CCH</strong> region is three symbols long.</td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Preliminary <strong>DL</strong> 3GPP LTE VoIP Capacity Results from May 2007. Case 1 (500 m ISD) and Case 3 (1732 m ISD) Deployment Scenario</td>
<td>36</td>
</tr>
<tr>
<td>3.2</td>
<td>Preliminary <strong>UL</strong> 3GPP LTE VoIP Capacity Results from May 2007. Case 1 (500 m ISD) and Case 3 (1732 m ISD) Deployment Scenario</td>
<td>37</td>
</tr>
<tr>
<td>3.3</td>
<td>VoIP capacity of IMT-A scenarios UMa, UMi, and RMa as reported to ITU by 3GPP member companies.</td>
<td>38</td>
</tr>
<tr>
<td>4.1</td>
<td>Selected IMT-Advanced Evaluation Methodology Scenario Parameters.</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>Distance Dependent LoS Probability.</td>
<td>47</td>
</tr>
<tr>
<td>4.3</td>
<td>Path Loss Model Parameters.</td>
<td>48</td>
</tr>
<tr>
<td>7.1</td>
<td>Attributes of a Candidate <strong>TB</strong>.</td>
<td>97</td>
</tr>
<tr>
<td>7.2</td>
<td><strong>TB</strong> Candidate Set Filters.</td>
<td>98</td>
</tr>
</tbody>
</table>


Bibliography


[103] 3GPP, "Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN) (Release 7),” Tech. Rep. 3GPP 25913, 3GPP, June 2005.


<table>
<thead>
<tr>
<th>ACRONYMS</th>
<th>EXPLANATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>3G</td>
<td>3rd Generation</td>
</tr>
<tr>
<td>3GPP</td>
<td>3rd Generation Partnership Project</td>
</tr>
<tr>
<td>3GPP2</td>
<td>3rd Generation Partnership Project 2</td>
</tr>
<tr>
<td>ACF</td>
<td>Auto Covariance Function</td>
</tr>
<tr>
<td>AM</td>
<td>Acknowledged Mode</td>
</tr>
<tr>
<td>AMC</td>
<td>Adaptive Modulation and Coding</td>
</tr>
<tr>
<td>ACK</td>
<td>Acknowledgement</td>
</tr>
<tr>
<td>ARQ</td>
<td>Automatic Repeat Request</td>
</tr>
<tr>
<td>ATIS</td>
<td>Alliance for Telecommunications Industry Solutions</td>
</tr>
<tr>
<td>BCH</td>
<td>Broadcast Control Channel</td>
</tr>
<tr>
<td>BLER</td>
<td>Block Error Rate</td>
</tr>
<tr>
<td>BS</td>
<td>Base Station</td>
</tr>
<tr>
<td>BSR</td>
<td>Buffer Status Report</td>
</tr>
<tr>
<td>C-RNTI</td>
<td>Cell Radio Network Temporary Identity</td>
</tr>
<tr>
<td>CCDF</td>
<td>Complementary Cumulative Distribution Function</td>
</tr>
<tr>
<td>CCE</td>
<td>Control Channel Element</td>
</tr>
<tr>
<td>CCH</td>
<td>Control Channel</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CEUSE</td>
<td>cell edge user spectral efficiency</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>CQI</td>
<td>Channel Quality Indicator</td>
</tr>
<tr>
<td>CRC</td>
<td>Cyclic Redundancy Check</td>
</tr>
<tr>
<td>CSE</td>
<td>cell spectral efficiency</td>
</tr>
<tr>
<td>DCI</td>
<td>Downlink Control Information</td>
</tr>
<tr>
<td>DL</td>
<td>downlink</td>
</tr>
<tr>
<td>DLL</td>
<td>Data Link Layer</td>
</tr>
<tr>
<td>DM RS</td>
<td>Demodulation Reference Signal</td>
</tr>
<tr>
<td>DTMC</td>
<td>Discrete Time Markov Chain</td>
</tr>
<tr>
<td>E-UTRAN</td>
<td>Evolved Universal Terrestrial Radio Access Network</td>
</tr>
<tr>
<td>EESM</td>
<td>Exponential Effective SINR Metric</td>
</tr>
<tr>
<td>eNB</td>
<td>Enhanced Node-B</td>
</tr>
<tr>
<td>EDF</td>
<td>Earliest Deadline First</td>
</tr>
<tr>
<td>EPC</td>
<td>Evolved Packet Core</td>
</tr>
<tr>
<td>ESP</td>
<td>Encapsulating Security Payload</td>
</tr>
<tr>
<td>EV-DO</td>
<td>Evolution-Data Optimized</td>
</tr>
<tr>
<td>FDD</td>
<td>Frequency Division Duplex</td>
</tr>
<tr>
<td>FIFO</td>
<td>First in First Out</td>
</tr>
<tr>
<td>GPRS</td>
<td>General Packet Radio Service</td>
</tr>
<tr>
<td>GTP</td>
<td>GPRS Tunnelling Protocol</td>
</tr>
<tr>
<td>GTP-U</td>
<td>GPRS Tunnelling Protocol User Plane</td>
</tr>
<tr>
<td>HARQ</td>
<td>Hybrid Automatic Repeat Request</td>
</tr>
</tbody>
</table>
HOL  Head of Line
HSDPA  High Speed Downlink Packet Access
HSPA  High Speed Packet Access
HSUPA  High Speed Uplink Packet Access
IAT  inter arrival time
IETF  Internet Engineering Task Force
IMS  IP Multimedia Subsystem
IMT-A  IMT-Advanced
InH  Indoor Hotspot
IoTN  Interference over Thermal Noise
IP  Internet Protocol
ISD  Inter-site distance
ITU  International Telecommunication Union
ITU-R  International Telecommunication Union
          Radiocommunication Sector
LA  Link Adaptation
L1  Layer 1
L2  Layer 2
LoS  line-of-sight
LTE  Long Term Evolution
LTE-A  Long Term Evolution Advanced
MAC  Medium Access Control
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS</td>
<td>Modulation and Coding Scheme</td>
</tr>
<tr>
<td>MIESM</td>
<td>Mutual Information Effective SINR Metric</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple Input Multiple Output</td>
</tr>
<tr>
<td>MMPP</td>
<td>Markov-Modulated Poisson Process</td>
</tr>
<tr>
<td>MMSE</td>
<td>Minimum Mean Square Error</td>
</tr>
<tr>
<td>MRC</td>
<td>Maximum Ratio Combining</td>
</tr>
<tr>
<td>NACK</td>
<td>Negative Acknowledgement</td>
</tr>
<tr>
<td>NDI</td>
<td>New Data Indicator</td>
</tr>
<tr>
<td>NLoS</td>
<td>non line-of-sight</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>OFDMA</td>
<td>Orthogonal Frequency Division Multiple Access</td>
</tr>
<tr>
<td>openWNS</td>
<td>open source Wireless Network Simulator</td>
</tr>
<tr>
<td>OTT</td>
<td>Over the Top</td>
</tr>
<tr>
<td>PCFICH</td>
<td>Physical Control Format Indicator Channel</td>
</tr>
<tr>
<td>PCI</td>
<td>Protocol Control Information</td>
</tr>
<tr>
<td>PDCP</td>
<td>Packet Data Convergence Protocol</td>
</tr>
<tr>
<td>PDF</td>
<td>probability density function</td>
</tr>
<tr>
<td>PDU</td>
<td>Protocol Data Unit</td>
</tr>
<tr>
<td>PER</td>
<td>Packet Error Rate</td>
</tr>
<tr>
<td>PF</td>
<td>Proportional Fair</td>
</tr>
<tr>
<td>PGF</td>
<td>Probability Generating Function</td>
</tr>
<tr>
<td>PHY</td>
<td>Physical</td>
</tr>
</tbody>
</table>
SC-FDMA  Single Channel Frequency Division Multiple Access
SCH  Synchronization Channel
SDU  Service Data Unit
SFR  Soft Frequency Reuse
SID  Silence Indicator
SIMO  Single Input Multiple Output
SINR  Signal to Interference and Noise Ratio
SISO  Single Input Single Output
SPS  Semi-Persistent Scheduling
SRS  Sounding Reference Signal
SUR  Satisfied User Ratio
TB  Transport Block
TCP  Transmission Control Protocol
TDD  Time Division Duplex
TDMA  Time Division Multiple Access
TIA  Telecommunications Industry Association
TTA  Telecommunications Technology Association of Korea
TTI  Transmission Time Interval
UDP  User Datagram Protocol
UE  User Equipment
UL  Uplink
UM  Unacknowledged Mode
UMa  Urban Macro
UMi  Urban Micro
UMTS  Universal Mobile Telecommunications System
UTRAN  Universal Terrestrial Radio Access Network
VoIP  Voice over IP
VoLTE  Voice over LTE
VPN  Virtual Private Network
VRB  Virtual Resource Block
WBL  Wide Band Loss
WiMAX  Worldwide Interoperability for Microwave Access
WINNER+  Wireless World Initiative New Radio Plus
WTSC  Wireless Technologies and Systems Committee
CURRICULUM VITAE

Name: Maciej Grzegorz Mühleisen
Date of Birth: 1st of October, 1982
Place of Birth: Wrocław, Poland
Nationality: German and Polish

06/2002  German Abitur at Pius Gymnasium, Aachen

10/2002 - 01/2008 Dipl.-Ing. in Computer Engineering (Technische Informatik) at RWTH Aachen University

08/2006 - 02/2007 Internship at Philips Research, Eindhoven

02/2008 - 06/2012 Research engineer at Chair of Communication Networks / ComNets Research Group, RWTH Aachen University

since 07/2012 Senior research engineer at Institute of Communication Networks (ComNets), Hamburg University of Technology
This thesis is the result of seven and a half years research at ComNets Research Group, RWTH Aachen University, Germany. During this time I had the pleasure to work with Prof. Bernhard Walke. I am deeply grateful to him for his support, the inspiring discussions and the freedom in research. Furthermore, I like to thank Prof. Andreas Timm-Giel from the Institute of Communication Networks (ComNets), Hamburg University of Technology, Germany for his support as second examiner.

I am very grateful to Prof. Ulrich Killat from Hamburg University of Technology and Prof. Hans Daduna from Universität Hamburg for providing feedback to the mathematic model developed within this thesis.

I would like to thank all my colleagues at ComNets Hamburg for fruitful scientific discussions, feedback and patience during the last months when I was not able to fulfill all my duties at the institute. First to mention is of course Prof. Andreas Timm-Giel who always assured I had sufficient resources to pursue my thesis. I would also like to thank the researchers from the Vehicular- and Mobile Communication Group for the great collaboration during the last three years: Raphael Elsner, Leo Krüger, Ming Li, Christoph Petersen, and Rene Steinrücken.

Big thanks go to everyone I had the pleasure to work with for four and a half years at ComNets Aachen: Guido Hiertz and Erik Weiss who were my first supervisors during my student project; Guido Gehlen and Erik Weiss who supervised me as a student assistant; Ralf Jennen who was the supervisor of my diploma thesis and my mentor at the beginning of my scientific career; everyone who contributed to the openWNS simulator, the tool that made this thesis possible, especially Daniel Bült-
mann, Ralf Jennen, Karsten Klagges, Swen Kühnlein, Sebastian Max, Ralf Pabst, Holger Rosier, Klaus Sambale, and Marc Schinnenburg; the people I had the pleasure to work with in the WINNER+ Project Deniel Bülmann, Karsten Klagges, Klaus Sambale, and Benedikt Wolz.

Furthermore I would like to thank the students who contributed to this work through theses, projects and as student assistants: Jairo Alfonso Garcia Luna, Syed Ghazanfar Salari, Christoph Goertz, Kay Henzel, Adalbert Kutz, Catalina Mejia Rodriguez, Khondoker Mohammad Rahamatullah, Abdellahi Ould Abidine, Hossein Tajbakhsh, Mario Ugarkovic, and Sonia Yuste Lledo.

I would further like to thank Prof. Carmelita Görg, Mohammad Siddique and Bernd-Ludwig Wenning for the collaboration during the PoSSuM research project.

Thank you Norbert Konkol, Felix Okon, Doro Pawelzick, Gabi Reimann, Heinz Rochhausen, Anne Schröder, and Rosi Söhnen, for being the backbone of ComNets Aachen.

When ever there is a lot of work to do there must be relief after (or during :-) ) that and this is what friends and hobbies are for. For this I would like to thank the following people clubs and organizations: Kristin Birke, Max Brammertz, Jens Breuer, Maurice Breuer, Lars Brinkmann, Sven Brinkmann, Anne Bunde, Carsten Butsch, Silvan Conle, Erwin Dautzenberg, Raphael Elsner, Katharina Feinhals, Bernd Fischer, Dennis Förster, Jairo Garcia Alfonso Luna, Axel Gotzes, Jan Habermann, Philip Hammerschmidt, Yana Heinisch, Dominik (Feivel) Herff, Nina Holtermann, Tim Holtermann, Philipp Hoymann, Patrick Keulers, Alex (Kiesel) Kersting, Michael Kloeser, Marco Kowol, Leo Krüger, Adalbert Kutz, Meike Laufenberg, Philipp Mathews, Christian Meirich, Lutz Mierbach, Lucien Müllejans, Christoph Petersen, Bettina Ryschka, Rene Steinrücken, Sebastian (Schnebi) Schmitz, Nina Thomas, Hendrik (Moby) Wunden, Katholische Studierende Jugend (KSJ) Aachen, especially KSJ Pius, RWTH Aachen Hochschulsport, especially the Snowsport-Team, Hamburger Ingenieure der Elektro- und
Informationstechnik (HITECH) e.V., Fachschaftsrat Elektrotechnik und Informationstechnik (FSR ET/IT) TU Hamburg-Harburg, Apollo, B9, Das Ding, Himmerich, La Paloma, Nightlife, Starfish and of course most important: Alemannia Aachen, the best club now and for ever. Liebe kennt keine Liga!

I would like to express my appreciation for my parents Jadwiga and Tomasz for always being there for me and supporting me throughout my thesis.

Finally but most importantly I would like to thank my wife Patrycja for her almost never ending patience, love and support during the last years and especially the last months.

Aachen, July 2015

Maciej Mühleisen