



HAL
open science

Enhancing and improving voice transmission quality over LTE network : challenges and solutions

Duy Huy Nguyen

► To cite this version:

Duy Huy Nguyen. Enhancing and improving voice transmission quality over LTE network : challenges and solutions. Networking and Internet Architecture [cs.NI]. Institut National des Télécommunications, 2017. English. NNT : 2017TELE0002 . tel-01534513

HAL Id: tel-01534513

<https://theses.hal.science/tel-01534513>

Submitted on 7 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**DOCTORAT EN CO-ACCREDITATION
TÉLÉCOM SUDPARIS - INSTITUT MINES-TÉLÉCOM
ET L'UNIVERSITÉ PIERRE ET MARIE CURIE - PARIS 6**

Spécialité: Informatique et Réseaux

École doctorale EDITE

Présentée par

Duy Huy NGUYEN

**Enhancing and improving voice transmission
quality over LTE network: Challenges and
Solutions**

Soutenue le 24/02/2017 devant le jury composé de:

Jalel BEN OTH-MAN	Professeur, Université Paris 13	Rapporteur
Fabrice VALOIS	Professeur, Institut National des Sciences Appliquées (INSA) de Lyon	Rapporteur
Marcelo DIAS DE AMORIM	Professeur, Université Pierre et Marie Curie (Paris 6)	Examineur
Véronique VEQUE	Professeure, CNRS-CentraleSupélec-Université Paris-Sud	Examinatrice
Pierre DUHAMEL	Professeur, CNRS-CentraleSupélec-Université Paris-Sud	Examineur
Hang NGUYEN	Directrice d'études, Télécom SudParis	Encadrante de thèse
Éric RENAULT	Maître de conférences, Télécom SudParis	Directeur de thèse

Thèse numéro : 2017TELE0002

Acknowledgments

I would like to express my gratitude, appreciation and sincere thanks to my advisors Prof. Hang Nguyen and Prof. Éric Renault for their excellent guide, useful discussions, theoretical and technical suggestions and for their continuous encouragement during the whole period of my PhD thesis.

I am grateful to Mrs Xayplathi Lyfoung, Veronique Guy and Valérie Mateus and the other members of Télécom SudParis and Télécom École de Management for their help for administrative procedures.

I am greatly indebted to the Vietnamese government, Ministry of Education and Training for the scholarship to study abroad. I also thank Prof. Hang Nguyen, Prof. Éric Renault, Prof. Maryline LAURENT, Zahia ZENDAGUI and Télécom SudParis for the financial support to attend the conferences and for the working conditions.

I would like to thank all my colleagues and friends at Télécom SudParis and at Télécom École de Management for their support and the sharing of life's experiences. Their advices already encouraged me to strive towards my goal.

Finally, I deeply thank the constant encouragement and support for my whole family including my parents, my sisters and brothers, my cousins, especially my loved wife and son. I dedicate this thesis to all the members of my family.

Abstract

LTE is a broadband wireless standard derived from the UMTS network. LTE also known as one of the beyond 3G wireless network technology is designed to have a greater delivery service for multimedia, voice and video applications to end users. Transmitting these services in a good quality over the wireless networks is a big challenge to service providers, especially in case of real-time services such as voice and video calls. Originally, LTE was seen as a completely IP cellular system just for carrying data, so operators would be able to carry voice either by reverting to 2G/3G systems or by using VoIP in one form or another. VoLTE is an enhanced technology for voice and video transmission over LTE network. It is an IMS-based specification thus it enables the system to be integrated with the suite of applications that will become available on LTE. With VoLTE, the voice traffic goes over the high-speed data network instead of its voice network. In this way, VoLTE is expected to provide better sound quality, shorter latency and lower packet loss.

This thesis aims to present some new approaches to enhance, improve and predict speech transmission quality over LTE networks. Since LTE is All-IP network, the deployment of VoLTE is rather complex and ensuring VoLTE quality is a big challenge. For this reason, many different solutions have been deployed in the LTE systems. The proposed solutions offered in this thesis mainly focus on the following fields:

- First, we propose solutions to enhance LTE channel codec (coder and decoder). More specifically, at the transmitting side, we propose solutions for enhancing LTE channel coding. When any data including voice is transmitted over the air interface, it is coded by channel coding to protect them from environment effects such as noise, fading, etc. Channel coding is mandatory in LTE network and many different coders are used. In order to automatically choose the suitable channel code rate corresponding to each source codec mode and offer the best selection for all modes, we propose a dynamic rate adaptation algorithm for jointly source-channel code rate based on several given QoS parameters. This algorithm uses WB E-model for getting the R-factor. Then, this R-factor is mapped to MOS score. The MOS displays the user satisfaction for speech quality. Simulation results showed that the percentage of redundant bits generated by channel coding reduced significantly with an acceptable MOS reduction. The proposed algorithm has simple computational operations, so that, it can be applied to real-time applications such as voice calls for monitoring the speech quality. Otherwise, at the receiving side, we present a solution for enhancing LTE channel decoding. Specifically, we propose a new error correction function which is used in Log-MAP algorithm. This function has a very important role for decoding. It is based on the understanding of polynomial regression function. The proposed function has the approximated performance which is normally performed in BER (Bit Error Rate) and has a simpler computational complexity compared to the one of the original function of the Log-MAP algorithm. This allows reducing time consumption for channel decoding task and helps to decrease end-to-end delay for any data traffic, especially live stream such as voice or video calls, etc.
- Second, we propose solutions to improve downlink scheduling schemes in LTE networks. The proposed schedulers are called E-MQS, WE-MQS, and WE-MQS-VoIP

Priority which are based on user perception (i.e., MOS). In order to predict user satisfaction, we used E-model and WB E-model for narrowband audio users and wideband audio users, respectively. Both the E-model and the WB E-model predict voice quality without any reference to original signals. Thus they are very suitable for real-time services such as voice calls. Thanks to the presence of user perception in scheduling process, the system performance is increased significantly. The proposed schedulers have performance metrics such as delay, PLR, cell throughput, FI and SE improved significantly in comparison with several well-known schedulers including FLS, M-LWDF, and EXP/PF, etc. The performance evaluation is performed in a heterogeneous traffic with mobility in LTE-Sim software using featured voice codecs, including G.729 and AMR-WB for narrowband audio services and wideband audio services, respectively.

- Last, in order to evaluate the efficiency of the above proposed schedulers, we proposed new non-intrusive models for predicting voice quality in LTE networks. These models are the combination of the LTE-Sim software with E-model or WB E-model corresponding to narrowband audio quality or wideband audio quality. One of limitations of the E-model as well as the WB E-model is related to the way to measure their inputs exactly. In order to overcome this disadvantage, we propose to complement buffer jitter factor which is very important in evaluating voice quality into the E-model thus we have the extended or enhanced E-model. In order to predict wideband audio quality, as the LTE-Sim software only support G.729 codec which is used to simulate narrowband audio services, we propose to add the AMR-WB codec into this software. Although the proposed models do not have exactly the same user perception as the other intrusive models such as PESQ or PESQ-WB, however, they are very suitable for predicting quality of live stream such as voice calls because they do not refer to the original signals.

It can be concluded that through the thesis, the E-model and WB E-model are the backbones of the proposals. For the application of these two models, we have proposed several solutions to enhance, improve, and predict voice transmission quality in LTE networks. Most of the proposed solutions are implemented in the LTE-Sim software which allows simulating the whole LTE network that is rather similar to a real system. The preliminary results prove the applicability of our proposals to enhance LTE channel codec and MAC scheduling schemes as well as to measure user satisfaction in LTE networks.

Résumé

LTE est une norme sans fil à bande large provenant de la UMTS réseau. Egalement connu comme l'une des technologies au-delà de la 3G de réseau sans fil, LTE est conçu pour avoir un service de livraison pour les applications multimédia, de voix et de vidéo destinées aux utilisateurs finaux. Transmettre ces services dans une bonne qualité sur les réseaux sans fil est un grand défi pour les fournisseurs de services, spécialement pour les services en temps réel tels que les appels vocaux et vidéo. A l'origine, le LTE était considéré comme un système cellulaire complètement IP permettant uniquement de transporter des données. Les opérateurs seraient ainsi en mesure de transporter la voix soit en revenant au système 2G/3G ou en utilisant VoIP sous une forme ou une autre. VoLTE est une technologie améliorée pour la voix et la transmission vidéo sur le réseau LTE. C'est une spécification basée sur IMS, et permet ainsi au système d'être intégré à la suite d'applications qui seront disponibles sur le LTE. Avec le VoLTE, le trafic vocal va sur le réseau de données à haute vitesse à la place de son réseau de voix. De ce fait, VoLTE devrait fournir une meilleure qualité sonore, un temps de latence plus court et une perte de paquets inférieure.

Cette thèse vise à présenter de nouvelles approches pour améliorer et prédire la qualité de transmission de la parole sur les réseaux LTE. Etant donné que LTE est un réseau entièrement IP, le déploiement de VoLTE est assez complexe, et comment faire pour assurer la qualité de VoLTE est un grand défi. Pour cette raison, de nombreuses solutions différentes doivent être déployées dans les systèmes LTE. Les solutions proposées dans cette thèse portent principalement sur les points suivants:

- Tout d'abord, nous avons proposé de nouvelles solutions pour améliorer la chaîne de codage LTE. Plus précisément, du côté de l'émission, nous avons proposé des solutions pour améliorer le canal de codage LTE. Ainsi quand une donnée sur la voix est transmise à travers l'interface d'air, il est codé par codage en chaîne pour le protéger contre les effets de l'environnement tels que le bruit, l'altération, etc., Le codage en chaîne est indispensable dans le réseau LTE et différents codes sont utilisés. Afin de choisir automatiquement le taux de code de canal approprié correspondant à chaque mode de codec source et offrir le meilleur choix pour tous les modes, nous proposons un algorithme d'adaptation de débit dynamique pour débit de la chaîne de code source conjoint fondé sur plusieurs paramètres de qualité de service donné. Cet algorithme utilise WB E-model pour obtenir le facteur R. Ce facteur R est ensuite mappé au score MOS. Le MOS montre la satisfaction de l'utilisateur pour la qualité de la parole. Les résultats de simulation ont montré que le pourcentage de bits redondants générés par codage en chaîne réduit de manière significative avec une réduction MOS acceptable. L'algorithme proposé présente de simples opérations de calcul, de sorte qu'il peut être appliqué à des applications en temps réel telles que des appels vocaux pour surveiller la qualité de la parole. Par ailleurs, du côté de la réception, nous avons proposé une solution pour améliorer la chaîne de décodage LTE. Plus précisément, nous avons proposé une nouvelle fonction de correction d'erreur qui est utilisée dans l'algorithme Log-MAP. Cette fonction a un rôle très important pour le décodage. Elle est basée sur la compréhension de la fonction de régression polynomiale. La fonction proposée a la performance approximative qui est normalement réalisée en BER, et la complexité de calcul est plus simple par rapport à celui de la fonction originale de

l'algorithme Log-MAP. Cela permet de réduire le temps pour la tâche de décodage par chaîne et contribuera à diminuer le délai de mise en bout pour tout le trafic de données, particulièrement les flux vivants tels que les appels vocaux ou vidéo, etc.

- Deuxièmement, nous avons proposé des solutions pour améliorer les systèmes de planification de liaison descendante dans les réseaux LTE. Les planificateurs proposés sont E-MQS, WE-MQS ou WE-MQS-VoIP Priority qui sont basés sur la perception de l'utilisateur (i.e., MOS). Afin de prédire la satisfaction des utilisateurs, nous avons utilisés le E-model et le WB E-model pour les utilisateurs audio à bande étroite et à bande large respectivement. Tous les deux prédisent la qualité vocale sans faire référence aux signaux d'origine. Par conséquent, ils sont bien appropriés pour les services en temps réel tels que les appels vocaux ou vidéo. En cas de présence de la perception de l'utilisateur dans le processus de planification, la performance du système augmente de manière significative. Les planificateurs proposés ont les indicateurs de performance tels que le retard, le taux de PLR, le débit de la cellule, l'indice de l'équité et de l'efficacité spectrale améliorée de façon significative en comparaison avec plusieurs planificateurs bien connu, comme le FLS, M-LWDF, et EXP/PF, etc. L'évaluation de la performance est réalisée dans un trafic hétérogène avec mobilité dans le logiciel LTE-Sim en utilisant les codecs vocaux en vedette, y compris G.729 et AMR-WB pour les services audio à bande étroite et à large bande des services audio respectivement.
- Enfin, pour évaluer l'efficacité des planificateurs proposées ci-dessus, nous avons proposé de nouveaux modèles non intrusifs pour prédire la qualité de la voix sur les réseaux LTE. Ces modèles sont la combinaison du logiciel LTE-Sim avec E-model ou WB E-model correspondant à la qualité audio bandes étroite ou large. Une des limitations des E-model et WB E-model est liée à la façon de mesurer exactement leurs intrants. Afin de remédier à cet inconvénient, nous avons proposé de compléter le facteur de gigue de tampon qui est très important dans l'évaluation de la qualité de la voix dans E-model. Ainsi nous avons le E-model étendu ou amélioré. Pour prédire la qualité audio large bande, comme le logiciel LTE-Sim ne prend en charge que le codec G.729 qui est utilisé pour simuler des services audio à bande étroite, nous avons proposé d'ajouter le codec AMR-WB dans ce logiciel. Bien que les modèles proposés n'aient pas exactement pas la même perception de l'utilisateur que celle des modèles intrusifs, tels que PESQ ou PESQ-WB, ils sont cependant très appropriés pour prédire la qualité du flux en direct, tels que les appels vocaux, car ils ne se réfèrent pas aux signaux originaux.

On peut dire que, grâce à ce travail, le E-model et WB E-model sont des propositions cruciales. Pour l'application de ces deux modèles, nous avons proposé plusieurs solutions pour renforcer, améliorer et prévoir la qualité de transmission de la voix sur les réseaux LTE. La plupart des solutions proposées sont mises en œuvre dans le logiciel LTE-Sim qui permet de simuler l'ensemble du réseau LTE, qui est assez similaire à un système réel. Les résultats préliminaires démontrent l'applicabilité de nos propositions pour améliorer le codec de canal LTE et les systèmes de planification MAC ainsi que pour mesurer la satisfaction des utilisateurs dans les réseaux LTE.

List of Courses

1. Publish or perish (17, 24 June 2014, at Télécom ParisTech): 2 points
2. Writing and publishing research articles: Structure and strategies (29-30 January 2015, at Télécom SudParis): 2 points
3. Simplifying the writing of articles and presentations with LaTeX (24-25 March 2015, at Télécom ParisTech): 0 point
4. Publishing and peer review (13 April 2015, at Télécom ParisTech): 1 point
5. Scientific presentation course (23, 30 November 2015 and 7, 14 December 2015, at Télécom ParisTech): 3 points

List of Publications

- [1] Duy-Huy Nguyen, Hang Nguyen. *An improved Log-MAP algorithm based on polynomial regression function for LTE Turbo decoding*. In Proc. of 25th IEEE International Conference on Communications (ICC) 2015 - Workshop on Advanced PHY and MAC Techniques for Super Dense Wireless Networks: ICC'15 - Workshops 13, 08–12 June 2015, London, UK. DOI: 10.1109/ICCW.2015.7247502
- [2] Duy-Huy Nguyen, Hang Nguyen. *A dynamic rate adaptation algorithm for voice traffic over LTE network*. In Proc. of 8th IFIP/IEEE Wireless Days (WD), 23–25 March 2016, Toulouse, France. DOI: 10.1109/WD.2016.7461490
- [3] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *A new non-intrusive assessment method for VoLTE quality based on extended E-model*. In Proc. of 2nd Springer International Conference on Mobile, Secure and Programmable Networking (MSPN), 01–03 June 2016, Paris, France.
- [4] Manh-Cuong Nguyen, Hang Nguyen, Duy-Huy Nguyen, Eric Georgeaux, Philippe Mege, and Laurent Martinod. *Adaptive Physical Resource Block Design for Enhancing Voice Capacity over LTE network in PMR Context*. In Proc. of 27th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 04-07 September 2016, Valencia, Spain.
- [5] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *E-MQS – A new Downlink scheduler for Real-time flows in LTE network*. In Proc. of 84th IEEE Vehicular Technology Conference (VTC2016-Fall), 17–21 September 2016, Montréal, Canada.
- [6] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *WE-MQS: A new LTE Downlink Scheduling scheme for voice services based on user perception*. In International Journal of Computer Applications (IJCA). Volume 142, number 10, 2016. ISSN: 0975-8887. Publisher: Foundation of Computer Science (FCS), NY, USA. DOI: 10.5120/i-jca2016909948
- [7] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *WE-MQS-VoIP Priority: An enhanced LTE Downlink Scheduler for voice services with the integration of VoIP priority mode*. In International Journal of Advanced Computer Science and Applications (IJACSA). Volume 7, issue 7, 2016. ISSN: 2158-107X (Print), 2156-5570 (Online). Publisher: The Science and Information (SAI) Organization, West Yorkshire, United Kingdom. DOI: 10.14569/IJACSA.2016.070776

- [8] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *Predicting VoLTE Quality using Random Neural Network*. In International Journal of Applied Information Systems (IJ AIS). Volume 11, number 3, 2016. ISSN: 2249-0868. Publisher: Foundation of Computer Science (FCS), NY, USA. DOI: 10.5120/ijais2016451587
- [9] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *A New Non-Intrusive Model for Measuring VoLTE Quality based on Wideband E-Model*. In Communications on Applied Electronics (CAE). Volume 5, number 8, 2016. ISSN: 2394-4714. Publisher: Foundation of Computer Science (FCS), NY, USA. DOI: 10.5120/cae2016652333
- [10] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *A New Channel- and QoS-Aware Scheduling Scheme for Real-Time Services in LTE Network*. In International Journal of Applied Information Systems (IJ AIS). Volume 11, number 4, 2016. ISSN: 2249-0868. Publisher: Foundation of Computer Science (FCS), NY, USA. DOI: 10.5120/ijais2016451601
- [11] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *Enhancing and improving voice transmission quality over LTE network: Challenges and Solutions*. PhD thesis summary. In Journées non thématiques RESCOM 2017, 12-13 January 2017, Valbonne, France. (presented)
- [12] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *Performance evaluation of E-MQS scheduler with Mobility in LTE heterogeneous network*. In 2017 IEEE International Conference on Communications (ICC 2017), 21-25 May 2017, Paris, France. (accepted)
- [13] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *Enhancing LTE channel codec based on wideband E-model and polynomial regression function*. In Wireless Networks. (submitted)
- [14] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *A new LTE downlink scheduler for voice traffics based on user perception and VoIP priority mode*. In Mobile Networks and Applications. (submitted)
- [15] Duy-Huy Nguyen, Hang Nguyen and Éric Renault. *A novel object non-intrusive model for predicting VoLTE quality*. In Annals of Telecommunications. (submitted)

List of Acronyms

2G	The Second Generation
3G	The Third Generation
3GPP	The Third Generation Partnership Project
4G	The Forth Generation
ACS	Add-Compare-Select
AMC	Adaptive Modulation and Coding
AMR	Adaptive Multi-Rate
AMR-WB	AMR Wideband
ANIQUE+	Auditory Non-Intrusive Quality Estimation Plus
ANN	Artificial Neural Networks
APL	Application Layer
ARP	Allocation Retention Priority
ARQ	Automatic Repeat Request
AWGN	Additive White Gaussian Noise
BATD	Buffer-Aware Traffic-Dependent
BEM	Bandwidth Expansion Mode
BER	Bit Error Rate
BET	Blind Equal Throughput
BPSK	Binary Phase-Shift Keying
C/I	Carrier-to-Interference
CBRM	Circular Buffer Rate Matching
CDMA	Code Division Multiple Access
CMOS	Complementary Metal-Oxide-Semiconductor
CODEC	Coder and Decoder
CQI	Channel Quality Indicator
CRC	Cyclic Redundancy Check

CS	Circuit Switching
CSFB	Circuit-switched Fall Back
DHS	Dynamic Hybrid Scheduler
DL	Downlink
DL-SCH	Downlink Shared CHannel
DRX	Discontinuous Reception
DS	Dynamic Scheduler
EnodeB/eNB	Evolved Node B
EPC	Evolved Packet Core
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
EXIT	EXtrinsic Information Transfer
EXP/PF	Exponential/Proportional fair
FDD	Frequency Division Duplex
FDPS	Frequency Domain Packet Scheduler
FEC	Forward Error Correction
FGMM	Fuzzy Gaussian Mixture Model
FI	Fairness Index
FIFO	First In First Out
FLS	Frame Level Scheduler
FNN	Fuzzy Neural Network
FTP	File Transfer Protocol
GBR	Guaranteed Bit-Rate
GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service
HARQ	Hybrid ARQ
HD	High Definition
HOL	Head Of Line
HSDPA	High Speed Downlink Packet Access
HSS	Home Subscriber Server

IMS	IP Multimedia Subsystem
IP	Internet Protocol
ITU	International Telegraph Union
ITU-T	ITU Telecommunication Standardization Sector
JB	Jitter Buffer
LLR	Log Likelihood Ratio
Log-MAP	Logarithmic MAP
LSQPP	Largest-Spread QPP
LTE	Long Term Evolution
LTE-Sim	LTE Simulator
LUT-Log-BCJR	Look-UpTable-Log-BCJR
LWDF	Largest Weighted Delay First
MAC	Medium Access Control
MAP	Maximum A Posteriori
Max-Log-MAP	Maximum Log-MAP
MBR	Maximum Bit Rate
MCH	Multicast CHannel
MCS	Modulation and Coding Scheme
MLBS	Mean Loss Burst Size
M-LWDF	Maximum-Largest Weighted Delay First
MME	Mobility Management Entity
MNO	Mobile Network Operator
MOS	Mean Opinion Score
MQS	Maximum Queue Size
MSQPP	Maximum-Spread QPP
MT	Maximum Throughput
NU	Number of User
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiplexing Access

PCCC	Parallel Concatenated Convolution Code
PCH	Paging Channel
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PDSCH	Physical Downlink Shared Channel
PDU	Protocol Data Unit
PESQ	Perceptual Evaluation of Speech Quality
PESQ-WB	Perceptual Evaluation of Speech Quality Wideband
P-GW	Packet Data Network Gateway
PHY	Physical
PLP	Perceptual Linear Predictive
PLR	Packet Loss Rate
POLQA	Perceptual Objective Listening Quality Assessment
PRB	Physical RB
PS	Packet switching
PSS	Priority Set Scheduler
QAM	Quadrature amplitude modulation
QCI	QoS Class Identifier
QL	Queue Length
QoE	Quality of Experience
QoS	Quality of Service
QPP	Quadratic Permutation Polynomial
QPSK	Quadrature Phase-Shift Keying
RAD-DS	Required Activity Detection with Delay Sensitivity
RAN	Radio Access Network
RB	Resource Block
RCIC	Rate-Compatible Insertion Convolutional
RF	Radio Frequency
R-factor	Transmission Quality Rating

RLC	Radio Link Control
RM	Rate Mactching
RNC	Radio Network Controller
RNN	Random Neural Network
RoHC	Robust of Header Compression
RR	Round Robin
RRC	Radio Resource Control
RRM	Radio Resource Management
RTP	Real-time Transport Protocol
RTT	Round Trip Time
SAE	System Architecture Evolution
SC-FDMA	Single-Carrier FDMA
SDU	Service Data Unit
SE	Spectral Efficieny
SGSN	Serving GPRS Support Node
S-GW	Serving Gateway
SID	Silent Insertion Descriptor
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Soft Input Soft Output
SMS	Short Message Service
SNR	Signal-to-Noise Ratio
SOVA	Soft Output Viterbi Algorithm
SPS	Semi Persistent Scheduling
SRVCC	Single Radio Voice Call Continuity
SVC	Scalable Video Coding
SV-LTE	Simultaneous voice over LTE
SVR	Support Vector Regression
TBS	Transport Block Size
TDD	Time Division Duplex

TDM	Time Division Multiplexing
TDPS	Time Domain Packet Scheduler
TTI	Transmission Time Interval
UDP	User Datagram Protocol
UE	User Equipment
UL	Uplink
UL-SCH	Uplink Shared CHannel
UMTS	Universal Mobile Telecommunications System
VLSI	Very-Large-Scale Integration
VoIMS	Voice over IMS
VoIP	Voice over IP
VoLTE	Voice over LTE
WB	Wideband
WCDMA	Wideband CDMA
WFQ	Weight Fair Queuing
WWW	World Wide Web

Table of contents

Introduction	5
Motivation	5
Contributions and outline	8
Chapter 1: Overview of LTE network and E-model	8
Chapter 2: State-of-the-art	8
Chapter 3: The proposed solutions based on wideband E-model and polynomial regression function for enhancing LTE channel codec	8
Chapter 4: The proposed LTE Downlink packet schedulers for voice traffics based on user perception and VoIP-priority mode	9
Chapter 5: The proposed object non-intrusive models for predicting voice quality in LTE networks	10
Chapter 6: Conclusions and perspectives	10
1 An overview of LTE technology and E-model	11
1.1 Overview of LTE technology	12
1.1.1 LTE architecture	12
1.1.2 Quality of Service	12
1.1.3 Enhanced technologies	13
1.2 VoLTE	15
1.2.1 Strategies for VoLTE	15
1.2.2 VoLTE source codec	18
1.2.3 VoLTE radio protocol stack	18
1.3 Overview of the E-model and the WB E-model	24
1.3.1 E-model - Speech quality assessment for narrowband audio	24
1.3.2 WB E-model - Speech quality assessment for wideband audio	26
1.4 Summary	27
2 State-of-the-art	29
2.1 LTE Turbo codec	30
2.1.1 LTE Turbo coding	30
2.1.2 LTE Turbo decoding	31
2.2 LTE Downlink scheduling schemes	35
2.2.1 Channel-unaware scheduling strategies	35
2.2.2 Channel-aware and QoS-unaware scheduling strategies	36

2.2.3	Channel-aware and QoS-aware scheduling strategies	38
2.2.4	Dynamic and Semi-persistent Scheduling for VoIP support	39
2.2.5	Energy-aware	40
2.3	QoE evaluation approaches for voice user	40
2.3.1	Standard models	41
2.3.2	Non-standard models	43
2.4	Summary	45
3	The proposed solutions based on wideband E-model and polynomial regression function for enhancing LTE channel codec	47
3.1	General idea	48
3.2	The proposed algorithm for minimizing redundant bits generated by channel coding	49
3.2.1	The calculation of the delay impairment factor	49
3.2.2	The calculation of the equipment impairment factor	50
3.2.3	The proposed joint source-channel code rate adaptation algorithm	52
3.2.4	Simulation results and Performance evaluation	53
3.3	Enhancing channel decoding algorithm based on polynomial regression function	57
3.3.1	Decoding algorithms in Turbo codes	57
3.3.2	The proposed Log-MAP algorithm based on polynomial regression function	59
3.3.3	Simulation results and Performance evaluation	63
3.4	Summary	65
4	The proposed LTE Downlink packet schedulers for voice traffics based on user perception and VoIP-priority mode	67
4.1	General idea	68
4.2	LTE downlink scheduling	69
4.2.1	Simple scheduling diagram	69
4.2.2	The correlated scheduling algorithms	70
4.3	The calculation of input parameters of the E-model and the WB E-model	72
4.3.1	An extension of the E-model	72
4.3.2	Calculation of the input parameters of the WB E-model	73
4.4	The proposed scheduling schemes	73
4.4.1	The proposed scheduling scheme based on user perception for narrow-band audio services	73
4.4.2	The proposed scheduling scheme for wideband audio services	74
4.5	Simulation environment and Performance evaluation	80
4.5.1	Simulation environment	80
4.5.2	Performance measures	82
4.5.3	Performance evaluation	84
4.5.4	Comparison of three proposed scheduling schemes	94
4.6	Summary	95
5	The proposed object non-intrusive models for predicting voice quality in LTE networks	97
5.1	General idea	98

5.2	The proposed object non-intrusive models for predicting voice quality in LTE networks	98
5.2.1	The implemented model for predicting narrowband audilo quality . . .	98
5.2.2	The implemented model for predicting wideband audio quality	99
5.3	Simulation environment and Performance evaluation	101
5.3.1	Simulation environment	101
5.3.2	Performance evaluation of the proposed model for predicting narrow-band audio quality	101
5.3.3	Performance evaluation of the proposed model for predicting wide-band audio quality	105
5.4	Summary	111
6	Conclusions and Perspectives	113
6.1	Conclusions	113
6.2	Perspectives	117
	Bibliography	118
A	Effects of the MQS on system performance	133
B	Performance evaluation of the E-MQS scheduler in LTE heterogeneous with mobility	135

List of Figures

1.1	LTE architecture deployment	12
1.2	The bearer concept	13
1.3	TTI Bundling of 4 TTIs	14
1.4	Example of Semi-Persistent Scheduling	15
1.5	Different strategies for LTE voice service provision	16
1.6	Voice in Legacy networks	17
1.7	VoLTE with IMS	17
1.8	The implemented protocol stack in the LTE-Sim software	19
1.9	Resource grid	21
1.10	Structure of the Turbo encoder (the dotted lines are for trellis termination)	22
1.11	Structure of the Turbo decoder	23
1.12	Turbo decoder with the rate matching module	24
1.13	Reference connection of the E-model	25
3.1	$I_{e,wb}$ vs. Packet loss for nine modes of AMR-WB codec	51
3.2	A proposed simple framework of voice traffic over LTE network	52
3.3	MOS vs. AMR-WB mode and channel code rate when p_r is fixed equal to 10%	55
3.4	MOS vs. AMR-WB mode and channel code rate when p_r is randomly changed in range of 0..15%	56
3.5	AMR-WB mode vs required bandwidth when p_r is fixed equal to 10% (left figure) and when p_r is randomly changed in range of 0..15% (right figure)	57
3.6	The proposed approximated function for the correction function in the Log-MAP algorithm	60
3.7	The comparison of the approximations of the correction function	61
3.8	The simulation model	62
3.9	BER performance of the proposed Log-MAP algorithm and the other Log-MAP based algorithms with the block size $N = 1024$	63
3.10	BER performance of the proposed Log-MAP algorithm and the other Log-MAP based algorithms with the block size $N = 512$	64
3.11	BER performance of the proposed Log-MAP algorithm and the other Log-MAP based algorithms at low SNRs with the block size $N = 1024$	64
3.12	BER performance of the proposed Log-MAP algorithm with $n = 2, 3, 4$ and the Log-MAP algorithm at low SNRs with the block size $N = 1024$	65

4.1	Simplified model of a packet scheduler	70
4.2	Implemented model of the proposed scheduling scheme	75
4.3	Implemented model of the proposed scheduling scheme	78
4.4	The proposed scheduling algorithm	81
4.5	Implemented model of the proposed scheduling scheme	82
4.6	Delay vs number of VoIP user	85
4.7	PLR vs number of VoIP user	86
4.8	Cell throughput vs number of VoIP user	87
4.9	FI vs number of VoIP user	87
4.10	Spectral efficiency vs number of user	88
4.11	Delay vs number of VoIP user	89
4.12	Packet loss rate vs number of VoIP user	90
4.13	Throughput vs number of VoIP user	90
4.14	Fairness Index vs number of VoIP user	91
4.15	Spectral efficiency vs number of user	91
4.16	Delay vs number of VoIP user	92
4.17	Packet loss rate vs number of VoIP user	93
4.18	Throughput vs number of VoIP user	93
4.19	Fairness Index vs number of VoIP user	94
4.20	Spectral efficiency vs number of VoIP user	94
5.1	The proposed model for predicting narrowband audio quality	99
5.2	The proposed model for predicting wideband audio quality	100
5.3	MOS vs number of VoIP user at the speed of 30 km/h	103
5.4	MOS vs number of VoIP user at the speed of 120 km/h	103
5.5	Effects of Delay on voice quality at the speed of 30 km/h	104
5.6	Effects of Delay on voice quality at the speed of 120 km/h	104
5.7	Effects of PLR on voice quality at the speed of 30 km/h	105
5.8	Effects of PLR on voice quality at the speed of 120 km/h	106
5.9	MOS vs number of VoIP user with AMR-WB mode 0	107
5.10	MOS vs number of VoIP user with AMR-WB mode 1	108
5.11	MOS vs number of VoIP user with AMR-WB mode 2	109
5.12	Average MOS vs the number of VoIP user	109
5.13	Effects of Delay and the number of VoIP user on voice quality	110
5.14	Effects of PLR and the number of VoIP user on voice quality	110
A.1	The effect of the MQS on delay of VoIP flow	133
A.2	The effect of the MQS on PLR of VoIP flow	133
A.3	The effect of the MQS on cell throughput of VoIP flow	134
B.1	Delay vs VoIP users at user speed of 120 km/h	135
B.2	PLR vs VoIP users at user speed of 120 km/h	136
B.3	Cell throughput vs VoIP users at user speed of 120 km/h	136
B.4	FI vs VoIP users at user speed of 120 km/h	137
B.5	SE vs the number of user at user speed of 120 km/h	137

List of Tables

1.1	LTE service classes with QoS requirements	14
1.2	Packet sizes of AMR-WB codec modes	18
1.3	R-factor and MOS with corresponding user satisfaction	26
3.1	The detailed results of figure 3.3	53
3.2	The detailed results of figure 3.4	56
4.1	Values of $I_{e,wb}$ and B_{pl} according to AMR-WB modes	73
4.2	Basic simulation parameters	83
4.3	Comparison of three proposed schedulers	95
5.1	Basic simulation parameters	102
5.2	The detailed results of Figures of 5.5 - 5.8	106
5.3	The detailed results of Figures of 5.13 - 5.14	111

Introduction

Motivation

An LTE mobile network is based on an All-IP network. Thus, it does not support circuit switching method which is used to supply voice call services in 2G/3G networks. In order to deploy a voice communication service, additional technologies are required. The VoLTE service was developed to provide voice and video communication and SMS on the LTE networks. According to [1], until April 09 2015, in the world, there were 393 commercially launched LTE networks in 138 countries and 16 operators commercially launched VoLTE-HD voice in 7 countries.

VoLTE is a VoIP-based multimedia service, in which, voice calls and video conference services can be provided. The basic difference between VoLTE and VoIP over LTE is that VoLTE traffic is guaranteed with given QoS parameters and is delivered over dedicated bear which has the highest priority while VoIP traffic is delivered over default bearer which only preserves “best effort” QoS.

With VoLTE, the end user can enjoy voice and LTE data speeds simultaneously on a single carrier. Since operators are currently relying on their 2G/3G CS networks for voice services, there is no way to offer voice and LTE data simultaneously without the use of dual-transceiver devices capable of simultaneously access to both 2G/3G network and LTE network [2]. While there are operators who have chosen such an approach, known as SV-LTE, it represents a challenge when it comes to device complexity, cost and power consumption. Operators who are using their 2G/3G networks for voice traffic are foregoing offering the SV-LTE and instead making use of the CSFB feature which combines the 2G/3G CS voice network with the LTE data network. Paging messages for voice calls from 2G/3G network are delivered to a device currently on the LTE network allowing that device to postpone data services on the LTE network in order to switch over to 2G/3G CS network to accept the voice call. While the simultaneous voice and LTE data capability is lost and network complexity is increased, these devices can be based on cost and power-efficient single transceiver architectures.

When deploying VoLTE, mobile network operators have to face many difficulties and challenges. It is important to examine the challenges that VoLTE brings to LTE. Traditionally, a voice service was carried over a CS network that allocated dedicated trunk and radio resources for the entire duration of the call. Those resources were dedicated to the call and not used for any other purposes. On the contrary, VoLTE uses PS for carrying the voice traffic end to end, as well as for the signaling, over an all-IP network. In packet switching, resources are requested and allocated on demand; they are not held for the duration of the call. This makes packet switching much more efficient than circuit switching because it can dynamically assign resources as needed. However, it also makes the underlying bearer less

predictable in terms of availability. In order to overcome this issue, mechanisms are defined in LTE specifications, such as SPS and TTI bundling, so that, VoLTE can be supported to deliver the same level of service quality as traditional circuit switch voice. For mobile operators, in order to ensure that VoLTE provides the same high level service that 2G/3G CS voice provides today, three following benchmarks have to be met:

- Call completion and call retention rates of 99% or better
- Total end-to-end delay of less than 300 ms
- Voice quality that is as good as or better than 3G circuit switch voice (as measured by PESQ, POLQA, and MOS).

Since 3GPP has standardized IMS based VoLTE for voice over LTE and SRVCC for call continuity, IMS adaption has been very slow mainly because of cost issues. VoLTE integrates VoIP, LTE radio network (i.e., E-UTRAN), LTE core network (i.e., EPC), and the IMS to support voice services. VoLTE is an aggregation of multiple technologies, protocols and implementation scenarios [3]. IMS is designed for call session control not only for LTE, but other network technologies as well, including UMTS, CDMA2000, WiFi and even wired networks. So there must be the combination of interworking VoLTE with traditional circuit-switched voice.

According to [3], the above challenges can be divided into three categories as follows:

- **(1) Challenges of Technology:**

Due to the fact that VoLTE is based on IMS, there are many new protocols which are related to IMS such as IPv6, SigComp, IPSec and P-headers that make matters worse. The integration of the LTE protocol stack with the IMS control layer is to be taken care of and end-to-end IMS signaling must be tested over the LTE access network. In addition, implementation of mobility between the packet switched LTE and the circuit switched networks is also an issue.

- **(2) Challenges of Implementation:**

Standards for voice services over LTE based on the 3GPP IMS architecture are still maturing. It would take time for subscribers based on LTE to be anywhere close to 2G/3G networks. Therefore, it is expected that the operators would look for temporary solutions before moving on to a full-fledged IMS architecture.

- **(3) Challenges of User satisfaction:**

Performance of the network for crucial real-time services need to be tested with real-time audio and video quality measurement tools. Network impairment simulation may be carried out to test voice call quality by inserting errors in the application data stream. Depending on the VoLTE solution deployed by the operator, there may be delay in call set up and degradation in quality. The service providers would therefore face a difficult choice in the short term and long term. Since IMS-based VoLTE has the support of 3GPP, the operators need to have a plan towards full IMS services.

VoLTE is expected to provide high quality speech, since LTE is based on an all-IP network, it requires voice calls to traverse the network utilizing VoIP. VoIP calls struggle to guarantee service levels, that's exactly what's required if MNOs are going to match the service quality level of existing CS calls. As a result, ensuring the VoLTE quality is a big challenge. Problems of VoLTE quality fall into three categories as follows:

- **(1) Call setup failures:** VoLTE utilizes IMS to setup the VoIP sessions which is extremely complex.
- **(2) End-to-end QoS:** Congested cells make havoc of VoIP calls, therefore, the dedicated bearer to deliver guaranteed bandwidth even when the rest of the cell is congested must be monitored.
- **(3) Cell Edge Issues:** LTE networks are still wireless networks and have the same RF issues as other wireless networks. Therefore, when a VoLTE user is on the edge of the cell, interference and noise often increases above acceptable thresholds.

MNOs have automated LTE call trace and bandwidth verification tools to manage the first two categories above, RF optimization is still very much a manual task. Today, VoLTE is optimized in the same manner voice calls have been optimized for years, VoIP calls are much more susceptible likely to higher levels of bit errors caused by noise and interference that change constantly on the edge of LTE cells. Existing optimization methods such as drive testing and walk around are not sufficient enough to meet the quality threshold required for VoLTE calls. The problem is: VoLTE users on the edge of the cell can experience a worse experience than VoLTE users in the rest of the cell.

Based on the above analyses, we see that ensuring voice call quality in VoLTE is a big challenge. When voice traffic is delivered over LTE network, it is affected by many network impairment factors such as delay, packet loss, jitter, etc. Especially, when it is transmitted over a noisy channel, it is distorted by noise, interference, etc. So its quality is decreased. In order to protect data when it is transmitted over a wireless channel, channel coding is used at the Physical layer. In order to compress audio data at the Application layer, source coding is used. Actually, the quality of VoLTE depends on many factors such as source codec, network architecture, channel codec, bandwidth and configurations of operators, etc. New optimization solutions are required in order to enhance voice transmission, improve and automatically monitor VoLTE call quality for every user in real-time. The platform needs to provide real-time incremental network adjustments to solve these issues with a real-time feedback mechanism that verify results.

For all those reasons, this thesis mainly proposes solutions to enhance and to improve voice transmission quality over LTE networks. Our proposals mainly concentrate on dressing some challenges of the third group of VoLTE challenges and the second and the third categories is dedicated to VoLTE quality issues.

This thesis is organized into six Chapters with three main Chapters corresponding to three contribution parts. The first part is related to the channel codec in LTE networks. It presents several solutions to enhance channel coding as well as channel decoding. The second part aims to improve scheduling schemes at MAC layer to reach an enhancement of QoS for voice services. Three scheduling schemes are made in contributions to this part. The third part focuses on the solutions for measuring voice quality over LTE networks. To archive this goal, two object non-intrusive models are proposed. The structure of the thesis is described as follows.

Contributions and outline

Chapter 1: Overview of LTE network and E-model

The objective of this chapter provide a brief overview of the related contents such as LTE technology, VoLTE, E-model as well as WB E-model. Firstly, we briefly introduce to LTE architecture and characteristics. We describe the main entities of this architecture and its main functions. Then we briefly introduce to QoS requirements in LTE networks. In particular, since LTE is an All-IP network, each service type has a specify bearer. This means each bearer has its QoS requirements. Next, we describe some enhanced technologies used in LTE systems such as Transmission Time Interval Bundling and Semi-Persistent Scheduling. Also in this chapter, an overview of VoLTE is introduced. Specifically, we focus on deployment strategies for VoLTE. Then, the radio protocol stack of VoLTE is also described. We also briefly describe all layers in this stack protocol to clear the process of voice transmission in LTE networks.

Next, this chapter also describes concisely the voice source codec (AMR-WB) which is proposed to complement into the LTE-Sim software in some our contributions. The E-model and WB E-model are also described in this chapter because they are backbones of the proposals in this thesis.

Chapter 2: State-of-the-art

In this chapter, we present a literature review on state-of-the-art of the issues which will be solved in this thesis. The goal is to find out new trends for solving those problems. We first present a literature review on channel coding and decoding, especially using WB E-model for joint source-channel coding adaptation rate. Next, several channel decoding algorithms which are based on MAP algorithm are described. This work allows finding out trends for channel coding and decoding in LTE networks. This is the base to propose new solutions in Chapter 3. Second, we present the state-of-the-art of scheduling strategies in LTE downlink direction. For this work, we want to find out a strategy that is suitable for voice services and recent trends of scheduling strategies. We also classify those scheduling strategies into groups based on their characteristics and roles. We also analyze “pros and cons” of each of them. All this information is really useful for developing and presenting our scheduling schemes in Chapter 4. Last, approaches of QoE measurement and their applicability in LTE network are discussed. Through this work, we can find out a suitable method for predicting voice quality over LTE networks. This is served for proposals in Chapter 5.

Chapter 3: The proposed solutions based on wideband E-model and polynomial regression function for enhancing LTE channel codec

In order to enhance voice transmission over LTE networks, channel coding plays a very important role. If channel coding is executed well, voice quality will be improved and vice versa. The process of channel coding as well as channel decoding affect directly on end-to-end delay as well as BER. This means the better channel codec, the higher voice quality. For this goal, we propose two solutions for enhancing voice transmission over LTE network. The first solution presents a dynamic adaptation algorithm of joint source-channel code rate for enhancing voice transmission over LTE network. In order to measure the speech quality (also known as QoE), we use the WB E-model. In this model, both end-to-end

delay and packet loss are taken into account. The goal of this proposal is to find out the best suboptimal solution for improving voice traffic over LTE network with some constraints on allowed maximum end-to-end delay, allowed maximum packet loss, and the minimum required bandwidth. The best suboptimal choice is channel code rate corresponding to each mode of the AMR-WB codec that minimizes redundant bits generated by channel coding with an acceptable MOS reduction. Further, this algorithm can be also integrated with rate control in AMR-WB codec to enhance selection of the required mode.

When data is coded at sender by channel coding then at receiver it is decoded by channel decoding algorithms. For this task, the channel decoding algorithms also play an important role to improve voice quality. This is the reason for us to present the second proposal. It is an improved Log-MAP algorithm for Turbo decoding. In the proposed algorithm, we exploit the understanding of polynomial regression function to approximately compute the logarithm term (also called error correction function) in the Jacobian logarithmic function. The goal of this idea is to replace the correction function in the log-MAP algorithm on another one with the approximated performance and the reduced computational complexity.

We also investigate several well-known channel decoding algorithms which are based on MAP algorithm utilized in Turbo decoding. This allows us to compare them to our proposed solution.

Parts in this chapter were published in [4] and [5].

Chapter 4: The proposed LTE Downlink packet schedulers for voice traffics based on user perception and VoIP-priority mode

In this chapter, we present proposed scheduling schemes in LTE downlink direction. The main idea of this Chapter is proceeded from Chapter 2. We see that, QoE is becoming a new trend for enhancing real-time voice services such as VoLTE. Hence, we would like to enhance MAC scheduler in LTE downlink with the presence of QoE. QoE is estimated by using the extended E-model and WB E-model for narrowband audio and wideband audio users, respectively. First, we present a new scheduling scheme (called E-MQS scheduler) which is the combination of MQS factor, the modifications of M-LWDF-based scheduling algorithm, and user perception for narrowband audio services. In order to predict user satisfaction, we use the extended E-model. These proposals were published in [6] and [7]. Besides, for VoLTE service, AMR-WB is mandatory, but most of existing softwares which allow simulating VoLTE service do not support this voice codec including the LTE-Sim. For this reason, we propose to complement the AMR-WB codec into the LTE-Sim software. This allows the LTE-Sim can simulate wideband audio services such as VoLTE. Then, we upgrade E-MQS scheduler to suit for the wideband audio services. In order to do this, the WB E-model is used to measure the user perception (so called WE-MQS scheduler). This proposed scheduling scheme was published in [8].

Due to LTE is fully based on IP network, VoLTE is also deployed in an All-IP network. This means VoLTE is also a VoIP service but its QoS requirements are guaranteed by the mobile network operators. VoLTE is a special service that is very sensitive to network impairments such as delay, PLR, jitter, etc. Therefore, it needs to have the special priority. For this goal, we proposed to integrate the VoIP-Priority mode into our proposed schedulers to enhance VoIP flows (so called WE-MQS-VoIP Priority scheduler). This proposal was published in [9].

In addition, we also investigate several other well-known downlink packet schedulers such

as FLS, EXP/PF, M-LWDF and M-LWDF-based algorithms and compare their performance to our proposed schedulers.

Chapter 5: The proposed object non-intrusive models for predicting voice quality in LTE networks

In order to guarantee user perception, MNOs always have to monitor voice quality to adjust network impairments instantly. VoLTE is a new service in LTE network and based on IMS. So its deployment is very complex. Thus, how to monitor and evaluate VoLTE quality is a big challenge for MNOs. For this goal, we present two non-intrusive models for predicting voice quality in LTE systems. The first model is used for narrowband audio services whereas the second one is used for wideband audio services. Both of them are the combination of the LTE-Sim software with the extended E-model and WB E-model for measuring narrowband audio quality and wideband audio quality, respectively. The main idea is the usage of the extended E-model and WB E-model which are represented in Chapter 3 to predict voice quality. The numerical results show that these models may predict voice quality not as exact as the intrusive models such as PEQS or PESQ-WB, but they do not require any reference to original signal. They are thus very suitable for predicting real-time speech quality such as VoLTE. When user satisfaction can be automatically monitored, MNOs can adjust network impairments, choose suitable schedulers, thus, the better MCS can be used for enhancing LTE channel codec. The proposed models can be also useful for researchers in laboratories to evaluate VoLTE quality for many different scenarios which are configured in the LTE-Sim software. These two proposals were published in [10] and [11].

Chapter 6: Conclusions and perspectives

This chapter concludes our contributions in this thesis. Besides the obtained results, we also offer some limitations of our proposals. This is the base to continue to find out new research directions for my work in the future. Some new perspectives are proposed in this chapter.

Chapter 1

An overview of LTE technology and E-model

Contents

1.1	Overview of LTE technology	12
1.1.1	LTE architecture	12
1.1.2	Quality of Service	12
1.1.2.1	The bearer	12
1.1.2.2	QoS Class Identifier	13
1.1.3	Enhanced technologies	13
1.1.3.1	TTI Bundling	13
1.1.3.2	Semi-Persistent Scheduling	14
1.2	VoLTE	15
1.2.1	Strategies for VoLTE	15
1.2.2	VoLTE source codec	18
1.2.3	VoLTE radio protocol stack	18
1.2.3.1	PHY layer	19
1.2.3.2	MAC layer	20
1.2.3.3	LTE channel coding	21
1.3	Overview of the E-model and the WB E-model	24
1.3.1	E-model - Speech quality assessment for narrowband audio	24
1.3.2	WB E-model - Speech quality assessment for wideband audio	26
1.4	Summary	27

In this Chapter, we present a brief overview of the general LTE architecture standardized by the 3GPP specifications, VoLTE, and the E-model. First, we describe an overview of LTE technology in Section 1.1. Concepts such as EPS, E-UTRAN are discussed. We extend a general overview of the QoS architecture in LTE network. In this subsection, we highlight on how the QoS is handled. Next, an overview of VoLTE is described in Section 1.2. Several main layers in VoLTE radio protocol stack are discussed. We describe detailedly VoLTE source codec, MAC and PHY layers which are closely related to our proposals in this thesis. Third, in Section 1.3, we present an overview of the E-model and the WB E-model which are the backbones of our contributions in this thesis. Last, we summarize the main content in this Chapter such as shown in Section 1.4.

1.1 Overview of LTE technology

1.1.1 LTE architecture

According to [12], a typical LTE network consists of three main components such as UE, E-UTRAN and EPC. Figure 1.1 illustrates an example of deployment for LTE network. Point-to-multipoint (P2M) mode is the popular deployment by mobile operators where one eNodeB (or eNB) serves many UEs or terminals. Each eNB is a base station that controls the mobiles in one or more cells. The brief description of these components are described as follows [3]:

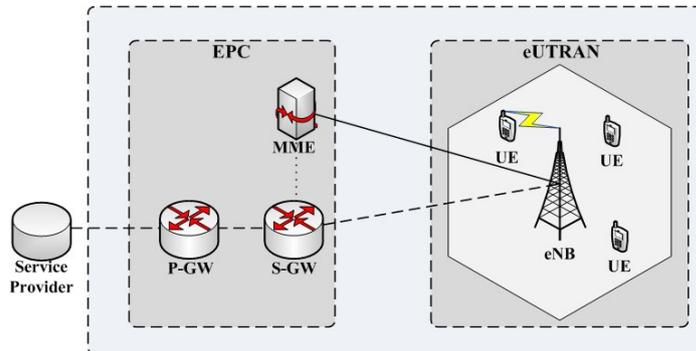


Figure 1.1: LTE architecture deployment

- **User Equipment (UE):** Includes functionalities of a mobile terminal that is responsible for call functions, a terminal equipment for data stream and Universal Subscriber Identity Module (USIM). The USIM stores network identity and user information.
- **Radio Access Network (RAN):** It is also called the E-UTRAN. It handles communication between the UE and the EPC and consist of the base stations called eNodeBs. The radio related functions consist of resource management, admission control and security. LTE has a flatter architecture because it has no RNC which is in 3G network and the feature resulting in better efficiency. When the UE roams into another eNodeB area, complete UE state is transferred to the other eNodeB.
- **The Evolved Packet Core (EPC):** It is also called SAE. It is the new All-IP core defined by 3GPP in Release 8. It performs functions such as network access control, mobility management, security and network management. The subscriber information is stored in the HSS. The MME control the setup and release of connections between the user and the packet data network. Is also does UE authentication location registration and handover using information from the HSS. The P-GW perform the function of GGSN and SGSN of the 3G network, i.e. connectivity to the IP network. The S-GW routes the data between the base stations and P-GW.

1.1.2 Quality of Service

1.1.2.1 The bearer

The QoS in LTE introduces a central element called bearer. A bearer identifies packet flows that receive a common QoS treatment between the terminal and the GW. All packet flows

mapped to the same bearer receive the same packet-forwarding treatment (e.g. scheduling policy, queue management policy, rate-shaping policy, link-layer configuration, etc.). In LTE, all the applications are mapped to GBR and non-GBR bearers as shown in Figure 1.2. Non-GBR bearers are also known as default bearers, and GBR-bearers are also known as dedicated bearers. A GBR guarantees a minimum bit rate requested by an application. In GBR, packet loss is assumed does not occur because of overflow of buffer, but services in non-GBR may encounter packet loss issue. Bearers are established, deleted and modified at the gateway by an entity called 'Policy Controller'. The Policy Controller makes its decisions based on QoS parameters such as QCI, ARP, MBR, GBR [13]. GRB bearers are established on demand basis since these bearers block transmission resources by reserving them. Otherwise, non-GRB may exists for long time since this is non-blocking resource transmission [12]. Non-GBR bearers do not guarantee any particular bit rate, and are typically used for applications as WWW, e-mail, chat, FTP, etc.

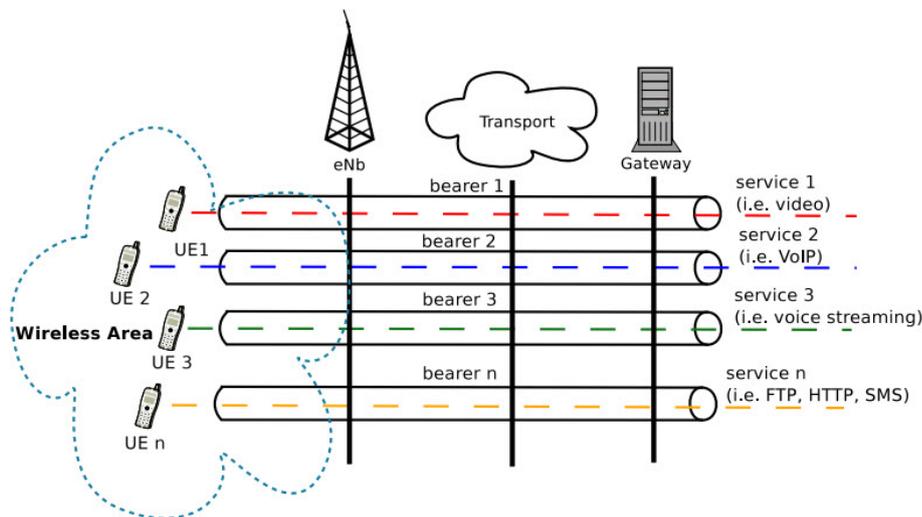


Figure 1.2: The bearer concept

1.1.2.2 QoS Class Identifier

In order to choose a bearer, the QCI is used. A bearer consists of information of barer type, priority, packet delay budget, and packet loss rate. LTE architecture supports hard QoS, each bearer has its separate QoS requirements as in Table 1.1. The ARP is used to decide whether a bearer establishment or modification request can be accepted or must be rejected due to resource limitations. The MBR is the bit rate that traffic on the bearer must not exceed, and the GBR is the bit rate that the network guarantees to users.

1.1.3 Enhanced technologies

1.1.3.1 TTI Bundling

Transmission Time Interval (TTI) Bundling is used to optimize the UL coverage for VoLTE. LTE defined 1 ms subframes as the TTI which means scheduling occurs every 1 ms. A one milisecond (ms) transmission interval combined with an 8 ms RTT, see Figure 1.3. TTI

Table 1.1: LTE service classes with QoS requirements

Resource Type	Priority	Packet Delay Budget (ms)	Packet Error Loss Rate	Example services
Guaranteed Bit Rate (GBR)	2	100	10^{-2}	Conversational voice
	4	150	10^{-3}	Conversational video (live streaming)
	3	50	10^{-3}	Real-time gaming
	5	300	10^{-6}	Non-conversational video (buffered stream)
Non-GBR	1	100	10^{-3}	IMS signaling
	6	300	10^{-6}	Video (buffered streaming) TCP-based (e.g. www, e-mail, chat, FTP, P2P sharing, progressive video, etc.)
	7	100	10^{-6}	Voice, Video (live streaming, Interactive Gaming)
	8	300	10^{-3}	Video (buffered streaming) TCP-based (e.g. www, e-mail, chat, FTP, P2P sharing, progressive video, etc.)
9	10^{-6}			

Bundling provides a way to increase the aggregation time without having to segment the packet. Small TTIs are good for reducing round trip latency, but do introduce challenges for UL VoIP coverage. TTI bundling was introduced in Rel-8 which combined four subframes spanning 4 ms. This allowed for a single IP header over a bundled 4 ms TTI that greatly improved the subframe utilization (from 1/8 to 1/2) and thus the coverage (by more than 3 dB). The drawbacks of TTI Bundling consist of dependence on user equipment, limited transmission choices, increased RTT and increased RRC signaling.

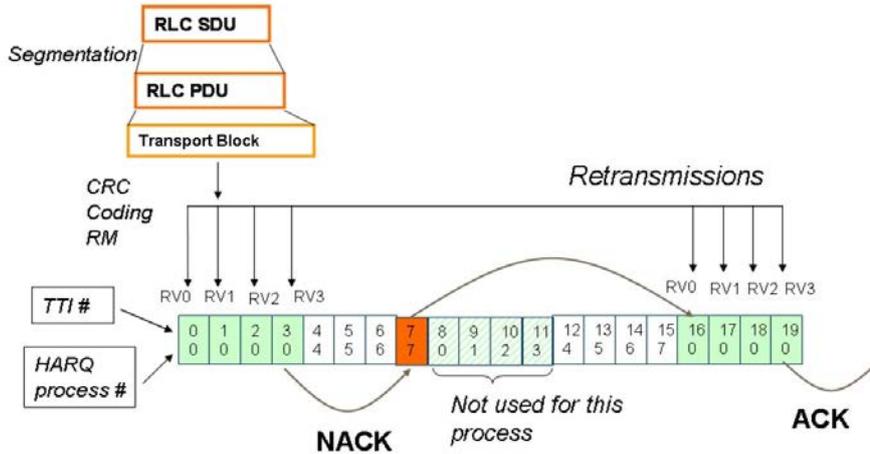


Figure 1.3: TTI Bundling of 4 TTIs

1.1.3.2 Semi-Persistent Scheduling

In LTE, VoLTE uses the AMR-WB such as source codec. This codec has 9 different bit rates. A dynamic scheduler (DS) typically offers grants (information on new resource allocation sent on the downlink control channel) for each new voice packet that needs to be transmitted. This can quickly lead to a bottleneck on the PDCCH and limit VoLTE capacity [2]. Since for a given AMR-WB codec mode, the voice packet size is fixed and the

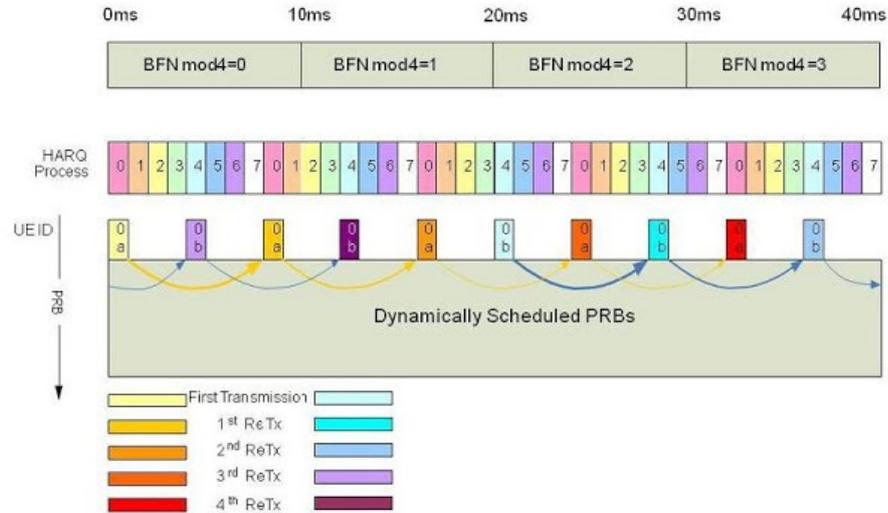


Figure 1.4: Example of Semi-Persistent Scheduling

packets are generated every 20 ms while in order to sustain voice quality, silent insertion descriptor (SID) packet arrives every 160ms. So that, 3GPP introduces SPS as an approach to utilize this information and improve grant utilization, see Figure 1.4. SPS grants are issued only at the beginning of an active or inactive period and the scheduler pre-allocates a set of dedicated resources in time and frequency for some of the initial transmissions. This pre-allocation can lead to significant grant savings.

The main drawback of SPS is the resource allocation must not be the best match because channel or interference conditions may change over the process of the activity or inactivity period. This presents a tradeoff between grant savings versus performance loss due to absence of dynamic link adaptation.

1.2 VoLTE

1.2.1 Strategies for VoLTE

VoLTE is expected to become the mainstream solution for providing voice services in commercial LTE networks. VoLTE integrates VoIP, LTE radio network (i.e., E-UTRAN), LTE core network (i.e., EPC), and the IMS to support voice services. Voice service in LTE can be provided both by CS network or LTE network. There is no impact to legacy CS network if voice service remains in CS. LTE system is an All-IP network without CS domain. VoIP service can be supplied on LTE system with IMS based voice call control functionality. In order to deploy VoLTE, mobile operators can use the following strategies:

- **Strategy 1: Voice over CS**

In this strategy, voice service keeps in legacy CS network while LTE system provides only high speed data service. This strategy includes two solutions: VoIMS and SRVCC to provide IMS-based VoLTE and support PS/CS interworking between LTE and 2G/3G. This strategy is also standardized by 3GPP. This can be the target solution

for VoLTE.

- **Strategy 2: Voice over LTE**

In this strategy, LTE system provides VoIP service and high speed data service simultaneously. IMS is used for voice call control. Voice service should be supplied in CS network when out of LTE coverage. This strategy includes three solutions: The first solution is CS Fallback (LTE first) if voice is needed, reselecting from LTE to 2G/3G to provide, the second solution is CS Fallback (2G/3G first) with firstly select 2G/3G, if data service is needed, reselecting from 2G/3G to LTE based on some conditions. These two solutions are also standardized by 3GPP. And the last one is Multi-mode Dual-standby Single-USIM Terminal when the UE access 2G/3G and LTE simultaneously, voice over 2G or 3G, data service over LTE or 2G/3G. This is a network independent solution.

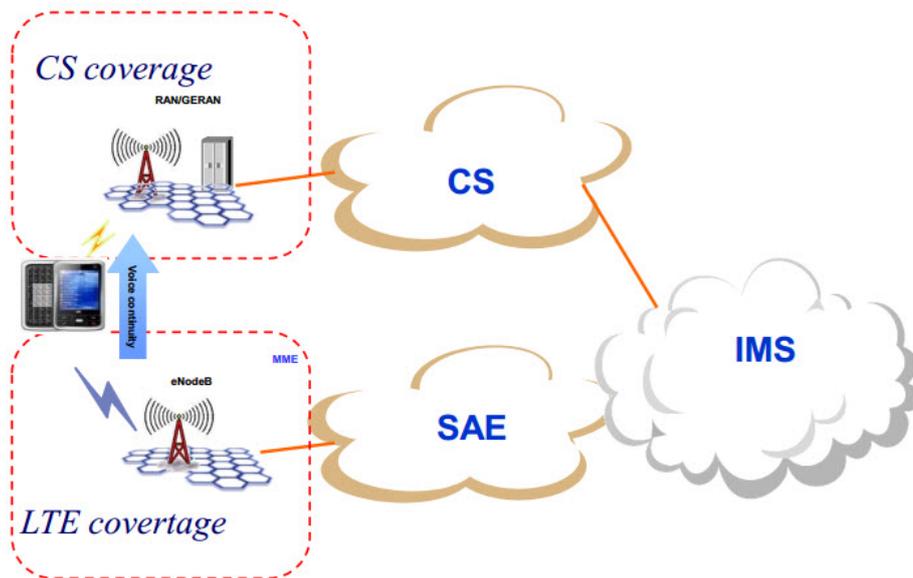


Figure 1.5: Different strategies for LTE voice service provision

In order to demonstrate clearly on different strategies for VoLTE, and how to VoLTE interworking with legacy networks, we describe the general principle of Voice over LTE as follows:

- Legacy networks provide different services via CS and PS domains where CS domain supplies voice and SMS services and PS domain provides data service as shown in Figure 1.6.
- When VoLTE is deployed based on IMS, VoIP service and high data service are simultaneously offered over LTE. Due to lack of call control function, thus, EPC network needs IMS network for call control and service provision. IMS network provides voice continuity from LTE to CS by anchoring voice call. While LTE terminal uses the same MSISDN in LTE and CS networks as shown in Figure 1.7.

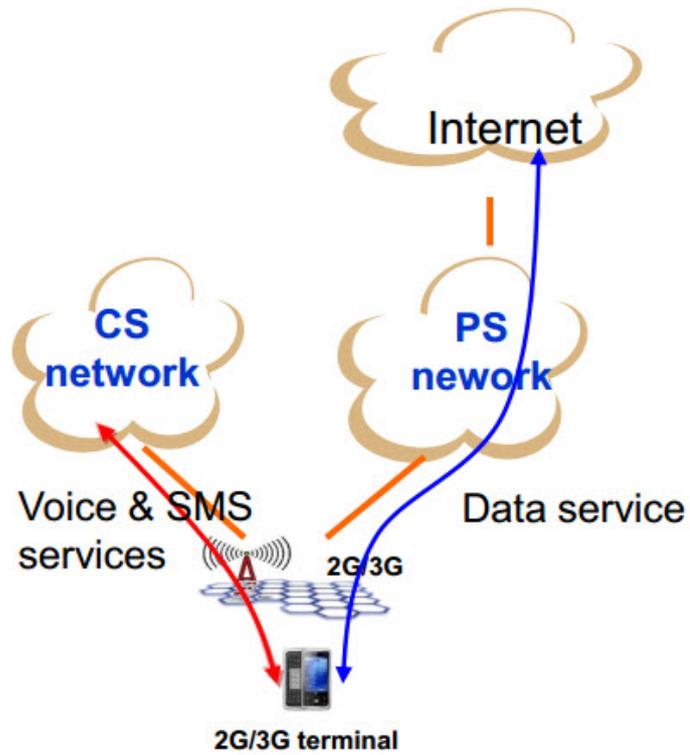


Figure 1.6: Voice in Legacy networks

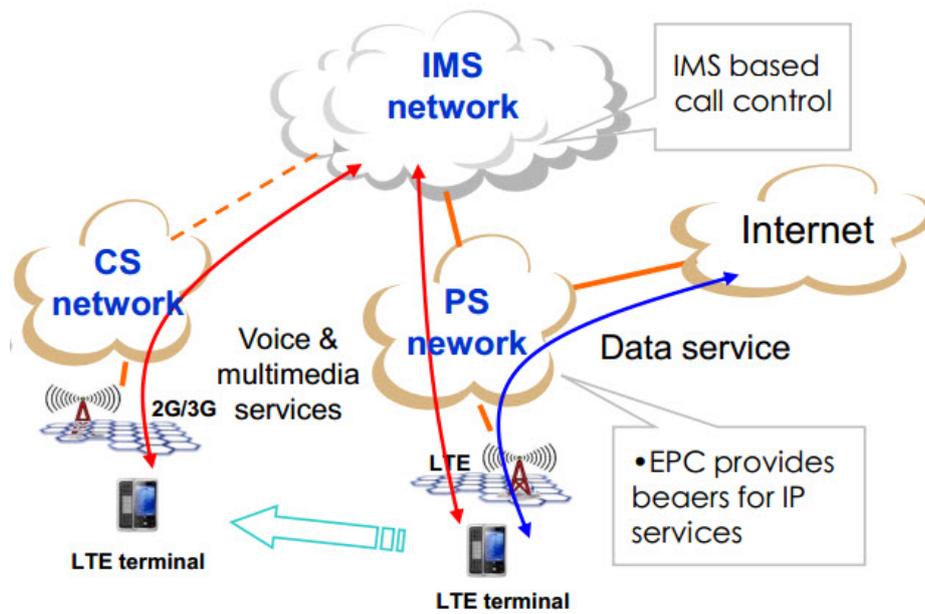


Figure 1.7: VoLTE with IMS

1.2.2 VoLTE source codec

VoLTE uses AMR-WB as its voice source codec. This is mandatory. AMR-WB codec is a speech codec which has been developed by ETSI (the European Telecommunications Standards Institute) and applied in the 3GPP LTE network for voice compression and decompression. It is fully described in [14]. AMR-WB codec uses a sampling rate of 16 kHz, which covers 50-7000 Hz audio bandwidth. It has 9 different codec modes (from mode 0 to mode 8) corresponding to 9 source bitrates in range of 6.6-23.85 Kb/s. Each of them generates an encoded 20 ms speech frame and switches among them every 20 ms. The bits in the encoded speech frame are ordered according to their subjective importance. These bits are divided into three classes with reducing perceptual importance: Class A, Class B and Class C. Total bits of each class depend on codec mode. AMR-WB packet size depends on the bitrate (mode) is described such as in Table 1.2 [15].

Table 1.2: Packet sizes of AMR-WB codec modes

Parameter	AMR-WB bitrate (kbps)								
	23.85	23.05	19.85	18.25	15.85	14.25	12.65	8.85	6.6
Payload size (bits)	477	461	397	365	317	285	253	177	132
Frame size (bits)	488	472	408	376	328	296	264	192	144
RTP header (bits)	96	96	96	96	96	96	96	96	96
Packet size (bits)	584	568	504	472	424	392	360	288	240

In LTE network, AMR-WB codec is configured into 3 configurations [16] as follows:

- Configuration A (Config-WB-Code 0): 6.6, 8.85, and 12.65 Kb/s (Mandatory multi-rate configuration)
- Configuration B (Config-WB-Code 2): 6.6, 8.85, 12.65, and 15.85 Kb/s
- Configuration C (Config-WB-Code 4): 6.6, 8.85, 12.65, and 23.85 Kb/s

These configurations are used to simplify the negotiation of bitrate between the user equipment and the base station, thus will simplify the implementation and testing. The remaining bitrates can still be used for other purposes in mobile networks. In order to choose a bitrate, the receiver measures quality of radio channel. The CQI is used for this purpose. It is defined as an equivalent C/I ratio. The C/I ratio is then compared to a set of predefined thresholds to decide which mode to be used. Switching among modes in a configuration depend on the rate control algorithm in AMR-WB codec. The criterion for mode switching is threshold value of C/I ratio. These threshold values depend on the channel condition, frequency hopping scheme, network configuration and other factors. Furthermore, network conditions change over time, so that, even well-selected adaption thresholds will not be best.

1.2.3 VoLTE radio protocol stack

In this thesis, we usually use the LTE-Sim software to simulate our proposals, thus, in this section, we present the protocol stack of VoLTE deployed in the LTE-Sim. LTE-Sim is an open-source framework which is developed by Giuseppe Piro and his colleagues [17]. It is freely available for scientific community. It is used to simulate entire LTE network. There are many researchers who used LTE-Sim to simulate their proposals such as scheduling strategies, radio resource optimization, frequency reuse techniques, the adaptive modulation

and coding module, user mobility, and etc. for both downlink and uplink directions and in both multicell/multiuser environments. The implemented protocol stack of LTE-Sim is represented in Figure 1.8 for both user-plane. It's clear that it is nearly similar to a real LTE system.

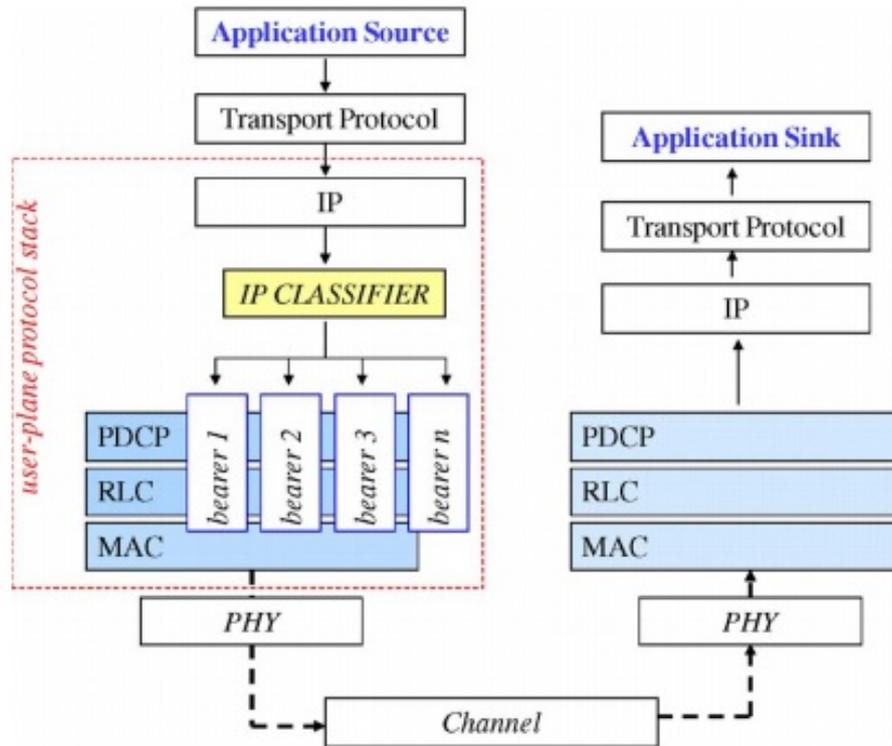


Figure 1.8: The implemented protocol stack in the LTE-Sim software

Figure 1.8 can be briefly described as follows: When a voice traffic flow transmitted over the LTE-Sim, it is encapsulated sequentially with network protocols. For the downlink direction, the VoLTE packet uses transport protocols of RTP, UDP and IP. It is then packetized with radio protocols such as PDCP, RLC and MAC, and PHY layer before it is transmitted over the air interface. At each layer, the corresponding header size is added. LTE-Sim supports both IPv4 and IPv6 protocols with header sizes are 40 and 60 bytes, respectively while the voice payload is about 32 bytes, thus, to reduce the overhead, RoHC is deployed at PDCP layer. The IP header is then compressed by RoHC down to only 1-4 bytes, normally 3 bytes.

In this thesis, our proposals mainly focus on the APL, MAC and PHY layer, thus, we will present a brief introduction to them. For the APL layer, we use some characteristics of AMR-WB codec in our proposals, so this codec is described in Section 1.2. Description of MAC and PHY layers is depicted as the following subsections:

1.2.3.1 PHY layer

PHY (Physical) layer plays an important role for transmitting both data and control information between the eNB and UEs. The most important features of this layer are OFDM and

the support of FDD and TDD as radio frame structure. OFDM systems divide the available bandwidth into many narrower sub-carriers and transmit the data in parallel streams. Since data is transmitted in parallel rather than serially, OFDM symbols are generally much longer than symbols on single carrier systems of equivalent data rate. PHY layer utilizes OFDMA for downlink direction and SC-FDMA for uplink one. Fully description of PHY layer is presented in the 3GPP specifications [18], [19].

OFDMA. OFDMA allows data to be directed to or from multiple users on a subcarrier-by-subcarrier basis for a specified number of symbol periods. Although the LTE specifications describe both FDD and TDD to separate UL and DL traffic, market surveys demonstrate that the major deployment in LTE systems will be FDD [20]. OFDMA is an excellent choice of multiplexing scheme for LTE downlink. Although it involves added complexity in terms of resource scheduling, it ensures high performances in terms of efficiency and latency. In OFDMA, users are allocated a specific number of sub-carriers for a predetermined amount of time. These are so-called PRBs. PRBs have both time and frequency domain. Allocation of PRBs is processed by a scheduling function at the eNB.

Link adaptation, Modulation and Coding. The AMC is a powerful technique used by 4G technologies such as LTE and WiMAX to enhance the robustness of the communication due to the high variation of channel conditions. This is attained by deploying a robust MCS, i.e. transmitting at low data rates when the channel is poor and increasing the data rate using a more efficient MCS when the channel conditions are good. The modulation techniques are supported by LTE including BPSK, QPSK, 16QAM and 64QAM. All MCS values supported by LTE are qualified in the 3GPP specifications [21].

Frame Structure. In OFDMA, the general frame structure is used with FDD. Another alternative frame structure is TDD. LTE frames are 10 ms in length. Each frame is divided into 10 subframes, each subframe is 1.0 ms in length. Each subframe is further divided into two slots with a length of 0.5 ms. Each includes either 6 or 7 OFDM symbols depending on whether the normal or extended cyclic prefix is deployed [22].

The total number of available sub-carriers depends on the whole transmission bandwidth of the system. The LTE specifications define parameters for system bandwidths in range of 1.25..20 MHz. A PRB consists of 12 consecutive sub-carriers for one slot (0.5 ms) in length. A PRB is the smallest element of resource allocation assigned by the eNB scheduler. The structure of radio frame in LTE is shown in Figure 1.9.

1.2.3.2 MAC layer

MAC (Medium Access Control) layer is responsible for multiplexing and demultiplexing data between the PHY layer and RLC layer. This layer includes logical channels that are connected to physical channels for transmission of data between the PHY and MAC layers. The key functions of the MAC layer consists of scheduling of radio resources among UEs, random access procedure, uplink timing alignment, discontinuous reception, and scheduling information transfer [23], [24]. The MAC layer in LTE provides a medium-independent interface to the PHY layer and is designed to support the PHY layer by focusing on efficient radio resource management. The MAC layer provides data transfer services on logical channels. A set of logical channel types is defined for different types of data transfer services as offered by the MAC layer.

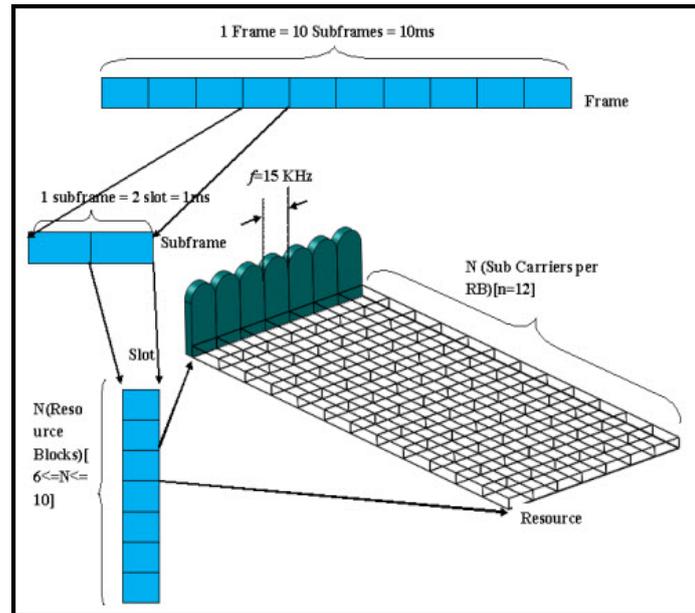


Figure 1.9: Resource grid

1.2.3.3 LTE channel coding

One of the most important issues in digital communication systems is error correction. In order to protect the voice packet when it is delivered over a noisy channel, some error correction technologies are included. Error correction can be divided in two main categories including ARQ and FEC. ARQ is used to request retransmission of the data packets if errors are detected. This will be done until the received packets are error-free or a maximum number of retransmissions is reached.

For FEC technique, redundant bits are added to the transmitted data bits to detect and to correct errors. FEC has two main kinds consisting of Block codes and Convolutional codes. In block codes, the input data blocks which can be considered as vectors are multiplied by a generator matrix resulting in a code word vector is generated. The convolutional codes have a similar structure like FIR (Finite Impulse Response) filters and operate bitwise. The convolutional codes have memory which means the coded output bit depends not only on the current bit but also on the m previous bits, in which m is the number of registers in the convolutional encoder. In LTE, both block codes and convolutional codes are used. One of the featured Block codes is CRC which is a cyclic linear block code is used for HARQ as an error detecting technique. There are two CRC schemes for a Physical Downlink Shared Channel (PDSCH) in LTE: $gCRC24A$ and $gCRC24B$. Both of them have a 24 parity bits length, but work with different cyclic generator polynomials. The $gCRC24A$ focuses on a transport block, while the $gCRC24B$ concentrates on the code block, which is the segmentation of a transport block when the size of a transport block is larger than the upper limit (6144 bits). Control channels and data channels in LTE use convolutional and Turbo codes, respectively, where the Turbo code is an enhancement of convolutional codes to obtain the nearest Shannon performance.

In this thesis, two of our proposals are about channel codec in LTE. We proposed the solutions for enhancing LTE channel coding and decoding. Hence, in this section, we will

briefly introduce to the Turbo codes in LTE networks.

1.2.3.3.1 Turbo encoder The 3GPP Turbo code is a systematic Parallel Concatenated Convolution Code (PCCC) with two 8-state constituent encoders (the same as in UMTS) and one turbo code internal interleaver (different from UMTS) [25]. It was first introduced by Berrou et al [26], and until now it is still one of the most powerful error correcting codes that obtains the performance closest to the Shannon capacity. The standard turbo code rates are 1/3, 3/4, and 4/5, where code rate 1/3 is the original code rate. Turbo code rate is chosen based on CQI (Channel Quality Indicator) index [27]. CQI index includes 16 values, where value of 0 is not used. Values of CQI index from 1 to 6, 7 to 9, and 10 to 15 are corresponding to Turbo code rates of 1/3, 3/4, and 4/5. Each CQI index is mapped to a SINR (Signal-to-Interference-plus-Noise Ratio) value. SINR also has 15 values (from -6.7 dB to 22.7 dB) [28] counted by the receiver and sent to the transmitter. Channel coding will map the SINR value to the corresponding CQI index, and then chooses the corresponding channel code rate.

Turbo code is also applied in many modern wireless communication standards, such as HSDPA [29] and LTE [25]. The Turbo encoders are based on Recursive Systematic Convolutional (RSC) codes and their generator polynomial is given by $G = [1, g_0/g_1]$, in which $g_0 = [1011]$ (feedback) and $g_1 = [1101]$ (feedforward). The structure of the Turbo encoder used in LTE can be illustrated in Figure 1.10 [30].

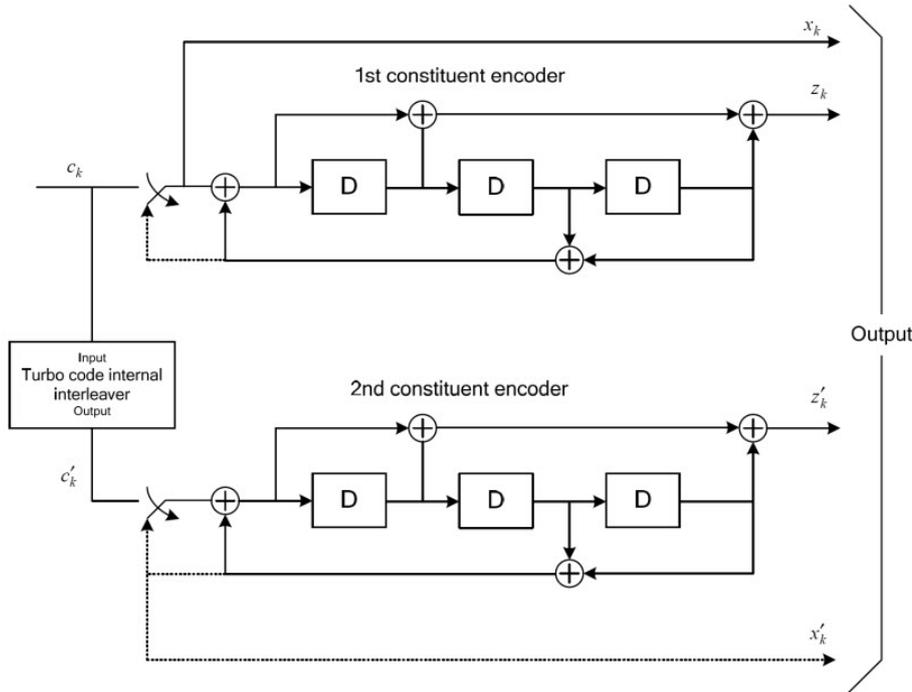


Figure 1.10: Structure of the Turbo encoder (the dotted lines are for trellis termination)

As shown in Figure 1.10, the output of the LTE Turbo encoder including three parts, a systematic bit and two parity bits. The systematic bit (X_k) is the untouched input bit. The first parity bit (Z_k) is the output of the first convolutional encoder with the original input (C_k) and the second parity bit (Z'_k) is the output of the second convolutional encoder

after the internal interleaver of the input bit (C'_k) as its input. For trellis termination the tail-bits (X'_k) are inserted. Turbo coding is used in LTE for following channels: UL-SCH (Uplink Shared CHannel), DL-SCH (Downlink Shared CHannel), PCH (Paging CHannel) and MCH (Multicast CHannel).

The internal contention-free interleaver. It is one of the main components of the LTE Turbo encoder. The main difference between the Turbo encoders in LTE and UMTS is that the interleaver in LTE is in contrast to the one in UMTS contention-free. Contentions occur, when parallel working processes try to write or read to/from the same memory address simultaneously. Since the two SISO MAP decoder engines of the Turbo decoder use such processes, the contention-free concept becomes survival for designing the Turbo encoders internal interleaver efficiently.

There were two candidates for the LTE internal interleaver: Almost Regular Permutation (ARP) and Quadrature Permutation Polynomial (QPP), which are very similar. However QPP is chosen for LTE because it offers more parallelism. The QPP interleaver for a block size of K is defined as follows: $\pi(i) = (f_1i + f_2i^2) \bmod K$, where i is the input index, $\pi(i)$ is the output index, and f_1 and f_2 are permutation parameters, which can be get from the standard.

1.2.3.3.2 Turbo decoder The Turbo decoder in LTE is based on two SISO decoders, which work together in an iterative manner. Each SISO decoder has two inputs called a normal input and an a-priori Log Likelihood Ratio (LLR) input, and two outputs called an a-posteriori LLR and an extrinsic LLR. In each iteration the (de)interleaved version of the extrinsic output of a decoder is used as the a-priori information for the other decoder. Typically after 4 to 8 iterations, the a-posteriori LLR output can be used to obtain the final hard decision estimates of the information bits. The structure of the LTE Turbo decoder can be illustrated in Figure 1.11 [30].

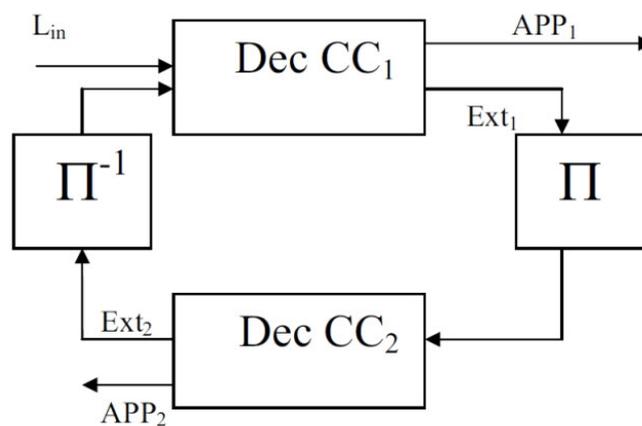


Figure 1.11: Structure of the Turbo decoder

1.2.3.3.3 Rate matching Since the mother code rate of LTE Turbo encoder is $1/3$, thus, to get other code rates, repetition or puncturing has to be performed, which both are done by a rate matching module. The rate matching module includes three so-called sub-block interleavers for the three output streams of the Turbo encoder core, a bit selection, and

pruning part which is realized by a circular buffer. The sub-block interleaver is based on the classic row-column interleaver with 32 columns and a length-32 intra-column permutation. The bits of each of the three streams are written row-by-row into a matrix with 32 columns (number of rows depends on the stream size). Dummy bits are padded to the front of each stream to completely fill the matrix. After a column permutation, bits are read out from the matrix column-by-column. The column permutation of the sub-block interleaver is given as following: [0, 16, 8, 24, 4, 20, 12, 28, 2, 18, 10, 26, 6, 22, 14, 30, 1, 17, 9, 25, 5, 21, 13, 29, 3, 19, 11, 27, 7, 23, 15, 31]. For example, if column 1 has to be read originally, after permutation column 16 will be read.

The Turbo encoder with rate matching module is illustrated in Figure 1.12 [30].

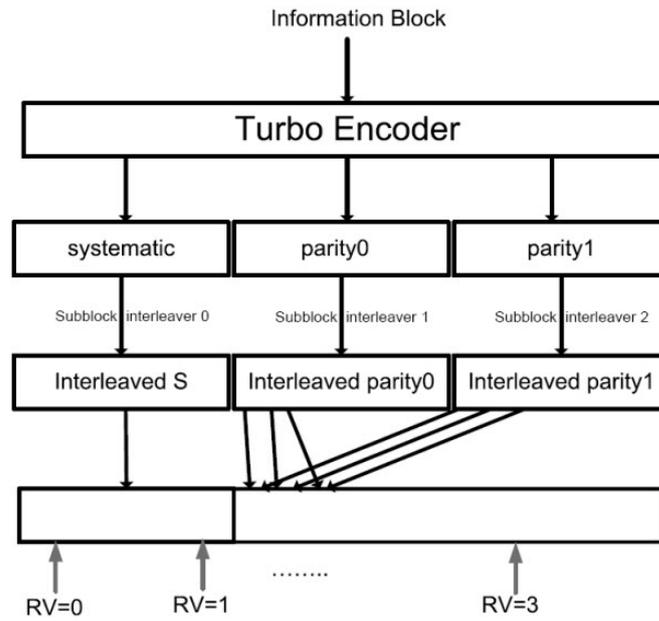


Figure 1.12: Turbo decoder with the rate matching module

1.3 Overview of the E-model and the WB E-model

1.3.1 E-model - Speech quality assessment for narrowband audio

E-model is a computational model developed and standardized by ITU-T [31]. It is used to estimate the MOS for narrowband audio quality. The reference connection of the E-model is illustrated in Figure 1.13.

As shown in Figure 1.13, the reference connection is split into a send side and a receive side. The model estimates the conversational quality from mouth to ear as perceived by the user at the receive side, both as listener and talker. The transmission parameters used as an input to the computation model. Values for room noise and for the D-factors are processed separately in the algorithm for the sender and receiver and may be of different amounts. The parameters SLR, RLR, and N_c are referred to a defined 0 dBr point. All other input parameters are either considered as values for the overall connection, such as overall OLR, i.e., the sum of SLR and RLR, number of qdu, equipment impairment factors I_e and

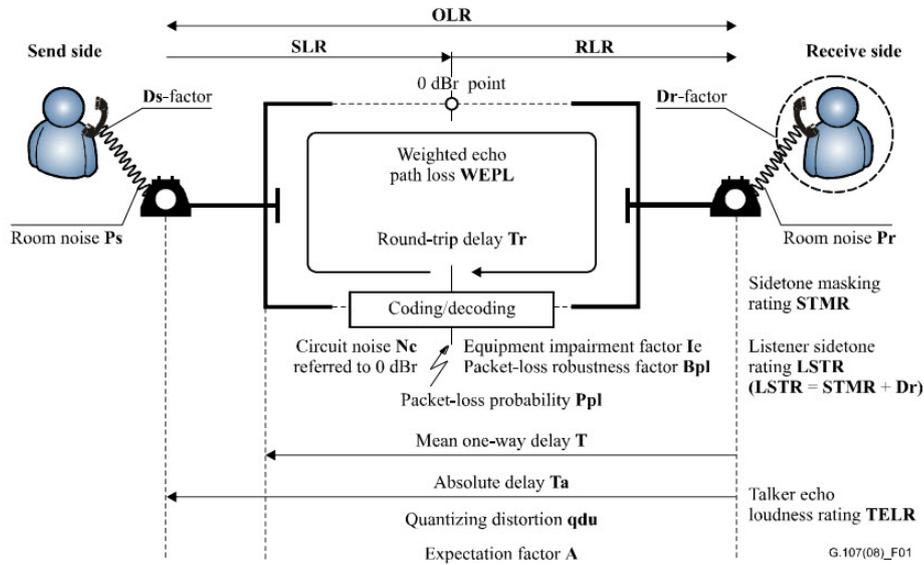


Figure 1.13: Reference connection of the E-model

advantage factor A , or referred to only for the receive side, such as $STMR$, $LSTR$, $WEPL$ used for the calculation of listener echo and $TELR$. There are three different parameters associated with transmission time. The absolute delay T_a represents the total one-way delay between the sender and receiver and is used to estimate the impairment due to excessive delay. The parameter mean one-way delay T represents the delay between the receiver (in talking state) and the point in a connection where a signal coupling occurs as a source of echo. The round-trip delay T_r only represents the delay in a 4-wire loop, where the 'double reflected' signal will cause impairments due to listener echo.

The output of the model is Transmission rating factor (R-factor). The values of this R-factor are in range of 0-100 where 100 is the best and 0 is the worst quality. And then, it is mapped to the corresponding MOS value. The standard R-factor in the E-model is defined as follows:

$$R = R_0 - I_s - I_d - I_{ef} + A \quad (1.1)$$

In which:

- R_0 : The basic signal-to-noise ratio which consists of noise sources such as circuit and room noise.
- I_s : The simultaneous impairment factor, it is the sum of all impairments which may occur more or less simultaneously with the voice transmission. In this model, the default value is set to 0.
- I_d : The delay impairment factor, representing all impairments due to delay of voice signals.
- I_{ef} : The equipment impairment factor, capturing the effect of signal distortion due to low bit rates of the codec and packet losses of random distribution.

- A : The advantage factor, capturing the fact that some users can accept a reduction of quality due to the mobility of cellular networks. In this model, this factor is set to 0.

In above factors, I_d and I_{ef} are affected by end-to-end delay and packet loss, respectively, while R_0 and I_s do not depend on network performance. The details of theoretical calculations, default values and permitted ranges of them are described in [31].

After setting the default values for the E-model, Equation (1.1) can be rewritten as follows:

$$R = R_0 - I_d - I_{ef} \quad (1.2)$$

The R-factor is then translated into the MOS as follows [31]:

$$MOS = \begin{cases} 1, & \text{if } R < 0 \\ 1 + 0.035 \times R + 7 \times 10^{-6} \times R \times (R - 60) \times (100 - R), & \text{if } 0 \leq R \leq 100 \\ 4.5, & \text{otherwise} \end{cases} \quad (1.3)$$

The relation between R-factor, user perception, and MOS is described in the Table 1.3.

Table 1.3: R-factor and MOS with corresponding user satisfaction

R	User satisfaction	MOS
$90 \leq R < 100$	Very satisfied	4.3-5.0
$80 \leq R < 90$	Satisfied	4.0-4.3
$70 \leq R < 80$	Some users dissatisfied	3.6-4.0
$60 \leq R < 70$	Many users dissatisfied	3.1-3.6
$50 \leq R < 60$	Nearly all users dissatisfied	2.6-3.1
$R < 50$	Not recommended	< 2.6

The Equation (1.2) is the base to compute the R factor. The calculation of I_d and I_{ef} used in this thesis will be described in the next Chapters to observe conveniently.

1.3.2 WB E-model - Speech quality assessment for wideband audio

WB E-model is a computational model developed and standardized by ITU-T [32]. It is used to estimate the MOS for wideband audio quality. The reference connection of the WB E-model is similar to the E-model.

The output of the model is R_{wb} -factor. The values of this R_{wb} -factor in range of 0-129. And then, it is mapped to the MOS. The R_{wb} -factor in the WB E-model is defined as follows:

$$R_{wb} = R_{0,wb} - I_{s,wb} - I_{d,wb} - I_{e,eff,wb} + A \quad (1.4)$$

In which:

- $R_{0,wb}$: The basic signal-to-noise ratio.
- $I_{s,wb}$: The simultaneous impairment factor, it is the sum of all impairments which may occur more or less simultaneously with the voice transmission. In this model, this factor is set to 0.

- $I_{d,wb}$: The delay impairment factor, representing all impairments due to delay of voice signals.
- $I_{e,eff,wb}$: The equipment impairment factor, capturing the effect of signal distortion due to low bit rates of the codec and packet losses of random distribution.
- A : The advantage factor, capturing the fact that some users can accept a reduction of quality due to the mobility of cellular networks. In this model, this factor is set to 0.

In above factors, $I_{d,wb}$ and $I_{e,eff,wb}$ are affected by end-to-end delay and packet loss, respectively, while $R_{0,wb}$ and $I_{s,wb}$ do not depend on network performance. The details of theoretical calculations, default values and permitted ranges of them are described in [32].

For $R = R_{wb}/1.29$, R factor is then mapped to the MOS using Equation (1.3), and then, the MOS is mapped to the satisfaction level of the users. According to [33], for the wideband audio, the value of $R_{0,wb}$ factor in equation (1.1) equals 129, thus, equation (1.1) can be rewritten as follows:

$$R_{wb} = 129 - I_{d,wb} - I_{e,eff,wb} \quad (1.5)$$

It is similar to the E-model, the Equation (1.5) is the base to compute the R_{wb} factor. The calculations of $I_{d,wb}$ and $I_{e,eff,wb}$ used in this thesis will be also described in the next Chapters to conveniently observe.

1.4 Summary

This Chapter presents an overview of the LTE network and the E-model. We introduce the general LTE architecture and its main components such as EPC, E-UTRAN, eNB and the general QoS architecture based on 3GPP specifications. Herein, we mainly focus on the QoS architecture and the MAC and PHY layers located at the eNB. Regarding the PHY layer, we describe its downlink frame structure called OFDMA and its sub-channelization. For the MAC layer, we detail the MAC air interface and its main functions. Since the radio resource allocation is performed at the MAC scheduler interface, we concentrate on this part to better understand the physical resource allocation and scheduling. In addition, the QoS architecture used by LTE is explained highlighting particular concepts such as bearer. However, state-of-the-art related to QoS support at MAC layer, specifically resource allocation techniques will be further discussed in the next Chapter. We also describe AMR-WB codec which is mandatory for VoLTE. This codec is used in our proposals to simulate wideband audio flows. As mentioned before, the E-model and WB E-model are the backbones in our proposals thus in this Chapter, we also detail the featured characteristics of them.

Chapter 2

State-of-the-art

Contents

2.1	LTE Turbo codec	30
2.1.1	LTE Turbo coding	30
2.1.2	LTE Turbo decoding	31
2.2	LTE Downlink scheduling schemes	35
2.2.1	Channel-unaware scheduling strategies	35
2.2.1.1	First In First Out	36
2.2.1.2	Round Robin	36
2.2.1.3	Blind Equal Throughput	36
2.2.1.4	Weight Fair Queuing	36
2.2.1.5	Largest Weighted Delay First	36
2.2.2	Channel-aware and QoS-unaware scheduling strategies	36
2.2.2.1	Maximum Throughput	37
2.2.2.2	Proportional Fair	37
2.2.2.3	Throughput to Average	37
2.2.2.4	Joint Time-Frequency domain	37
2.2.2.5	Buffer-aware	38
2.2.3	Channel-aware and QoS-aware scheduling strategies	38
2.2.3.1	Guaranteed Data Rate	38
2.2.3.2	Guaranteed Delay Requirements	38
2.2.4	Dynamic and Semi-persistent Scheduling for VoIP support	39
2.2.5	Energy-aware	40
2.3	QoE evaluation approaches for voice user	40
2.3.1	Standard models	41
2.3.1.1	E-model and wideband E-model	41
2.3.1.2	P.563 model	42
2.3.1.3	ANIQUE+ model	42
2.3.1.4	P.564 model	42
2.3.2	Non-standard models	43
2.3.2.1	Gaussian Mixture-based models	43

2.3.2.2	Neural networks-based models	43
2.3.2.3	Exponential functions-based models	44
2.3.2.4	Hybrid models	44
2.4	Summary	45

In this Chapter, we present a literature review on some topics that related to our proposals in this thesis. As mentioned before, in this thesis, we have proposed the solutions for enhancing and improving voice transmission quality in LTE network. The proposals mainly focus on LTE channel coding, scheduling schemes in downlink direction and prediction of voice quality. For these goals, in this Chapter, we present the literature review on Turbo coding and decoding, downlink scheduling strategies, and QoE requirements. In particular, Section 2.1 represents a review on LTE Turbo coding and decoding, downlink scheduling strategies are revised in Section 2.2, in Section 2.3, we present a revision of QoE requirements, and the Summary is shown in Section 2.4.

2.1 LTE Turbo codec

2.1.1 LTE Turbo coding

As mentioned in Chapter 1, LTE uses Block codes and Convolutional codes for channel coding purposes. The convolutional codes are used for control channels whereas Turbo codes are utilized for data channels. Turbo code is an enhanced convolutional code and is widely used in wireless communications. In this thesis, we focus on data channels, thus, in this section, we only present the literature review on Turbo code in LTE networks.

Authors in [34] presented the largest-spread and maximum-spread QPP interleavers (called LSQPP and MSQPP, respectively) to improve the performance of LTE QPP (Quadratic Permutation Polynomial) interleaver. Compared to LTE QPP interleaver, those algorithms have low computing complexity and better performances from simulation results. The authors concluded that, the BER performance of the MSQPP interleaver is better compared to the LTE and LS QPP interleavers and the MSQPP is suitable for LTE standard.

In [35], the authors presented two attractive turbo interleaver candidates (QPP and ARP) for LTE turbo coding. These interleavers are perfectly suited for LTE as they facilitate efficient high throughput turbo decoding required to support the high data rates envisioned for LTE systems. Some main features of these interleavers include flexible parallelism, efficient memory organization, and efficient support for Radix-4 decoding. The QPP interleaver held a slight edge over ARP due to its maximum contention-free property and its fully algebraic structure; hence it was selected for LTE turbo coding. Simulation results indicated that QPP and ARP designs for LTE have a performance equivalent to or better than the existing WCDMA turbo code.

Authors in [36] presented circular buffer rate matching (CBRM) algorithms for LTE turbo code of the WCDMA-based air interface. In order to enhance performance at high code rates, systematic bit puncturing is incorporated in conjunction with the CBRM. The RM algorithm is further optimized based on the algebraic properties of the QPP interleavers and the 8-state recursive systematic convolutional code of the LTE turbo code.

In [37], the authors first presented a detailed analysis of RCIC (Rate-Compatible Insertion Convolutional) Turbo codes. The authors showed that these codes outperform rate-compatible Turbo codes using bit repetition in terms of frame error rate and convergence

speed. Furthermore, their EXIT charts will be derived from the EXIT chart of the mother Turbo code. This EXIT chart analysis enables a novel semi-analytical optimization tool for RCIC Turbo codes and the efficient EXIT chart comparison of competing codes, avoiding complex simulations. Finally, a HARQ scheme for LTE, based on RCIC Turbo codes, is presented which obtains a higher system throughput in comparison with the standard LTE solution.

Authors in [38] presented an efficient method to implement high throughput processing is proposed. The method is attained by using recursive relationship among polynomials. The proposed method covers implementation of interleaver in both directions of forward state metric and backward state metric. In inclusion, with the proposed method, Turbo codes can be implemented in parallel processing with low latency for high data rate applications.

In [39], the authors proposed an efficient architecture for the QPP interleaver, called the Add-Compare-Select (ACS) permuting network. The proposed architecture can be used both as the interleaver and deinterleaver. This means it has a high-speed low-complexity hardware interleaver/deinterleaver for turbo decoding. It does not require memory or QPP inverse to perform deinterleaving and was fully implemented and tested both on a Virtex-6 FPGA as well as in a 0.18 μm CMOS process. Finally, it can be used for any block size.

In order to improve rate matching (RM) in LTE Turbo code, authors in [40] presented an efficient RM algorithm to improve the performance of turbo codes, specially at high code rates. The proposed algorithm based on interlaced puncturing not only eliminates most of these performance instabilities and reaches good performance over wide range of code rates but also brings relatively little computational complexity. However, the performance analysis of the computational complexity in software and hardware implementation has been not considered.

In [41], the authors proposed a new architecture of Turbo code encoder based on 3GPP standard. This architecture is developed by implementing optimized 8-level parallel architecture, dual RAM in turbo code internal interleaver, recursive pair wise matching, and efficient 8-level index generator in turbo code internal interleaver. In order to guarantee the functionality of the proposed algorithm and architecture, In conclusion, the proposed architecture successfully increases the speed of encoder 16 times faster compared to conventional architecture with size smaller than 50 %.

For joint source-channel coding, authors in [42] proposed a new joint source and channel decoding algorithm for 4G wireless systems. The goal of the algorithm is the correct reception of the correlated bits using joint and list decoders, after that correctly decoded values are used for improvement of non correlated bits. Performance of the new decoder is shown for LTE turbo code and AMR-WB vocoder, but can be applied to other sources. The achievable gain by the new decoder compared to the conventional turbo scheme is 0.15 dB, that is equal to 0.15-0.4 MOS speech quality improvement in comparison with the Scaled Max-Log-MAP and the conventional joint source-channel decoders.

2.1.2 LTE Turbo decoding

As mentioned in Chapter 1, Turbo decoding algorithms are based on the Trellis detection. The algorithms of this group can be classified into two categories as follows:

- **Sequence detection.** Including Viterbi algorithm, then SOVA (Soft Output Viterbi Algorithm) and SOVA-based algorithms. This group is used for the convolutional codes.

- **Symbol-by-symbol detection.** Consisting of MAP (Maximum A Posteriori) algorithm, then Log-MAP, Max-Log-MAP, and Log-MAP based algorithms. This group is utilized for Turbo codes.

In this section, we first review on Turbo decoders used in LTE networks, then we mainly focus on the Log-MAP based algorithms because they are relevant to our proposals in this thesis.

Turbo decoder is used in LTE networks to decode the encoded signals by Turbo encoder. There have been many papers proposed to modify the architecture as well as to improve performance of Turbo decoder. Authors in [43] presented a design of turbo decoder VLSI architecture based on 3GPP-LTE standard. In order to obtain the LTE peak data rate of 326.4 Mbps, Max-log-MAP radix-4 decoding algorithm and parallelization of the MAP decoder are used. The data dependency between each MAP decoder and hazard that comes from the usage of parallelization is explained. The modification made on the data flow and architecture of the branch/transition metric cache memory to increase the concurrency of parallelization method and to eliminate the hazard is introduced. Proposed design has a maximum throughput of 347.8 Mbps and FPGA implementation shows that the system can attain maximum frequency of 102.57 MHz. A fixed-point implementation of LTE turbo decoder with Max-Log-MAP algorithm was introduced in [44]. Also modular arithmetic was introduced to the state metric calculation and the scale process was introduced to extrinsic information calculation. With small hardware complexity, the resource overhead was reduced, and the working speed was improved. The results show that the decoding performance is much better than Max-Log-MAP algorithm and is close to the Log-MAP algorithm, with a small degradation less than 0.1 dB.

In [45], the authors proposed low-complexity energy-efficient Turbo decoder architecture. Turbo codes have recently been considered for energy-constrained wireless communication applications, since they have low transmission energy consumption. However, for reducing the overall energy consumption, Look-UpTable-Log-BCJR (LUT-Log-BCJR) architectures having low processing energy consumption are required. The proposed architecture reaches a low area and hence a low energy consumption. In this method, the LUT-Log-BCJR algorithm is decomposed into its most fundamental ACS operations. The architecture was validated by implementing an LTE turbo decoder and 71% energy reduction was obtained. Authors in [46] proposed an ungrouped backward recursion technique for the computation of backward state metrics. MAP decoder based on this technique can be extensively pipelined and retimed to attain higher clock frequency. The state metric normalization technique is employed in the design of an add-compare-select-unit (ACSU). This has reduced critical path delay of the decoder architecture. Turbo decoders with 8 and 64 parallel MAP decoders in 90nm CMOS technology is also designed and implemented. VLSI implementation of an 8 parallel turbo-decoder has achieved a maximum throughput of 439 Mbps and 0.11 nJ/bit/iteration energy-efficiency. Similarly, 64x parallel turbo-decoder has achieved a maximum throughput of 3.3 Gbps and 0.079 nJ/bit/iteration of energy-efficiency. These high-throughput decoders meet peak data rates of 3GPP LTE and LTE-Advanced standards.

Authors in [47] presented a turbo decoder chip supporting all 188 block sizes in 3GPP LTE standard. The design allows 1, 2, 4, or 8 SISO decoders to concurrently process each block size, and the number of iteration can be adjusted. In addition, a three-stage network is utilized to connect multiple memory modules and multiple SISO decoders. After fabricated in 90 nm process, the 2.1 mm^2 chip can achieve 129 Mb/s with 219 mW for the 6144-bit

block after 8 iterations.

In [48], authors presented the energy-efficient implementation of a high performance parallel radix-4 turbo decoder, which is designed to support multiple 4G mobile network standards such as Mobile WiMAX and LTE. The authors proposed a new hardware architecture that can share hardware resources for the two standards. It mainly includes eight retimed radix-4 SISO decoders to obtain high throughput and a dual-mode parallel hardware interleaver to support both ARP and QPP interleavers defined in the two standards. A prototype chip supporting both the WiMAX and LTE standards is fabricated in a 0.13 μm CMOS technology with eight metal layers. The decoder core occupies 10.7 mm^2 and can exhibit a decoding rate of more than 100 Mb/s with eight iterations while achieving an energy efficiency of 0.31 nJ/bit/iteration.

Authors in [49] presented a LTE compliant Turbo code decoder which provides a throughput of 150 Mb/s at 6.5 decoding iterations and 300 MHz clock frequency with a power consumption of about 300 mW. The decoder has been integrated in an industrial SDR chip in 65 nm low power CMOS process. The architecture has a very good scalability for further throughput demands. Special emphasis was put on the problem of acquisition in highly punctured LTE Turbo codes with code rates up to 0.95. The authors considerably reduced the high acquisition length needed for this code rate by implementing NII (next iteration initialization) in addition to the acquisition.

In [29], authors presented a design of a parallel turbo-decoder for the 3GPP LTE standard. The analysis of the throughput/area tradeoffs associated with parallel turbo decoders have shown that radix-4 in combination with eight M-BCJR instances are essential for the LTE peak data rate in 0.13 μm CMOS technology. Parallel and interleaved access to the memories at high throughput was achieved through the development of a master-slave Batcher network. Optimizations in the radix-4 M-BCJR unit finally led to a high performance and low area turbo decoder architecture. In addition, in order to set a record in turbo decoding throughput, both ultra low power and cost-effectiveness have been demonstrated by the presented implementation concept.

In [50], the authors presented a combined 3GPP LTE and HSDPA compliant turbo decoder which provides a throughput of 300 Mb/s at 6.5 decoding iterations and 350 MHz clock frequency with a power consumption of 452 mW. It supports decoding of highly punctured HSDPA and LTE turbo codes using a combination of acquisition and NII. The decoder occupies only an area of 1.46 mm^2 in a 40 nm low power CMOS process. A huge reduction of the memory required for the conflict resolution tables is obtained by using a hybrid method and compression of the permutation vectors.

Authors in [51] presented a highly-parallel architecture for the decoding of LTE/LTE-Advance Turbo codes. Based on the algebraic constructions, the QPP interleaver offers contention-free memory accessing capability which enables parallel Turbo decoding by using multiple MAP decoders working concurrently. They proposed a low-complexity recursive architecture for generating the QPP interleaver addresses on the fly. The QPP interleavers are designed to operate at full speed with the MAP decoders. The proposed architecture has scalable parallelism and can be sewed for different throughput requirements. With this architecture, a throughput of 1.28 Gbps is achievable with a core area of 8.3 mm^2 in a 65 nm CMOS technology.

A 650 mW Turbo decoder was presented for CAT4 LTE terminals in [52]. The design is based on algorithm/architecture co-optimization targeting a good tradeoff between performance and throughput. The results show that at feasible silicon cost, the parallel ar-

chitecture reduces significantly the decoding latency while allowing a link-level performance that is close to the traditional serial decoder to be attained.

Authors in [53] presented a parallel architecture of a turbo decoder using QPP interleaver. The proposed design can allow 1, 2, 4, or 8 SISO decoders to handle each block with configurable iterations. In order to support all data transmissions in the parallel design, a multistage network with low complexity is also used. In addition, a robust path metric initialization is given to improve the performance loss in small blocks and high parallelism. After fabrication in the 90 nm process, the 2.1 mm^2 chip can obtain 130 Mb/s with 219 mW for the 6144-bit block size and 8 iterations. The authors concluded that the proposed method and architecture facilitate the parallel turbo decoder implementation to attain both higher throughput and flexibility.

In [54], the authors presented the first LTE advanced compliant LTE turbo code decoder with a throughput of 2.15 Gb/s at frequency of 450 MHz and area of 7.7 mm^2 in a 65 nm process node with worst case P&R constraints. The proposed decoder can perform 6 full iterations at a large window size of 192 at full throughput resulting in a highly competitive communications performance.

Authors in [55] presented an analysis of LTE Turbo decoding algorithm including the MAP and improved MAP algorithm which based on SISO. The ability of MAP algorithm and improved MAP algorithm are analyzed over the length of data for code, iterative times and complexity of decoding. The authors concluded that, under the same conditions, the performance of these algorithms can be sorted by: MAP > Log-MAP > Max-Log-MAP. The performance of MAP algorithm is almost the same as the Log-MAP algorithm, but the computational complexity of Log-MAP algorithm is lower than MAP. Max-Log-MAP algorithm is a simplified Log-MAP algorithm, but BER is higher than Log-MAP. Two main factors affect the performance of Turbo decoding including: the length of Turbo coding and number of iterations.

As mentioned before, the symbol-by-symbol Log-MAP algorithm is an optimal algorithm for iterative decoding in white Gaussian noise [56]. However, this algorithm is executed in logarithmic domain or reading data from a big table will spend much time and logarithmic operations are not easy to implement in hardware. In order to reduce the high complexity of the optimal algorithm, its sub-optimal variants were proposed, such as the lookup table Log-MAP, Max-Log-MAP [57] and SOVA [58], which are utilized in practice to meet the tradeoff between performance and complexity. In these algorithms, the Max-Log-MAP algorithm has the least computational complexity, but it has the worst BER performance in comparison with the Log-MAP algorithm. It has a performance degradation about 0.4 dB [57] so will reduce about 10 % capacity of the system. Therefore, in order to improve the performance of the Max-Log-MAP algorithm while keeping the acceptable complexity, many proposals have been devoted in literatures [59], [60], [61], [62], [63]. In order to more easily observe, these Log-MAP based algorithms will be detailed in Chapter 3.

From the above analyses, it can be seen that, there are not many proposals for the LTE Turbo coder. The proposed solutions mainly concentrate on the modifications of the interleaver, rate matching, and joint source-channel coding whereas joint source-channel coding is a cross-layer optimization solution that is a trend to improve the system performance. Otherwise, for the LTE Turbo decoder, the proposed solutions mainly focus on the configurable architectures and the improved algorithms based on the Log-MAP. These are main research directions for improving Turbo decoding in LTE networks.

2.2 LTE Downlink scheduling schemes

LTE is a mobile network which has high data rate, low delay and fully packet-based. This means to improve the capability of legacy system by increasing data rates and extending superior QoS for various multimedia applications. Basic components of LTE network include a powerful eNodeB (eNB) station and several UEs in addition to a gateway [64]. The eNB combines with core network through several standard complicated protocols. Basic packet scheduling is carried out by the network operator in both UE and eNB station for both uplink as well as downlink. However, according to the 3GPP, there are no firm specifications for scheduling technique in LTE network. One of the most important modules of packet scheduling is RRM which decides users that would transmit their data on the air interface. The packet scheduling should integrate fairness in terms of throughput as well as the service policies to which users subscribe [65]. In order to meet different QoS requirements for these groups, several packet scheduling algorithms have been proposed.

In this section, we present literature review on different allocation strategies introduced for LTE systems, highlighting pros and cons related to each solution. According to [66] and [67], the scheduling strategies for LTE downlink are divided into five groups as follows:

1. **Channel-unaware strategies:** which were firstly introduced in wired networks and are based on the assumption of time invariant and error-free transmission media. In LTE networks, they are typically used in conjugation with channel-aware strategies to improve system performance
2. **Channel-aware/QoS-unaware strategies:** which allocate resources with optimal algorithms by taking into account the channel conditions. The channel quality can be estimated from CQI reports which help the scheduler to estimate the channel quality perceived by each UE and serves as an indication of the data rate which can be supported by the downlink channel
3. **Channel-aware/QoS-aware strategies:** As presented before, QoS differentiation in LTE is managed by associating a set of QoS parameters to each flow. Minimum required performance can be ensured by the scheduler if it knows the values of QoS parameters, either in terms of guaranteed data rates or of delivery delays.
4. **Semi-persistent scheduling for VoIP support:** which allow optimizing the performance of dynamic scheduling when VoIP traffic is present in a LTE heterogeneous network.
5. **Energy-aware strategies:** which limit energy waste and hence extend the battery life of UE. They can be applied to both eNB and UE. Power consumptions can be limited through DRX procedures and the persistent allocation.

The review on literature level of the above groups is represented in the following subsection.

2.2.1 Channel-unaware scheduling strategies

The possibility of direct application of scheduling algorithms in this group in LTE is not realistic, they are typically used jointly with channel-aware approaches to improve system performance.

2.2.1.1 First In First Out

This is the simplest case of channel unaware allocation policy serving users according to the order of resource requests, exactly like a First In First Out (FIFO) queue [68]. This means user which arrived at the eNB buffer first served first. This method is very simple, but is not efficient and not fair. Its priority metric depends on the current time t and the time instant T_i when request was issued by i -th user.

2.2.1.2 Round Robin

It performs fair sharing of time resources among users. Its priority metric is similar to the one defined for FIFO with the difference that, in this case, T_i refers to the last time when the user was served. This approach is not fair in terms of user throughput because in wireless networks, does not depend only on the amount of occupied resources, but also on the channel conditions. This technique is not practical in LTE because different terminals have different service with different QoS requirements [69].

2.2.1.3 Blind Equal Throughput

This technique aims at providing throughput fairness among all the users. To counteract the unfair sharing of the channel capacity, the Blind Equal Throughput (BET) scheduler uses a priority metric which considers past average user throughput [70]. This means users with lower average throughput in the past having the higher priority and and more resource blocks allocated for users with the bad channel quality. This allow achieving throughput fairness but at the cost of spectral efficiency.

2.2.1.4 Weight Fair Queuing

In this approach, the packets are grouped in various queues and each queue is assigned a weight to determine the fraction of the total bandwidth available to the queue [71]. This technique is a modification of the RR approach. It guarantees that flows with larger packets are not allocated more bandwidth than flows with smaller packets, it also supports variable-length packets. The Weighted Fair scheduling assigns the bandwidth for each service based on the weight assigned to each queue and not based on the number of packets.

2.2.1.5 Largest Weighted Delay First

In order to avoid packet drops, each packet has to be received within a certain delay deadline in Guaranteed delay services. It incorporates the information about the specific packet timing, when the packet was created and its deadline while calculating the priority metric [72]. Its priority metric depends on waiting time of the packet at the head of the line, drop probability, and target delay for each user. This technique does not take into account channel conditions, thus, it is poor in throughput.

2.2.2 Channel-aware and QoS-unaware scheduling strategies

These strategies allocate resources with optimal algorithms by taking into account the channel conditions. The channel quality can be estimated from CQI reports which help the

scheduler to estimate the channel quality perceived by each UE and serves as an indication of the data rate which can be supported by the downlink channel.

2.2.2.1 Maximum Throughput

This approach takes into account advantage of multiuser diversity to maximizing system throughput by assigning each RB to the user that can obtain the maximum throughput in the current TTI [70]. This also allows maximizing cell throughput but it performs unfair resource sharing since users with poor channel conditions will only get a low percentage of the available resources. For a practical scheduler, there need be a trade-off between the maximum throughput and the maximum cell throughput.

2.2.2.2 Proportional Fair

This technique can improve the fairness among users without losing the efficiency in terms of average throughput. The UEs are classified according to the priority metric which is defined as the ratio of the instantaneous to average throughput. Then the scheduler assigns resources to UE with the highest priority. It allows finding a balance between requirements on fairness and spectral efficiency. The Proportional Fair (PF) was designed specifically for the non real-time services, thus, does not assure any QoS requirement such as delay, jitter and latency. The priority metric is a combination between MT and BET approaches.

Several extensions of this technique have been proposed in literature. Authors in [73] evaluated the performance of PF scheduling for both time and frequency domain in OFDMA wireless systems. The performance is investigated for a broad range of the traffic load and the number of subbands. In [74], authors presented an optimization problem for PF to maximize the obtained throughput of LTE system. Herein, a multiuser scheduler with PF was proposed. A suboptimal PF scheduler, which has much lower complexity at the cost of throughput reduction was also proposed. The simulation results showed that the proposed PF scheduler provides a superior fairness performance with a modest loss in throughput, as long as the user average SINRs are fairly uniform.

2.2.2.3 Throughput to Average

This algorithm [75] tries to divide the available resources among all users. The priority metric performs averaging of resources evenly among users. The achievable throughput in the current TTI is used as normalization factor of the achievable throughput on the considered k -th RB. It is clear that the higher the overall expected throughput of an user is, the lower will be its metric on a single RB.

2.2.2.4 Joint Time-Frequency domain

This is a two-step technique for distributing radio resources [76] as follows: (1) a TDPS selects a subset of active users in the current TTI among those connected to the eNB, and (2) RBs are physically allocated to each user by a FDPS. The final allocation decision is the outcome of the decisions that work in series. The main advantage of such a partitioning is that the computational complexity at the FDPS is reduced, due to the number of candidate users for resource allocation decreases. This technique is to obtain fair sharing of time resources among users and to achieve a good trade-off between spectral efficiency and fairness.

2.2.2.5 Buffer-aware

This approach takes into account buffer management to avoid packet losses. Authors in [77] presented the Buffer-Aware Traffic-Dependent (BATD) scheme to deal with the packet dropping probability due to a receiver buffer overflow; they aim at keeping this probability as low as possible whereas guaranteeing high total system throughput and a certain level of fairness. BATD makes usage of buffer status information reported by the UE to the eNB and of traffic statistics for setting dynamic priorities associated to each MAC queue.

2.2.3 Channel-aware and QoS-aware scheduling strategies

As mentioned previously, QoS differentiation is managed by associating a set of QoS parameters to each flow. Minimum required performance can be guaranteed by the scheduler if it knows the values of QoS parameters, either in terms of guaranteed data rates or of delivery delays. Several typical approaches in this group are represented as follows:

2.2.3.1 Guaranteed Data Rate

Authors in [78] proposed a QoS oriented both time and frequency domain that focuses on GBR considerations. For the time domain, the Priority Set Scheduler (PSS) has been proposed to select users with the highest priority. Specifically, users with flows below their target bitrate form a high priority set. The remainder of the users forms a lower priority set. Users belonging to first and second sets are managed by using BET and PF algorithms, respectively. Once a number of candidate users has been selected by the TDPS, the FDPS allocates available resources through the PF scheduled metric. In [79], authors proposed a DHS based on two basic components corresponding to a guaranteed and a dynamic delay based rate allocation policy. The resources are allocated to the user with the highest priority. This is similar to the users with the lower priority if the RBs left free and so on. Authors in [80] presented a MAC scheduling technique with two stages: Time Domain (TD) and Frequency Domain (FD) schedulers. The TDPS differentiates the users according to their QoS characteristics whereas FDPS assigns the RBs among the priority users. The incoming packets are classified based on QCI corresponding to GBR and non-GBR sets and their priority order. Next, the FDPS orderly assigns the best RB to each user in the GBR set, updating the achieved bitrate. When all users in the list have reached their target bit-rate, if RBs are still available, the scheduler assigns them to users in the non-GBR list using PF metric.

2.2.3.2 Guaranteed Delay Requirements

Real-time flows have more constraints of delay than non real-time flows resulting in the reduction of influence of error correction. These scheduling strategies aim at guaranteeing the bounded delay falling in the category of the QoS-aware schemes.

The M-LWDF combines both channel conditions and the state of the queue with respect to delay in making scheduling decisions [81]. It guarantees that the probability of delay packets does not exceed the discarded bound below the maximum allowable packet loss ratio. The scheduler allocates resources to the user with the maximum priority metric is computed based on the HOL packet delay of the user, the channel capacity with respect to flow and the QoS differentiating factor. This is main scheduling which is strongly related

to our proposed scheduling schemes in Chapter 4. Hence, in order to observe easily, this scheduler will be described in Chapter 4.

Authors in [82] proposed an approach called EXP/PF which is a QoS-aware extension of PF that can support both non real-time and real-time flows at the same time. For each flow, there is a unique priority metric. This means the EXP/PF takes into account both the characteristics of PF and of an exponential function of the end-to-end delay. Real-time users are prioritized over non real-time ones when their HOL packet delays are reaching the delay deadline. The exponential term is closer to 1 if HOL delays of all users are as same as. This scheduler is used in this thesis to compare the performance to our propose scheduling schemes. Hence, it is also detailed in Chapter 4. EXP rule [83] can be considered as modified form of the EXP/PF. It also has separate priority metrics for real-time and non real-time flows. Authors in [84] evaluated the performance of EXP rule in comparison with the one of PF and L-MWDF schedulers. A variant of the EXP rule (called EXPQW) was proposed which assigns weights to the eNBs based on their queue length and waiting time. This approach uses a combination of the exponential rule for waiting time and queue length and other scheduling rules. In [85], authors proposed a two level resource allocation scheme which enhances the QoS for multimedia services. It corresponds to a procedure that combines cooperative game theory, a virtual token mechanism, and the EXP rule algorithm. This technique offers a significant performance gain over the EXP rule is achieved in terms of both PLR and fairness index. LOG rule technique was proposed in [86] with a new priority metric defined based on tuneable parameters, spectral efficiency, and HOL delay. In order to enhance real-time flows in LTE networks, a two-level downlink scheduling called FLS was proposed in [87]. At the highest level, a discrete time linear control law is applied every LTE frame. The total amount of data that real-time flows should transmit in 10 ms is calculated before considering their delay constraints. When FLS completes its task, the lowest layer scheduler works every TTI. The lower PF algorithm allocates radio resources by considering bandwidth requirements of FLS to flows hosted by UEs having the best channel quality. Specifically, the lowest layer scheduler decides the number of TTIs/RBs where each real-time source will actually transmit its packets. The resources left free by real-time flows are assigned to non real-time flows. In this thesis, we also use this scheduler to compare the performance to the one of our proposed scheduling schemes. Therefore, more detailed description of the priority metric will be represented in Chapter 4.

2.2.4 Dynamic and Semi-persistent Scheduling for VoIP support

In [88], authors proposed a dynamic packet scheduling for VoIP traffic in LTE downlink. The goal is to optimize the performance of dynamic scheduling when VoIP users are present in a heterogeneous traffic. The proposed algorithm is divided into both time and frequency domain. In time domain, at every TTI, the scheduler called as Required Activity Detection with Delay Sensitivity (RAD-DS) prioritizes each scheduled user according to the time domain metric. In the frequency domain, the scheduler allocates RBs to different users using the metric of the PF scheduler.

For enhancing VoIP traffics when its diversity is high, semi-persistent allocation solutions aiming at increasing the VoIP capacity of the network by maximizing the number of VoIP calls. Authors in [89] improved VoIP capacity of the network with the usage of semi-persistent scheme. The radio resources are divided in several groups of RBs. Each pre-configured block is associated only to certain users. Besides, RB groups are associated to

each user in consecutive TTIs. Resource allocation of each RB group to the associated UEs is performed dynamically. The proposed scheme reduces the control overhead with respect to the dynamic scheduling.

Authors in [90] proposed to couple VoIP users in pairs, thus, they share the same persistently allocated resources. The idea is that pre-allocated resources to each user pair are shared by the same users depending both on the channel conditions and on the experienced PLR.

2.2.5 Energy-aware

In LTE networks, energy consumption is heavy due to terrible processing load on UE. Energy preserving solutions limit energy waste and hence extend the battery life of UE through DRX procedures and the persistent allocation. In DRX, when there are no data transmissions, UE turns off its radio equipment to save energy.

In [90], authors introduced the light sleeping mode to improve the performance of DRX for QoS-aware traffic. The main idea is to turn off the power amplifier. Other components in transceiver cut down their power consumption whereas allowing fast wake-up. The proposed scheme reduces energy consumption whereas satisfying the delay constraints. Authors in [91] analysed the effect of different scheduling schemes from viewpoint of energy efficiency. The analyses showed that the Maximum Throughput scheme is more energy efficient than both PF and RR. In case of low traffic load, Bandwidth Expansion Mode (BEM) algorithm is used to obtain energy savings for the eNB [92]. The eNB transmission power is reduced by assigning a coding scheme with lower rate to each user.

2.3 QoE evaluation approaches for voice user

In communications systems, the perceived voice quality is usually represented as MOS. MOS can be attained by many methods. These methods are divided into two groups called subjective methods and objective ones.

1. **Subjective methods:** Humans listen to a live stream or a recorded file and rating it on a ratio of 1 (poor) to 5 (excellent) [93]. These methods have some disadvantages such as too expensive, time consuming and are not suitable for a large network infrastructure. Otherwise, objective methods have more advantages, they eliminate the limitations of subjective methods.
2. **Objective methods:** These methods are not based on customer surveys and are classified into two approaches: intrusive and non-intrusive ones.
 - **The intrusive methods** (e.g. Perceptual evaluation of speech quality (PESQ) [94], Perceptual evaluation of speech quality for wideband audio (PESQ-WB) [95]) are more exact and are widely utilized to predict aware voice quality. However, they are not suitable for real-time services such as VoIP because they require original signals to refer.
 - **The non-intrusive methods** (e.g. ITU-T E-model [31], ITU-T WB E-model [96]) are computational models that are used for transmission planning purposes. They are not as accurate as the intrusive approaches and they do not have complex mathematical operations. The obtained results from objective methods do

not always well relate to human perception. The main advantage of the non-intrusive methods are they predict voice quality without any reference to the original signals and they require less parameters than the intrusive methods. For several typical non-intrusive methods, authors in [97] proposed to use RNN to assess voice quality over internet. Voice quality assessment was predicted in [98] using RNN. Another non-intrusive method was proposed in [99] based on RNN for evaluating video quality in LTE network. Authors in [5] proposed to use WB E-model to predict VoLTE quality for minimizing redundant bits generated by channel coding. In [100], authors investigated the effects of PLR and delay jitter on VoIP quality to assess prediction errors of MOS for the E-model. A new model called “Packet-E-Model” was proposed in [101] to measure speech quality perception for VoIP in Wimax network. In [102], authors presented a voice quality measurement tool based on the E-model. A framework of objective assessment method for estimating conversational quality in VoIP was proposed in [103]. In [104], a simplified versions of the E-model were proposed to simplify the calculations and focus on the most important factors required for monitoring the call quality. According to our knowledge, at present, there are not any proposals which allow to predict VoLTE quality using LTE-Sim framework and extended E-model as well as wideband E-model.

VoLTE is a real-time service, so to evaluate directly its quality is a big challenge. It is clear that, for VoLTE service, the non-intrusive is the most suitable for evaluating voice quality. In this thesis, we proposed new models based on this method. Therefore, in this section, we only present literature review on the non-intrusive methods. This method includes standard models and non-standard ones such as described in the following subsection.

2.3.1 Standard models

2.3.1.1 E-model and wideband E-model

As described in Chapter 1, E-model is a computational model developed and standardized by ITU-T [105]. Originally designed as a telecommunication transmission planning tool, has become one of the most popular methods to evaluate the quality of a voice transmission system. This parametric model takes into account several tabulated transmission impairments, such as delay, echo, codec distortion, etc. E-model is used to evaluate voice quality for narrowband audio. The output of this model is Transmission Quality Rating (R-factor) with the scale of 0..100, then this factor is mapped to MOS which performs human perception. This model has some advantages and disadvantage as follows [106]:

- **Advantages:**

- Includes simple operations thus calculates quickly
- It does not refer to the original signals it is very suitable for predicting quality of VoLTE service.

- **Disadvantages:**

- Very stringent environments are required
- The process cannot be automated

- They are very costly and time consuming, which makes them unsuitable to be frequently repeated
- It does not offer exactly human perception such as intrusive methods.

It is similar to the E-model, the WB E-model is also fully described in Chapter 1. It is also developed and standardized by ITU-T [96]. It is used to measure quality of wideband audio. This is the difference of this model compared to the E-model. Besides, it also has advantages and disadvantages as ones of the E-model. The output of this model is R_{wb} -factor. Its values in range of 0..129. So it is converted to the R-factor based on Equation of $R = R_{wb}/1.29$ before mapped to user perception.

The problem of these models is how to determine its inputs. If this problem is solved, this model promises to be widely applied for measuring voice quality in LTE network. In this thesis, these models are the backbone of our proposals. The solutions including the extension or the enhancement for these models.

2.3.1.2 P.563 model

According to [106], this model describes a single-ended, signal-based method for objective speech quality assessment in narrowband audio. P.563 is able to predict the listening quality (MOS-LQO) in a perception-based scale considering the full range of distortions occurring in public switched telephone networks; therefore, this model allows real-time measurements to estimate the MOS of a voice call at any point of the path between users. However, the downside of this model is high computational complexity and requires many parameters. In order to enhance this mode, authors in [107] introduced an enhancement to this model to adapt its features to conditions of VoIP services. The authors explored mute-length, sharp-decline, and speech-interruptions as the most sensitive parameters to network impairments. These factors allow defining a new distortion class and a priority for this new class. This new class is complemented into the P.563 quality estimator. However, the authors did not provide any comparison between the accuracy obtained by the standard and the proposed model, thus, there need be additional work to demonstrate the validity of this enhancement.

2.3.1.3 ANIQUE+ model

The Auditory Non-Intrusive Quality Estimation Plus (ANIQUE+) model is an ANSI standard for non-intrusive, signal-based, estimation of narrowband speech quality. ANIQUE+ estimates the MOS-LQO of a voice call based on the functional roles of human auditory systems and the characteristics of human articulation systems [108]. The ANIQUE+ algorithm measures the overall distortion affecting the voice signal and maps this distortion to the MOS. This model obtains user perception better than P.563 model. The authors already evaluate the performance of PESQ, P.563 and ANIQUE+ models. The obtained results as follows: ANIQUE+ reached an accuracy of 93.9%, P.563 only reached 84.8% whereas PESQ obtained the best result with an accuracy of 95.3%.

2.3.1.4 P.564 model

This model is developed and is standardized ITU-T [109]. It is a conformance testing for voice over IP transmission quality assessment models. It has the provisional name is P.VTQ (Voice Transmission Quality). There are two different methodologies presented:

Telchemy's VQMon [110] and Psytechnics' PsyVoIP [111]. For the second method, it takes into account the particular features of the different manufactured edge-devices, such as VoIP phones or gateways, by using calibrated formulas and weighting coefficients to each specific device [106]. The development of the P.564 model specifies the minimum criteria for objective speech quality assessment models that predict the impact of observed IP network impairments on the one-way listening quality experienced by the end-user in IP/UDP/RTP-based telephony applications [109]. Originally specific to narrowband (3.1 KHz), the P.564 Recommendation also includes an extension for wideband (7 KHz) telephony.

2.3.2 Non-standard models

Models of this group are not standardized. For simplicity, we have divided this group into four categories as follows: (1) based on Gaussian Mixture Models (GMM), (2) based on neural network (RN), (3) based on exponential functions, and (4) hybrid models.

2.3.2.1 Gaussian Mixture-based models

Authors in [112] used GMMs to generate artificial reference models of speech behavior. This method compares distortion features introduced in these reference models to those affecting the real signal stream. Hence, a double-ended quality estimation algorithm is emulated. Next, authors introduced an enhancement to improve the quality estimation accuracy when noise suppression algorithms are incorporated. After that, the authors proposed a modification including additional information related to the transmission and coding schemes employed in the communication resulting in revealed better performance than P.563 with a significant reduction of processing time. Authors in [113], [114] proposed two different quality estimators corresponding to a quality estimation algorithm based on GMM and on Support Vector Regression (SVR) and an enhanced non-intrusive objective speech quality evaluation method based on Fuzzy Gaussian Mixture Model (FGMM) and Fuzzy Neural Network (FNN). In the first proposal, the authors used GMM to form an artificial reference model of the behavior of Perceptual Linear Predictive (PLP) features of clean speech. Consistency measures between the degraded speech and the reference model were utilized as indicators of speech quality. The effective least square SVR arithmetic was used to map the consistency values to the predicted MOS. In the second one, an improved version of the previous method was proposed based on FGMM and FNN. FGMM was employed, instead of GMM, to form the artificial reference model. FNN regression algorithm was used to map the consistency values to the predicted MOS. Both proposals had the results outperforming the P.563 model in some cases under voice codecs of G.711 and G.729.

2.3.2.2 Neural networks-based models

Neural Networks (NN) have been widely applied to estimate human perception. For this Several authors direction, several authors have developed quality estimation models for voice communications based on ANN and RNN. Authors in [115] analyzed effects on call quality of four different parameters including codec type, gender, loss pattern, and loss burstiness. In order to model the relationships between these factors and perceived speech quality, a NN model was developed to learn the non-linear mapping from these parameters to a MOS score. Authors in [116], [117] deeply studied on Pseudo-Subjective Quality Assessment (PSQA) technique. Their model an RNN-based quality assessment mechanic to assess the influence

of certain quality-affecting parameters, such as coding scheme, redundancy, packet loss rate, Mean loss Burst Size (MBS), and packetization interval on real-time listening quality. In [118], the authors proposed an enhancement of the previous model. In this work, the effects of delay, jitter, and FEC mechanisms were also taken into account. The authors concluded that three main parameters affecting the voice quality are PLR, coding bit-rate and FEC. The effects of delay or jitter is low. Authors in [119] also used RNN to capture the non-linear relation between network parameters that cause voice distortion and the perceived quality, thus, the proposed model so-called A_PSQA. It has two input parameters consisting of PLR and Mean Loss Burst Size (MLBS) of VoIP call. In this work, the effect of jitter is ignored. In order to train the RNN, authors developed a database of MOS scores for different speech samples transmitted under different loss conditions, characterized by the Gilbert model. These scores were obtained by using the PESQ model. After training the RNN, results showed very good correlation with PESQ, and outperformed the standard E-model and the IQX model.

2.3.2.3 Exponential functions-based models

Authors in [120] developed a model called IQX hypothesis. In this work, the authors assumed that an exponential functional relationship between QoE and QoS. Specifically, QoE is a function of QoS and factors of the codec. The authors concentrated on the packet loss probability to measure the quality of service using a well-known commercial VoIP application; an extension of this work was presented in [121] where delay and jitter were also taken into account to assess the QoS. The results showed good accuracy, verifying the exponential relationship between QoE and QoS. In [122], authors presented three parametric models obtained through regression analysis called Model A, Model B, and Model C, their unique input parameter is packet loss and the output is a MOS value in range of 1-5. The accuracy of these models was validated using PESQ, but no comparison with IQX or any other parametric models was performed.

2.3.2.4 Hybrid models

Hybrid model is the combination of the above presented models. This is a current trend to explore characteristics of the previous approaches and to joint them into a model. According to [106], the above models can be arranged into 3 groups as follows:

- **Signal-based models.** Allow estimating QoE by processing directly human speech, analyzing the distortion introduced in the voice signal. These models have shown to be sensitive to bursty packet loss and PLC algorithms.
- **Parametric models.** Allow estimating QoE based on assessment of different impairments introduced by network and encoding schemes, which make them sensitive to background noises or noise suppression strategies.
- **Packet-layer models.** Have not been fully developed, because of the difficulty of representing the complex interactions among the different impairments affecting the VoIP QoE just from parameters extracted from the packet headers.

For the hybrid model, authors in [123] extended the conventional parametric speech quality estimation models by considering the voice feature of lost packets. This model

builds a voice-aware speech quality model that allows accurately quantifying the effect of lost packets according to their voice property. In [124], authors presented a listening-only model based on both audio and voice distortions. Firstly, the above impairments are detected to assess the occurrences of packet loss, the loss pattern and the employed coding scheme, an analysis of the IP headers is conducted. Next, the impact of individual impairments is quantified. Lastly, the overall quality evaluation model was built by integrating individual impairments and employing the E-model. The results showed high correlation with PESQ, but several important impairments such as delay, jitter, echo, etc. were ignored. Authors in [125] extended the usage of their GMM proposals, by integrating packet header analysis. GMM is used to generate an artificial reference model that was compared to the transmitted speech signal. In addition, a parametrical analysis was carried out, evaluating the VoIP header. This model has the low computational complexity with 88% lower than the P.563 model.

Based on the above analyses, it can be seen that almost of QoE models do not take into account effects of delay and jitter buffer. Author in [106] concluded that to optimize the trade-off between voice quality and voice capacity, delay and jitter buffer are main parameters to manage. The QoE models need be improved as much as possible to have more accuracy.

2.4 Summary

In this Chapter, we present the literature review on the related topics in our proposals. We see that, for LTE channel coding, Turbo coding algorithm is rather good, thus, the proposed solutions for this are mainly focus on improving the interleaver or rate matching. One of trends for enhancing LTE channel coding is joint source-channel coding. This enhancement allows improving significantly decoding quality. For Turbo decoding, its main mission is to decode the coded bits by Turbo coding. In the decoding algorithms, a part which is very important is the error correcting function. Almost the proposed solutions for enhancing Turbo decoding focus on it. This is understandable because it decides not only the decoding efficiency but also the computational complexity of the decoding algorithm. For the scheduling techniques, almost the proposals concentrate on downlink direction. In the scheduling strategies, the category of Channel- and QoS-Aware is suitable for real-time services such as voice calls. QoE-based scheduling approaches are new trends and need to be developed in the future. The semi-persistent approaches are also suitable for VoIP service which needs to have a special priority when transmitted in an All-IP network such as LTE. In order to guarantee QoE of end users, voice call quality need to be monitored and evaluated to adjust network impairments opportunely. For real-time services such as voice calls, the object non-intrusive models are very suitable to predict voice quality because they do not require any reference to original signal.

Chapter 3

The proposed solutions based on wideband E-model and polynomial regression function for enhancing LTE channel codec

Contents

3.1	General idea	48
3.2	The proposed algorithm for minimizing redundant bits generated by channel coding	49
3.2.1	The calculation of the delay impairment factor	49
3.2.2	The calculation of the equipment impairment factor	50
3.2.3	The proposed joint source-channel code rate adaptation algorithm	52
3.2.4	Simulation results and Performance evaluation	53
3.3	Enhancing channel decoding algorithm based on polynomial regression function	57
3.3.1	Decoding algorithms in Turbo codes	57
3.3.1.1	The Log-MAP algorithm	57
3.3.1.2	The Log-MAP-based algorithms	58
3.3.1.3	The Non-linear Log-MAP algorithm	59
3.3.2	The proposed Log-MAP algorithm based on polynomial regression function	59
3.3.2.1	The proposed error correction function	59
3.3.2.2	The simulation model	61
3.3.3	Simulation results and Performance evaluation	63
3.4	Summary	65

This Chapter represents the solutions for enhancing LTE channel codec (Coder and Decoder). For the channel coder, we propose a joint source-channel code rate adaptation algorithm for reducing redundant bits which are generated by channel coding with a slight degradation of voice quality. Data is coded by channel coding at sending side and is decoded by channel decoder at receiving side. For the channel decoder, we propose a new function of error correction that plays an important role in decoding process. The goal of this proposal is to replace the error correction in the Log-MAP algorithm which has the best

BER performance but has high computational complexity. The proposed Log-MAP based algorithm has the BER performance that is closest to the Log-MAP but has the simpler complexity.

The remainder of Chapter is structured as follows: The general idea is described in Section 3.1. Section 3.2 represents the algorithm of joint source-channel code rate adaptation for minimizing the redundant bits generated by channel coding as well as its performance evaluation. The proposed Log-MAP based algorithm and its performance assessment are described in Section 3.3. In Section 3.4, we conclude the attained results of the proposals in this Chapter.

3.1 General idea

When voice service is transmitted over a wireless channel, it has to be coded by channel coding. This helps data to be protected by external effects such as noise, fading or interference, etc. For data channels in LTE network, Turbo code is used for channel coding purpose. In Turbo encoder, each information bit stream is a k -bits block. This block size is in range of 40-6144 bits. So each block can include one or several speech packets. A k -bits block is encoded into a n -bits codeword. So that, the number of redundant bits of LTE channel coding is $n - k$ bits and channel rate is k/n . As mentioned before, channel rate are chosen based on TBS from MAC layer and channel conditions. The MCS is chosen from a pre-defined table. This can lead to more redundancy in the case of channel quality sent back to the eNB from UE is not exactly. In addition, AMR-WB also adapts its bitrate according to channel quality while size of TBS depends on the source bitrate. For this reason, there should be a combination between source and channel code rate to minimize the redundancy of channel coding. Hence, new solutions for reducing redundant bits generated by channel coding is very necessary. This could be paid by a slight reduction of user perception. For this aim, an algorithm of dynamic source-channel code rate was proposed. The key idea of the proposed solution is that the channel rate is chosen based on user perception which is estimated and is sent back to the eNB. The proposed solution based on several previous proposed techniques to improve the speech quality delivered over a noisy channel. Examples include the proposals of [126], [127], [128], [129], [130], and [131]. But the closest work related to our proposal are represented in [126], [130], and [131]. The authors in [126] presented a dynamic joint source-channel coding rate adaptation algorithm for VoIP using AMR codec. The algorithm computes the optimal rates allocated to each frame for a set given QoS constraints. The aim of their paper is to find the tradeoff between packet loss recovery and end-to-end delay to maximize perceived speech quality. In [130], the authors proposed an optimization issue for supplying unequal error protection of speech frames according to their importance. An optimization framework for identifying the optimal joint source-channel code rate of each voice frame based on the frame perceptual importance was proposed in [131]. In that paper, the quality of the received speech signal is maximized. Our solution is the combination, extension, and adaptation of them for voice transmission in LTE network.

When data is coded by Turbo encoder at sending side, it has to be decoded by Turbo decoder at receiving side. In channel decoding algorithms, error correction function plays a very important role. Some well-known decoding algorithms for data channels are the MAP-based algorithms including the Log-MAP, and Max-Log-MAP whereas the Log-MAP has the best BER performance while the Max-Log-MAP has the worst performance. The

disadvantages of the Log-MAP is that its computational complexity is high. Therefore, a trade-off between BER performance and computational complexity is essential. This means the Log-MAP based solutions are trends and they have to ensure two goals as follows: (1) they have the BER performance that is closed to the one of the Log-MAP, and (2) they have the simpler computational complexity than the Log-MAP. The main idea of the proposed solution is the application of polynomial regression function to determine the approximated error correction function. This means a new error correction function was proposed to replace the one of the Log-MAP algorithm.

The proposed solutions in this Chapter are described in the following sections.

3.2 The proposed algorithm for minimizing redundant bits generated by channel coding

As mentioned in Chapter 1, in the WB E-model, the value of R_{wb} -factor depends on both impairment $I_{d,wb}$ and $I_{e,eff,wb}$ factors and they have the direct relationship with end-to-end delay and packet loss. Increasing end-to-end delay leads to decreasing the MOS and reversing. Therefore, finding out the suboptimal joint source-channel code rate solutions is very essential. So in this section, we present another viewpoint of choosing the suitable channel code rate corresponding to each mode of AMR-WB codec for minimizing the number of redundant bits generated by channel coding with an acceptable MOS reduction.

3.2.1 The calculation of the delay impairment factor

In order to calculate this factor, we have to compute the end-to-end delay. According to [132], the end-to-end delay (D_{e2e}) can be counted as follows:

$$D_{e2e} = D_{enc} + D_{network} + D_{play} \quad (3.1)$$

Where:

- D_{enc} : The delay time caused by encoding and packetizing at the AMR-WB encoder, $D_{enc} = k \times T \times f + l_a + D_{pack}$

With:

- T : The speech frame size, $T = 20$ ms.
- f : The number of frames of a speech packet.
- l_a : The look-ahead delay, $l_a = 5$ ms for all modes of AMR-WB codec.
- D_{pack} : The packetization delay for grouping f frames into one speech packet, i.e. $D_{pack} = (f - 1) \times T$.
- $D_{network}$: The sum of transmission delay, propagation delay and queuing delay at each hop h in the network path from the transmitter to the receiver. The transmission delay (T_h) is computed using the following equation [132]:

$$T_h = (n - k + 1) \times (T \times f \times R_s + H_{overhead}) \times \sum_h \frac{1}{B_h} + \sum_h (P_h + Q_h)$$

In which:

- R_s : Bitrate (mode) of AMR-WB codec.

- $H_{overhead}$: The number of overhead bits introduced by headers of RTP/UDP/IPV6, and PDCP/RLC/MAC, and 24 bits introduced by CRC at Physical layer.
- B_h : The bandwidth at hop h .
- P_h : The propagation delay depends on the distance from the source to destination. It is negligible if within a local area. For intra-continental calls, the propagation delay is in the order of 30 ms and for inter-continental calls, it can be as large as 100 ms [132]. It is clear that for a given voice connection, the only random component of voice delay (that is the only source of jitter) consists of queuing delay in the network [132], $Q = \sum_h Q_h$.
- D_{play} : The playback delay, voice packets are usually delayed in a jitter buffer and the fixed playback delay must be equal to at least two speech frame length [133], i.e. $D_{play} = 2 \times T$.

From above analyses, we get the end-to-end delay determined as the following equation:

$$D_{e2e} = k \times T \times f + l_a + (n - k + 1) \times (T \times f \times R_s + H_{overhead}) \times \sum_h \frac{1}{B_h} + \sum_h (Q_h + P_h) + 2 \times T + D_{rohc} + D_{harq} \quad (3.2)$$

With D_{rohc} : The delay of RoHC processing time at PDCP layer. According to [134], RoHC should not noticeably add to the end-to-end delay and according to [135], this delay is not very significant, approximately 10 - 67 μs /packet to compress and 12 - 51 μs /packet to decompress RoHC packets. Thus, in this study, we consider the delay caused by RoHC equal to 0. D_{harq} : The delay due to retransmission at MAC layer by HARQ. Each voice packet is retransmitted at least once. According to [136], because the RTT (round trip time) of HARQ is fixed and because of the higher priority for retransmissions, the HARQ delay is normally within 10 ms.

Normally, FEC also can cause the delay. However, several authors have pointed out that FEC does not introduce any delay excepting packet loss. So that, in this study, we have not considered the delay caused by FEC. The $I_{d,wb}$ factor is calculated according to the following equation:

$$I_{d,wb} = 0.024 \times D_{e2e} + 0.11 \times (D_{e2e} - 177.3) \times H(D_{e2e} - 177.3) \quad (3.3)$$

In which: $H(x)$ is the Heavyside function:

$$H(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{otherwise} \end{cases} \quad (3.4)$$

3.2.2 The calculation of the equipment impairment factor

In order to compute this factor, we have to count the packet loss rate after FEC schemes try to recover errors. We assume that the estimates for the packet loss rate P_{pl} on the end-to-end network path is available at time an adaptation decision is being made. In order count the P_{pl} , we use a random loss model, the relationship between the parameters (k, n)

of FEC schemes, “raw” packet loss rate on the end-to-end network path p_r and the packet loss rate P_{pl} is described as follows [134]:

$$P_{pl} = \sum_{i=n-k+1}^n \binom{n}{i} \times p_r^i \times (1 - p_r)^{n-i} \times \frac{i}{n} \quad (3.5)$$

Equation (3.5) shows that, we can attain the current p_r from the given measurement P_{pl} and a pair of (k, n) . This p_r value is counted once per adaptation period and is utilized in the proposed algorithm as shown in the next subsection.

After calculating the P_{pl} , the $I_{e,eff,wb}$ factor can be computed as follows: The $I_{e,eff,wb}$ is determined according to packet loss. In this study, packet loss probability is estimated at the receiver, after FEC. The output bits of AMR-WB encoder will be encoded by channel coding (Turbo code). The bits in class A, class B and class C of AMR-WB codec can be encoded with different channel code rates. In addition, 8-bits CRC code is applied to protect class A bits. In figure 3.1, the packet loss represents the average rate of speech frames for which CRC check fails in class A bits. It is determined for each mode of AMR-WB codec. According to [33], $I_{e,eff,wb}$ is determined as follows:

$$I_{e,eff,wb} = I_{e,wb} + (129 - I_{e,wb}) \times \frac{P_{pl}}{P_{pl} + B_{pl}} \quad (3.6)$$

In which:

- $I_{e,wb}$: The respective impairment factor without any packet loss.
- P_{pl} : Packet loss rate.
- B_{pl} : A codec-specific factor which characterizes its robustness against packet loss.

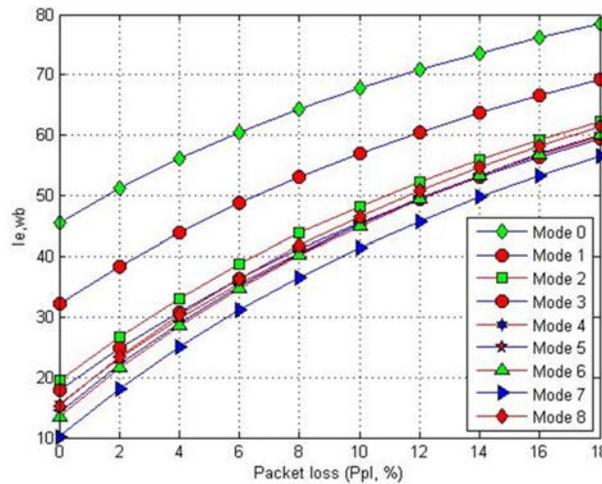


Figure 3.1: $I_{e,wb}$ vs. Packet loss for nine modes of AMR-WB codec

3.2.3 The proposed joint source-channel code rate adaptation algorithm

In this algorithm, we solve the problem of choosing suitable source and channel code rate within some constraints on maximum allowed end-to-end delay, maximum permitted packet loss and minimum required bandwidth for minimizing the number of redundant bits generated by channel coding with an acceptable MOS reduction. This algorithm is located at the transmitting side. The inputs of the algorithm include:

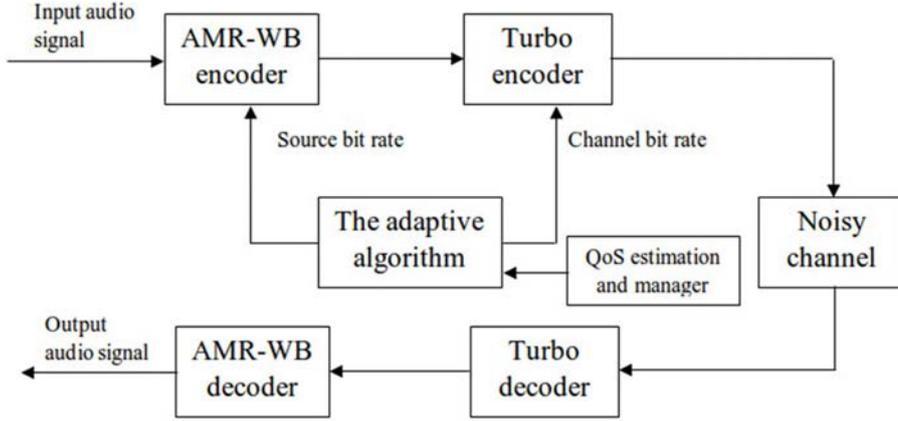


Figure 3.2: A proposed simple framework of voice traffic over LTE network

- QoS information: As path packet loss P_{pl} , path bandwidth B_h and congestion Q_h . These values can be obtained by QoS estimation module used in the network.
- QoS constraints: Maximum allowed end-to-end delay D_{max} , maximum permitted packet loss rate P_{max} and minimum required bandwidth BW_r .

The outputs of the algorithm will be a decision on the choice of source code rate R_s and channel code rate R_c . Besides the most suboptimal choice of channel code rate for every AMR-WB codec mode, the algorithm also offers the best suboptimal choice of channel code rate for all of AMR-WB codec modes. In order to describe the algorithm, we start by setting two the following constraints:

$$\begin{cases} D_{e2e} \leq D_{max} \\ P_{pl} \leq P_{max} \end{cases} \quad (3.7)$$

For each mode in AMR-WB codec modes, we change k and n over ranges $k = 1, 2, \dots, k_0$ and $n = k + 1, \dots, n_0$ where k_0 and n_0 are maximum available values of channel code rate in LTE network ($n_0 > k_0$). In order to find the most suitable suboptimal pair of (k_{subopt}, n_{subopt}) for all values of R_s , we define the MOS reduction ($MOS_r, \%$) and the percentage of decreased redundant bits ($G_r, \%$) as follows:

$$MOS_r = \frac{MOS(k_{best}, n_{best}) - MOS(k_{subopt}, n_{subopt})}{MOS(k_{best}, n_{best})} \times 100 \quad (3.8)$$

$$G_r = \frac{(n_{best} - k_{best}) - (n_{subopt} - k_{subopt})}{n_{best} - k_{best}} \times 100 \quad (3.9)$$

We see that the larger MOS_r , the lower speech quality, and the higher reduced redundant bits, the higher BER and this leads to the lower speech quality. So that, in order to find

out the tradeoff between MOS_r and G_r , we propose the criteria for this as follows: In the pairs of (k_{subopt}, n_{subopt}) of all modes, the algorithm firstly finds a pair which has the maximum value of G_r meeting $MOS_r \leq 1\%$ to ensure that this MOS is very close to the highest MOS. In the case there are many the same maximum values of G_r , the algorithm will secondly choose a pair of (k_{subopt}, n_{subopt}) which has the lowest value of the MOS_r . If there are many the same maximum values of MOS_r , then the algorithm will choose a pair of (k_{subopt}, n_{subopt}) which has the lowest value of BW_{subopt} where BW_{subopt} is the required bandwidth corresponding to the pair of (k_{subopt}, n_{subopt}) . The required bandwidth of each mobile user (BW_r) must be met the following condition:

$$BW_r \geq \frac{n}{k} \times (R_s + \frac{H_{overhead}}{T \times f}) \quad (3.10)$$

In which: all factors are described similarly as in the previous equations.

Finally, the proposed will offer the best suboptimal solution for minimizing the redundant bits generated by channel coding with an acceptable MOS reduction. The steps of the proposed algorithm are described in Algorithm 1.

The proposed algorithm will find the suboptimal solutions for each mode of AMR-WB codec and the best suboptimal solution on entire modes of AMR-WB codec. This solution meets some constraints on the allowed maximum end-to-end delay, the allowed maximum packet loss, and ensures the tradeoff between the MOS reduction and the percent of reduced redundant bits generated by channel coding.

3.2.4 Simulation results and Performance evaluation

In order to simulate the algorithm, we assume simulation parameters as follows:

- The number of speech frames per packet: $f = 1$
- The network path includes 15 hops, with 13 hops are the fast core network links: 10 hops at 622 Mb/s, and 3 hops at 1.8 Gb/s, and two are in the eNodeB cell at 50.4 Mb/s and 25 Mb/s, respectively.
- $P_h = 0.06$ ms per hop. Q_h is random between 0 and 1 ms.
- $D_{max} = 150$ ms. $P_{max} = 10^{-2}$
- $k_0 = 4$ and $n_0 = 5$

Table 3.1: The detailed results of figure 3.3

Mode	k_{best}	n_{best}	MOS_{best}	k_{subopt}	n_{subopt}	MOS_{subopt}	$MOS_r(\%)$	$G_r(\%)$	$BW_{subopt}(\text{kbps})$
0	1	5	3.51	3	5	3.47	1.14	50.00	17.26
1	1	5	3.98	3	5	3.95	0.75	50.00	21.10
2	1	5	4.33	3	5	4.32	0.23	50.00	27.59
3	1	5	4.35	3	5	4.33	0.46	50.00	30.32
4	1	5	4.40	3	5	4.39	0.23	50.00	33.05
5	1	5	4.43	3	5	4.42	0.23	50.00	37.15
6	1	5	4.45	3	5	4.43	0.45	50.00	39.88
7	1	4	4.48	3	5	4.47	0.22	33.33	45.34
8	1	4	4.42	3	5	4.40	0.45	33.33	46.70

Figure 3.3 and table 3.1 show the results in the case “raw” packet loss is fixed equal to 10%. This demonstrates that when the “raw” packet loss is fixed, the best MOS depends

Algorithm 1 A dynamic rate adaptation algorithm using WB E-model

Step 1: For all $R_s \in \{\text{mode } 0, \dots, \text{mode } 8\}$:

For $k = 1$ to k_0

For $n = k + 1$ to n_0

- Compute the end-to-end delay using equation (3.2).
- Compute the packet loss rate using equation (3.5).

Step 2: Find all triples $(R_{s,i}, k_i, n_i)$, $i = 1, \dots, t$ meeting the conditions in equation (3.7).

Step 3: Compute $I_{d,wb}$, $I_{e,eff,wb}$, R_{wb} , MOS , and BW_r using the equations (3.3), (3.6), (1.5), (1.3), and (3.10) respectively.

Step 4: Find the highest MOS of each AMR-WB codec mode corresponding to each pair of (k_{best}, n_{best}) .

Step 5: Find the suboptimal selection for each AMR-WB codec mode for reducing redundant bits of channel coding:

For $k = 1$ to k_0

For $n = k + 1$ to n_0

- Find the last pair of (k, n) which has the highest MOS, then mark $k_{subopt} = k$

For $k = 1$ to k_{subopt}

For $n = 1$ to n_0

if $n > k$ then

- Find the first pair of (k, n) which has the highest MOS, then mark $n_{subopt} = n$.
- Find the suboptimal MOS corresponding to each pair of (k_{subopt}, n_{subopt}) .

Step 6: Let $S = (R_{s,i}, k_i, n_i)$, $i = 1, \dots, 9$ denote the set of nine suboptimal solutions for nine AMR-WB codec modes. Finding the best suboptimal solution in set S which meets sequentially all the following constraints: (1) has the highest value of G_r , (2) has the lowest value of MOS_r , and (3) has the lowest value of BW_{subopt} . In the case three these conditions are met, we propose to choose the triple that has the lowest AMR-WB mode because normally the lower AMR-WB mode, the less required bandwidth.

on only two channel code rates, and one for the suboptimal MOS. Figure 3.3 also shows that with the “raw” packet loss is fixed equal to 10%, all of AMR-WB codec modes obtain the best MOS with channel code rate equal to 1/5 and 1/4, and the suboptimal MOS with channel code rate equal to 3/5. The results show that the suboptimal MOS is very close to the best MOS while the redundant bits generated by channel coding decreased significantly. In this case, the G_r of all triples have two different values (50% and 33.33%). There are only modes of 1 to 6 that have the highest values of G_r , thus, the algorithm will choose according to the lowest values of MOS_r . There are modes of 2, 4, and 5 that have the lowest values of MOS_r (0.23%), the algorithm will choose according to the third condition that has the lowest value of BW_{subopt} . So that, the triple $(R_{s,2}, 3, 5)$ is the best suboptimal solution because it has the lowest value of MOS_r with the suboptimal MOS equal to 4.32 while the best MOS equal to 4.33.

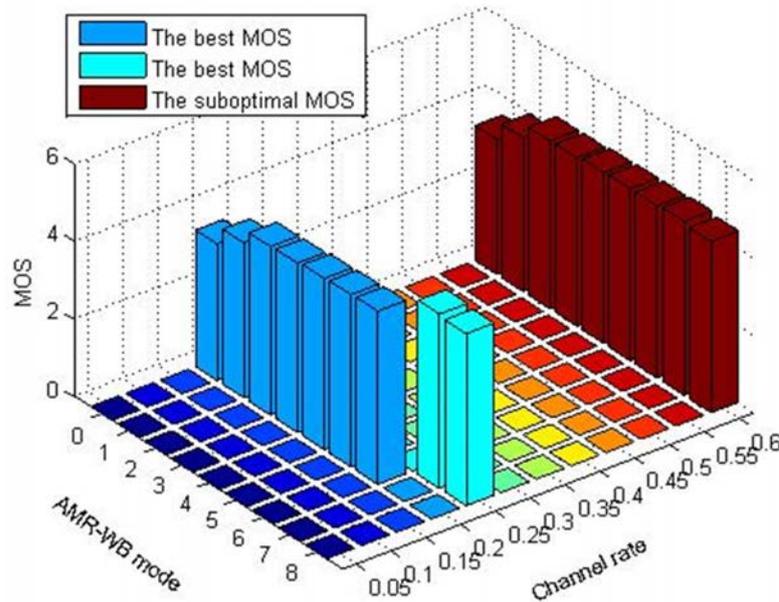


Figure 3.3: MOS vs. AMR-WB mode and channel code rate when p_r is fixed equal to 10%

Figure 3.4 and table 3.2 show the results in the case the “raw” packet loss is varied in an instant. These results demonstrate that when the “raw” packet loss is changed, the best MOS depends on only two specific channel code rates (1/4 and 1/5), while the suboptimal MOS depends on three channel code rates (3/5, 3/4, and 2/4). The results also show that there are 7 triples which have the same highest value of G_r (50%), those are two triples of $(R_{s,1}, 3, 5)$, $(R_{s,2}, 3, 4)$, $(R_{s,3}, 3, 5)$, $(R_{s,4}, 3, 5)$, $(R_{s,6}, 2, 4)$, $(R_{s,7}, 3, 5)$ and $(R_{s,8}, 3, 5)$ which meet the conditions. Similar to the case of the fixed p_r , the algorithm will firstly choose triples that have the highest values of G_r , secondly the lowest values of MOS_r and thirdly the lowest values of BW_{subopt} . So that, in this case, the best suboptimal solution is triple of $(R_{s,7}, 3, 5)$ which has the highest value of G_r (50%), the lowest values of MOS_r (0.22%), and the lowest values of BW_{subopt} (45.34) with the suboptimal MOS equal to 4.47 while the best MOS equal to 4.48.

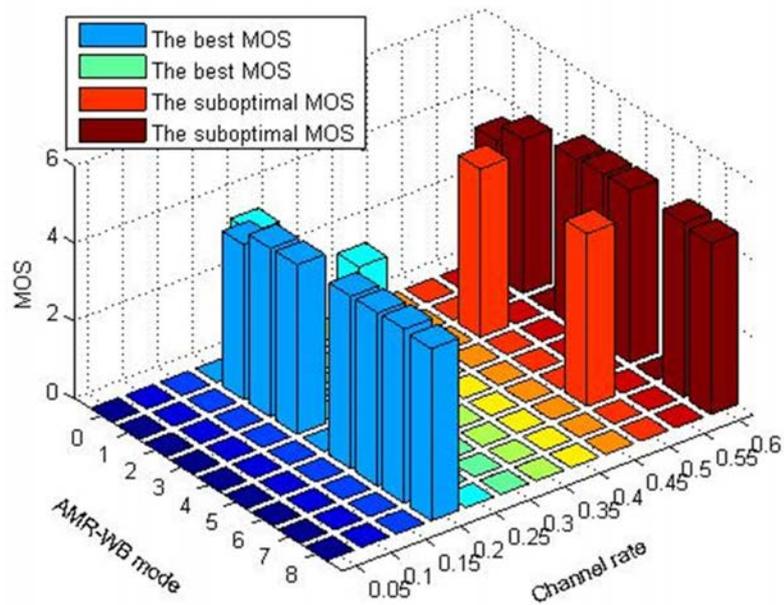


Figure 3.4: MOS vs. AMR-WB mode and channel code rate when p_r is randomly changed in range of 0..15%

Table 3.2: The detailed results of figure 3.4

Mode	k_{best}	n_{best}	MOS_{best}	k_{subopt}	n_{subopt}	MOS_{subopt}	$MOS_r(\%)$	$G_r(\%)$	$BW_{subopt}(kbps)$
0	1	4	3.51	3	5	3.47	1.14	33.33	17.26
1	1	5	3.98	3	5	3.95	0.75	50.00	21.10
2	1	5	4.33	3	4	4.32	0.23	50.00	33.11
3	1	5	4.35	3	5	4.33	0.46	50.00	30.32
4	1	4	4.40	3	5	4.39	0.23	50.00	33.05
5	1	5	4.43	3	5	4.42	0.23	33.33	37.15
6	1	4	4.45	2	4	4.44	0.22	50.00	47.85
7	1	5	4.48	3	5	4.47	0.22	50.00	45.34
8	1	5	4.42	3	5	4.40	0.45	50.00	46.70

In order to demonstrate that the proposed algorithm saving the bandwidth, we calculate the minimum required bandwidth for two cases of the best MOS and the suboptimal MOS. These values can be obtained from the equation (3.10) through two couples of (k_{best}, n_{best}) and (k_{subopt}, n_{subopt}) . Figure 3.5 shows the charts of the minimum required bandwidth of each mode of AMR-WB codec corresponding to two cases of the best MOS and the suboptimal MOS. It is clear that, when we choose according to the suboptimal condition, the system will save the bandwidth very significantly. This means the system will serve more mobile users at a moment.

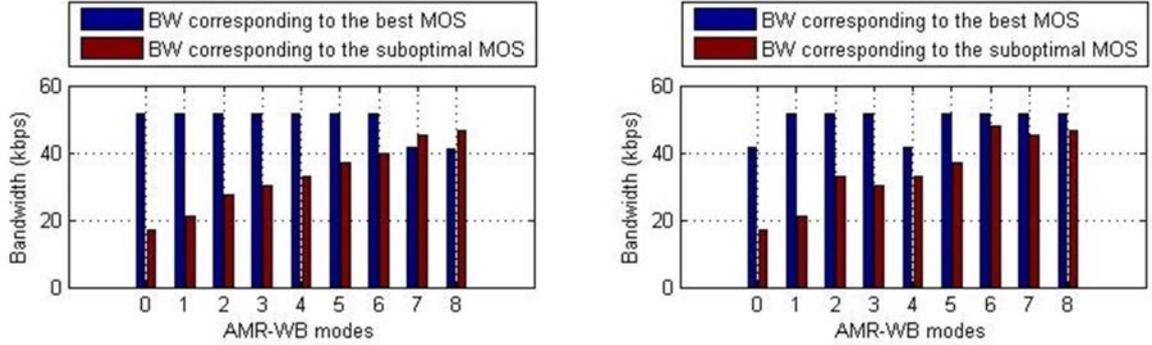


Figure 3.5: AMR-WB mode vs required bandwidth when p_r is fixed equal to 10% (left figure) and when p_r is randomly changed in range of 0..15% (right figure)

3.3 Enhancing channel decoding algorithm based on polynomial regression function

3.3.1 Decoding algorithms in Turbo codes

In this section, we briefly review some typical decoding algorithms in Turbo codes, we mainly focus on family of the MAP algorithms, consist of the optimal algorithm and some typical sub-optimal algorithms. Detailed description are represented in [57],[63].

3.3.1.1 The Log-MAP algorithm

The Log-MAP algorithm is a MAP algorithm which is implemented in logarithmic domain to reduce the computational complexity. It is an optimal algorithm for iterative decoding in Turbo decoder. The objective of the Log-MAP algorithm is to calculate the log-likelihood ratio [63]. The a priori information for the information bit u_k is computed as follows:

$$L(u_k) = \ln \left(\sum_{(s_{k-1}, s_k), u_k=+1} e^{\alpha_k^*(s_{k-1}) + \beta_{k+1}^*(s_k) + \gamma_k^*(s_{k-1}, s_k)} \right) - \ln \left(\sum_{(s_{k-1}, s_k), u_k=-1} e^{\alpha_k^*(s_{k-1}) + \beta_{k+1}^*(s_k) + \gamma_k^*(s_{k-1}, s_k)} \right) \quad (3.11)$$

In which:

- u_k is the information bits

- s_k and s_{k-1} denote the state at k th and $k-1$ th time instant, respectively.

In order to calculate the equation (3.11), we need to compute forward and backward recursive metrics, denoted as $\alpha_k(s_k)$ and $\beta_k(s_k)$. They are calculated as follows:

$$\alpha_k^*(S_k) = \ln(\alpha_k(S_k)) = \max_{s_{k-1} \in \sigma_{k-1}}^* (\gamma(S_{k-1}, S_k) + \alpha_{k-1}^*(S_{k-1})) \quad (3.12)$$

$$\beta_k^*(S_k) = \ln(\beta_k(S_k)) = \max_{s_{k+1} \in \sigma_{k+1}}^* (\gamma(S_k, S_{k+1}) + \beta_{k+1}^*(S_{k+1})) \quad (3.13)$$

In which:

- σ_{k-1} and σ_k are collection of all states at the moment $k-1$ and k , respectively
- γ is the branch metrics

Function $\max^*(.)$ in (3.12) and (3.13) is reckoned by using the Jacobian algorithm as follows:

$$\max^*(\delta_1, \delta_2) = \ln(e^{\delta_1} + e^{\delta_2}) = \max(\delta_1, \delta_2) + \ln(1 + e^{-|\delta_1 - \delta_2|}) \quad (3.14)$$

Where $\ln(1 + e^{-|\delta_1 - \delta_2|})$ is a correction function which corrects the error caused by the max approximation and makes the optimization for the Log-MAP algorithm. By replacing (3.12) and (3.13) into (3.14), we attain:

$$L(u_k) = \max[\beta_k^*(S_k) + \gamma(S_{k-1}, S_k) + \alpha_{k-1}^*(S_{k-1})] - \max[\beta_k^*(S_k) + \gamma(S_k, S_{k-1}) + \alpha_{k-1}^*(S_{k-1})] \quad (3.15)$$

3.3.1.2 The Log-MAP-based algorithms

The sub-optimal algorithms are based on the Log-MAP algorithm by replacing the correction function by an approximated function. Some typical sub-optimal algorithms consist of Max-Log-MAP [57], Constant Log-MAP [61], Linear Log-MAP [62] and Non-linear Log-MAP [63]. These algorithms try to reach the performance close to the Log-MAP algorithm with an acceptable complexity, obviously, they have the lower computational complexity than the Log-MAP algorithm. The approximated correction functions of these sub-optimal algorithms are expressed in following subsections:

3.3.1.2.1 The Max-Log-MAP algorithm The Max-Log-MAP algorithm has the least complexity because it omits the correction function. Hence, it is the simplest algorithm to implement but has the worst performance. The performance for the Max-Log-MAP algorithm gives up to a 10% performance drop [57] when compared to the Log-MAP algorithm. With the Max-Log-MAP algorithm, the correction function $f(x) = \ln(1 + e^{-x})$ with $x = |\delta_1 - \delta_2|$ is worked out as follows:

$$\ln(1 + e^{-|\delta_1 - \delta_2|}) \approx 0 \quad (3.16)$$

3.3.1.2.2 The Constant Log-MAP algorithm This algorithm is proposed by [60], [61], the correction function is approximated with the following principle:

$$\ln(1 + e^{-|\delta_1 - \delta_2|}) \approx \begin{cases} \frac{3}{8}, & \text{if } |\delta_1 - \delta_2| < 2 \\ 0, & \text{otherwise} \end{cases} \quad (3.17)$$

The Constant Log-MAP algorithm has a simple execution in hardware but with swap in performance.

3.3.1.2.3 The Linear Log-MAP algorithm In [62], the authors use the MacLaurin series to calculate the linear approximation for the correction function. It is observed that the correction function is effective when $f(x)$ is around zero. Thus, the MacLaurin series can be applied to approximate the correction function about zero. The authors propose the approximation for the correction term is given as:

$$\ln(1 + e^{-|\delta_1 - \delta_2|}) \approx \max(0, \ln 2 - \frac{1}{2} |\delta_1 - \delta_2|) \quad (3.18)$$

3.3.1.3 The Non-linear Log-MAP algorithm

This algorithm is proposed by the authors in [59]. It is the non-linear approximation for the correction term. It is inspired by observing the curve of the exact correction terms of Constant Log-MAP and Linear Log-MAP algorithms. The correction function is approximated as follows:

$$\ln(1 + e^{-|\delta_1 - \delta_2|}) \approx \frac{\ln 2}{2^{|\delta_1 - \delta_2|}} \quad (3.19)$$

3.3.2 The proposed Log-MAP algorithm based on polynomial regression function

3.3.2.1 The proposed error correction function

Although the Log-MAP algorithm has the best performance, its correction function carried out in logarithmic domain would bring some undesirable issues. According to [63], saving the results of the $\ln(1 + e^{-|\delta_1 - \delta_2|})$ in a lookup table would involve a quantization error caused by truncation of the input of the lookup. Another problem with the Log-MAP algorithm is that there are many lookup tables for a wide range of operating signal-to-noise ratios (SNRs), this will increase the hardware cost. Moreover, reading data from logarithmic tables is a time consuming process. So the correction function in the Log-MAP algorithm need be replaced by another function which has approximated performance but has simpler complexity. Therefore, we put forward a novel function to replace it. In order to find the new correction function, we exploit the understanding of polynomial regression function. The correction function is deployed into a form of polynomial function as follows:

$$f(x) = \ln(1 + e^{-x}) \approx a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (3.20)$$

In order to determine factors of $a_0, a_1, a_2, \dots, a_n$ and the suitable value of n , we have to determine a dataset consists of points presenting the relationship between $f(x)$ and x . According to [57], the values of x should only range between 0 and 5 to obtain ideal approximation. Thus, with $f(x) = \ln(1 + e^{-x})$, for $x = 0$ to 5 with step 0.1, we obtain a

dataset of 51 values of $f(x)$, respectively. This work is performed in Microsoft Excel, and then, we use the the Scatter method to represent the distribution of points of $f(x)$ [137]. In order to determine the approximated polynomial regression function, we use the Trendline method [138]. This is a method widely utilized in statistical probability field. It permits building a regressive function exactly from a set of known points.

Microsoft Excel supports the minimum value of the degree of polynomial regression function (n) of 2 and the maximum value is 6. The accuracy of the approximation is verified through the goodness of fit test against the exact Log-MAP curve with parameter of R-squared [139]. After changing the value of n from 2 to 6, we obtain parameters of R-squared are 0.9828, 0.9996 and 1s, respectively. However, we see that if choose $n = 2$ then we do not obtain the performance very close to the Log-MAP algorithm, otherwise, if we choose $n = 4, 5$ or 6 then will increase the computational complexity. Hence, we propose the value of $n = 3$. For this reason, we attain the approximated function for logarithmic term in the Log-MAP algorithm presented in Figure 3.7 and is given as:

$$f(x) = -0.0098x^3 + 0.1164x^2 - 0.474x + 0.6855 \quad (3.21)$$

As mentioned above, according to [57], in order to attain the ideal approximation, just

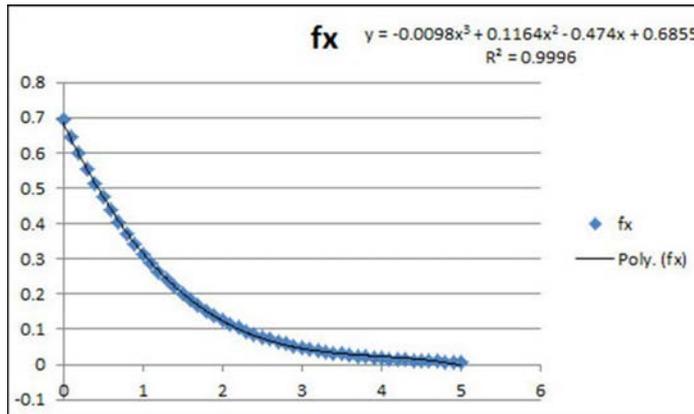


Figure 3.6: The proposed approximated function for the correction function in the Log-MAP algorithm

selecting the value of argument x between 0 and 5. Therefore, we can rewrite the proposed approximated correction function as follows:

$$\ln(1 + e^{-x}) \approx \begin{cases} -0.0098x^3 + 0.1164x^2 - 0.474x + 0.6855, & \text{if } x \leq 5 \\ 0, & \text{otherwise} \end{cases} \quad (3.22)$$

Through equation (3.22), we can calculate the logarithmic term in (3.14), and so, we attain the proposed Log-MAP algorithm.

In order to demonstrate that selecting the value of n is rational, we simulate the proposed Log-MAP algorithm with the different values of n . This is presented in section 3.3.3. The remaining approximated correction functions corresponding to $n = 2, 4, 5$ and 6 are presented as follows:

- With $n = 2$:

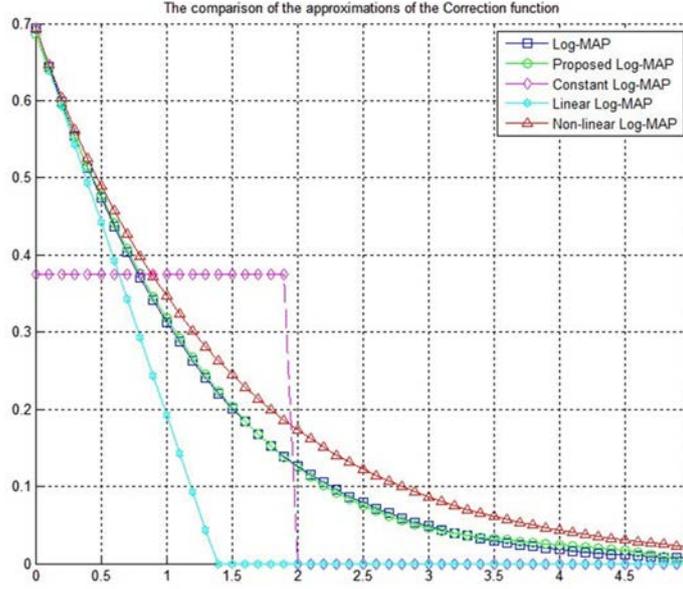


Figure 3.7: The comparison of the approximations of the correction function

$$\ln(1 + e^{-x}) \approx 0.0427x^2 - 0.3282x + 0.6277 \quad (3.23)$$

- With $n = 4$:

$$\ln(1 + e^{-x}) \approx 0.0012x^4 - 0.0216x^3 + 0.1539x^2 - 0.5148x + 0.6947 \quad (3.24)$$

- With $n = 5$:

$$\ln(1 + e^{-x}) \approx 7e^{-05}x^5 - 0.0003x^4 - 0.0178x^3 + 0.1469x^2 - 0.51x + 0.6941 \quad (3.25)$$

- With $n = 6$:

$$\ln(1 + e^{-x}) \approx -6e^{-05}x^6 + 0.001x^5 - 0.0049x^4 - 0.0041x^3 + 0.1302x^2 - 0.5021x + 0.6933 \quad (3.26)$$

The steps for determining the proposed error correction function is described in Algorithm 2.

3.3.2.2 The simulation model

In order to evaluate the performance of the proposed error correction function, we use the simulation model such as shown in Figure 3.8. The transport block which has dynamic size from MAC layer after is joined 24 parity bits by CRC will be fed to Turbo encoder. Next, this transport block is encoded and interleaved. And then, the encoded bit stream can be modulated by QPSK, 16QAM, or 64QAM depending on channel condition. The bit stream is then transmitted over a noisy channel such as AWGN or Rayleigh, etc. At the receiver, the process is inverse to the one at the sender. Specifically, the bit stream will be demodulated, de-interleaved and decoded.

Algorithm 2 Determining the approximated polynomial regression function

Put $f(x) = \ln(1 + e^{-x}), x \geq 0$

Step 1: For $x = 0$ to 5 step 0.1

- Calculating $f(x)$

Step 2: Determining point distribution of $f(x)$ in the Microsoft Excel using the Scatter method

Step 3: Determining the approximated polynomial regression function in the Microsoft Excel using the Trendline approach, and then determining R-squared ($R^2 \leq 1$)

Step 4: Changing the grade of the polynomial ($n = 2$ to 6) to determine n^* that the corresponding polynomial has the R-squared which is closest to 1

Step 5: The polynomial regression function which has the grade of n^* is the proposed error correction function

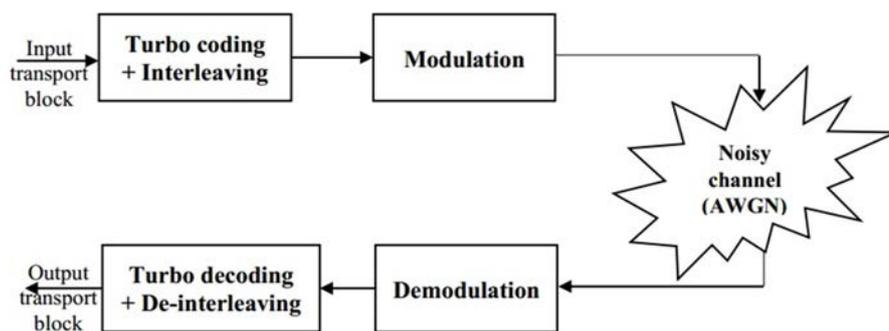


Figure 3.8: The simulation model

3.3.3 Simulation results and Performance evaluation

Figure 3.9 shows the simulated performance under AWGN channel for the proposed algorithm and the other Log-MAP-based algorithms, including Log-MAP, Max-Log-MAP, Constant Log-MAP, Linear Log-MAP and Non-linear Log-MAP algorithms. The Bit Error Rate (BER) performance is simulated in a rate-1/3, 8-states Turbo coded system with the transfer function for the constituent code $G = [1, \frac{1+D+D^3}{1+D^2+D^3}]$ [25]. The block size is $N = 1024$, using random interleaver and the maximum number of iterations for decoding was set to 5. As shown in Figure 3.9, the proposed algorithm has the BER performance closest to the Log-MAP algorithm and it outperforms the other sub-optimal algorithms.

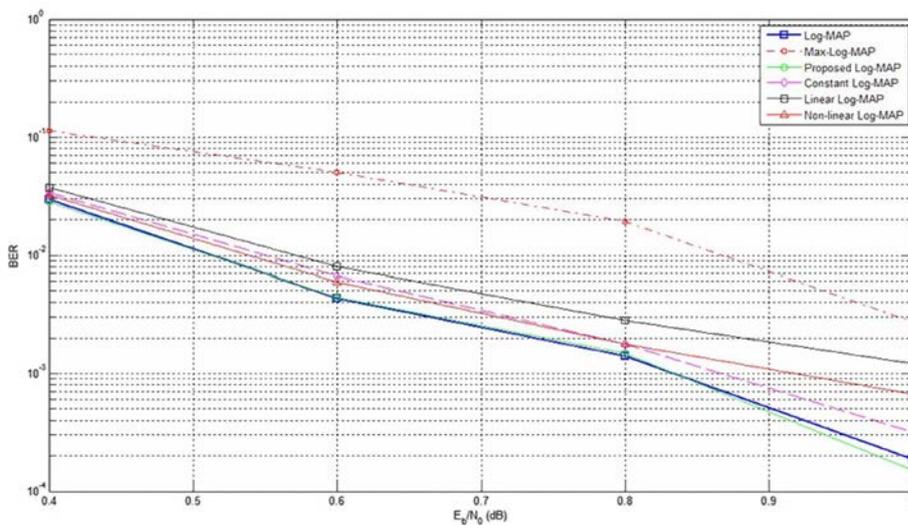


Figure 3.9: BER performance of the proposed Log-MAP algorithm and the other Log-MAP based algorithms with the block size $N = 1024$

Figure 3.10 shows the performance for the proposed algorithm, the Log-MAP algorithm and the other sub-optimal algorithms with the similar simulated parameters to figure 3.9, while its block size is $N = 512$. It is clear that although block size N is changed, the proposed algorithm still obtains the BER performance closest to the Log-MAP algorithm.

The simulated parameters utilized in Figure 3.11 is similar to Figure 3.9 but at low SNRs. The simulation result shows that the proposed algorithm has the BER performance nearly as identical as the Log-MAP algorithm, this is because the Log-MAP algorithm is sensitive to the SNR [140]. According to [139], at high SNRs, the performance of the Turbo code approaches the Max-Log-MAP algorithm and does not depend heavily on the correction function. Otherwise, at low SNRs, the decoder examines the a priori or extrinsic information from the previous decoder more. The distribution of argument x in the correction function $f(x)$ increased to areas close to zero where it is most effective. This is because the a priori is Gaussian distributed with increased number of iterations for decoding. The sensitivity of the Log-MAP algorithm to SNR is also more reported in encoders with more memory components [140].

The simulation results demonstrate that the proposed algorithm can be applied in real system in practice. It permits obtaining the closest performance with the computational complexity decreased in comparison with the Log-MAP algorithm because it eliminates the

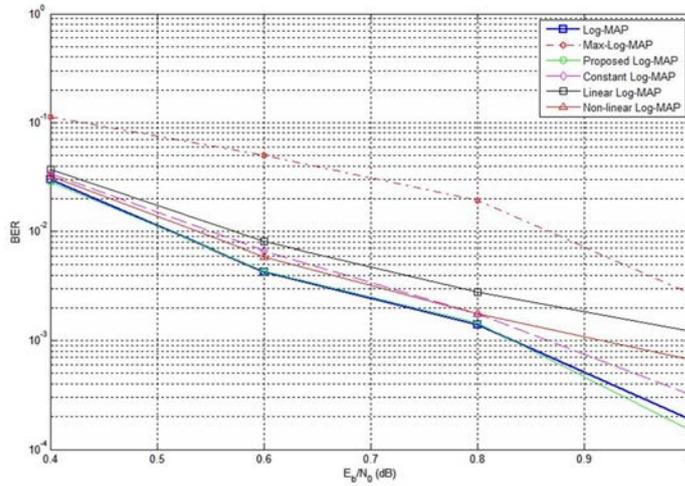


Figure 3.10: BER performance of the proposed Log-MAP algorithm and the other Log-MAP based algorithms with the block size $N = 512$

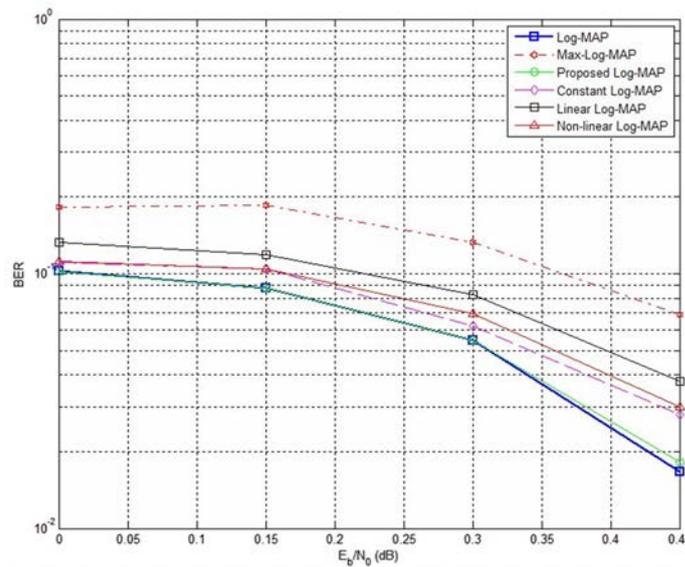


Figure 3.11: BER performance of the proposed Log-MAP algorithm and the other Log-MAP based algorithms at low SNRs with the block size $N = 1024$

logarithmic and exponential operations from the correction function.

Figure 3.12 shows simulation result of the different approximated functions corresponding to the various values of n . Because of R-squared parameters are all 1s with $n = 4, 5$ or 6 , so in this case, we only simulate BER performance with $n = 2, 3$ and 4 . The simulation result shows that with $n = 4$, BER performance is closest to the Log-MAP algorithm. In order to obtain good BER performance with an acceptable computational complexity, we see that choosing the value of $n = 3$ is suitable.

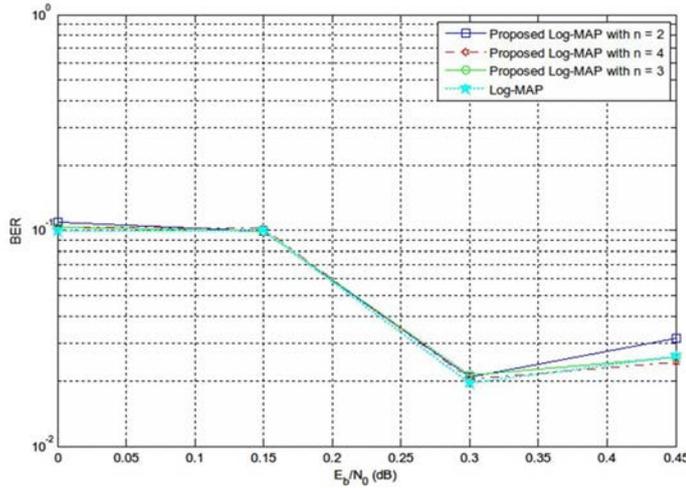


Figure 3.12: BER performance of the proposed Log-MAP algorithm with $n = 2, 3, 4$ and the Log-MAP algorithm at low SNRs with the block size $N = 1024$

3.4 Summary

In this Chapter, we first present an adaptive algorithm for dynamic joint source-channel code rate for voice traffic over LTE networks. The proposed algorithm permits choosing a suboptimal solution for minimizing the number of redundant bits generated by channel coding. The output of the algorithm is a pair of source and channel code rate resulting in minimizing the number of redundant bits generated by channel coding with an acceptable MOS reduction based on some constraints on the maximum allowed end-to-end delay, maximum permitted packet loss and minimum required bandwidth. The simulation results show that in the case of the fixed “raw” packet loss, AMR-WB codec mode 2 displays the suboptimal solution with channel code rate equal to $3/5$. Otherwise, when the varied “raw” packet loss, AMR-WB codec mode 7 displays the suboptimal solution with channel code rate equal to $3/5$. The simulation results also demonstrate that the proposed algorithm always finds out the suboptimal solution which meets some constraints on given QoS information for the tradeoff between speech quality and the number of redundant bits generated by channel coding. This means there is a slight reduction of MOS ($\leq 1\%$) when compared to the best MOS whereas we can obtain the percentage of decreased redundant bits generated by channel coding up to 50%. This will lead to saving the bandwidth and so that the system can serve more mobile users at the same time. Second, we present a novel approximated function replacing the one in the Log-MAP algorithm. This approximated function is deployed based on the understanding of the polynomial regression function. The proposed Log-MAP

based algorithm is a sub-optimal algorithm which achieves the performance closest to the Log-MAP algorithm with the much simpler computational complexity. Therefore, it can be easily implemented in hardware involving shift registers, multiplications, comparators and addition operations. The simulation results show that the proposed algorithm outperforms the other Log-MAP-based algorithms, particularly are superior to the Max-Log-MAP algorithm with slightly increased complexity.

Chapter 4

The proposed LTE Downlink packet schedulers for voice traffics based on user perception and VoIP-priority mode

Contents

4.1	General idea	68
4.2	LTE downlink scheduling	69
4.2.1	Simple scheduling diagram	69
4.2.2	The correlated scheduling algorithms	70
4.2.2.1	The FLS scheduler	70
4.2.2.2	The M-LWDF scheduler	71
4.2.2.3	The EXP/PF scheduler	71
4.3	The calculation of input parameters of the E-model and the WB E-model	72
4.3.1	An extension of the E-model	72
4.3.2	Calculation of the input parameters of the WB E-model	73
4.4	The proposed scheduling schemes	73
4.4.1	The proposed scheduling scheme based on user perception for narrowband audio services	73
4.4.2	The proposed scheduling scheme for wideband audio services	74
4.4.2.1	The proposed scheduling scheme based on user satisfaction	74
4.4.2.2	The proposed scheduling scheme based on user perception and the integration of VoIP-Priority mode	79
4.5	Simulation environment and Performance evaluation	80
4.5.1	Simulation environment	80
4.5.1.1	Traffic model	80
4.5.1.2	Simulation parameters	82
4.5.2	Performance measures	82
4.5.2.1	End-to-end delay	82
4.5.2.2	Packet Loss Rate	83
4.5.2.3	Cell throughput	83

4.5.2.4	Fairness index	84
4.5.2.5	Spectral efficiency	84
4.5.3	Performance evaluation	84
4.5.3.1	Performance evaluation of simulation scenario 1	84
4.5.3.2	Performance evaluation of simulation scenario 2	88
4.5.3.3	Performance evaluation of simulation scenario 3	92
4.5.4	Comparison of three proposed scheduling schemes	94
4.6	Summary	95

In this Chapter, we present the proposed scheduling schemes for voice traffics in LTE Downlink direction. The proposed schedulers called E-MQS, WE-MQS and WE-MQS are developed based on user perception by using the extended E-model and WB E-model for narrowband and wideband audio services, respectively. Besides, in order to enhance the priority for voice services, the VoIP-priority mode is used so-called WE-MQS-VoIP-Priority scheduler. The structure of this Chapter is organized as follows: Section 4.1 represents the problem formulation and justification of the proposals. In Section 4.2, we present simplified model of a packet scheduler and its operating principle as well as some related schedulers. The calculation of the input parameters of the E-model and the WB E-model is represented in Section 4.3. The proposed scheduling schemes for both narrowband and wideband audio users are described in Section 4.4. The simulation environment and performance evaluation are detailed in Section 4.5. In Section 4.6, we summarize the obtained results of the proposals in this Chapter.

4.1 General idea

As mentioned in Chapter 2, the scheduling strategies in the group of Channel- and QoS-Aware are very suitable for real-time services such as VoLTE. However, due to VoLTE is deployed in an All-IP network, and it is an emerge service, in the scheduling process, there should be the presence of many factors instead of only depending on HOL delay. We see that, the Maximum Queue Size (MQS) is a factor that should be present in the scheduling process. However, in the LTE-Sim, this factor is fixed equal to 0. This means the scheduling process does not take the MQS into account as an essential in the priority metric. However, in fact, the MQS should be a finite value because if the MQS value is infinite then the delay will increase and the congestion could be increased. Therefore, the MQS needs to be considered as an essential factor in the priority metric of scheduling algorithms. For this reason, we propose to complement the MQS into the scheduling process. Second, we propose to use user perception to make the scheduling decision. This idea is proceeded from Chapter 3 in which the WB E-model is used to determine user satisfaction. Hence, in this Chapter, we continue to apply that idea to the proposed schedulers. For the related work, there are several papers which mentioned it. Specifically, authors in [141], [142] proposed a new QoE-driven LTE Downlink Scheduling for VoIP Application that is based on QoE min (i.e. MOS score at least equals 3.5 for VoIP application) and they optimize number of user to access a cell. In [143], the authors proposed a cross-layer design scheme that jointly optimizes three different layers of wireless protocol stack, namely Application, MAC and Physical layers. The purpose of this paper is to maximize network resource utilization and user-perceived quality of service (QoE). These papers proposed a new LTE downlink

scheduler but not based on the E-model. The paper which is closest to our work described in [144]. In this paper, the authors proposed a new scheduling scheme for VoIP service in LTE networks by using the user satisfaction as a metric for their scheduler. The authors used the E-model to predict user perception via MOS score, and then this factor used in the metric for scheduling decision. However, in this paper, authors did not consider the impact of network jitter for E-model. In addition, in the priority metric of the scheduler, there is no the presence of the MQS, and the authors evaluated only for VoIP traffic. In our proposal, we develop the idea in [144] by extending the E-model and propose to consider the MQS as an essential and effective factor for the metric. We use the extended E-model to predict the MOS score and use this score as one of main factors in the metric.

Third, we extend the above proposed scheduling schemes for wideband audio services in LTE networks. As mentioned before, almost of our proposals are simulated in the LTE-Sim software. However, in this software, there is only G.729 codec supported. This means the LTE-Sim only supports to simulate narrowband audio services whereas VoLTE is a wideband audio service and VoLTE uses AMR-WB as its voice source codec. In order to overcome this limitation, we propose to complement AMR-WB into the LTE-Sim. This allows the LTE-Sim to simulate wideband audio services. And in order to predict user perception in this case, the WB E-model is adapted. Such that, for wideband audio services, the proposed scheduler still considers the MQS and user perception in its priority metric.

Last, since VoLTE is deployed in an All-IP network, there need to have a special priority for it. For this reason, in order to enhance VoIP users, we integrate the VoIP-Priority mode [145] into the proposed scheduler for modifying resource allocation method. The VoIP-Priority mode is only enhanced when there is VoIP user in the buffer. This allows scheduling VoIP users before any other user. In order to reduce the negative effects of VoIP-Priority mode on the system, the procedure of duration of VoIP-Priority mode is deployed. However, in this Chapter, we only integrate the VoIP-Priority mode into the proposed scheduler for wideband audio services. This work can be easily executed similarly to the case of narrowband audio services.

4.2 LTE downlink scheduling

4.2.1 Simple scheduling diagram

In this section, we present simplified model and operating principle of a packet scheduler in LTE network. According to [66], a simplified packet scheduler can be described such as in Figure 4.1.

The scheduling process in Figure 4.1 is described as follows:

- Each UE decodes the reference signals, calculates the CQI and sends it back to the eNB
- The eNB utilizes the CQI information for the allocation decisions and fills up a RB “allocation mask”
- The AMC module selects the best MCS that should be used for the data transmission by scheduled users
- The information about these users, the allocated RBs, and the selected MCS are sent to the UEs on the PDCCH

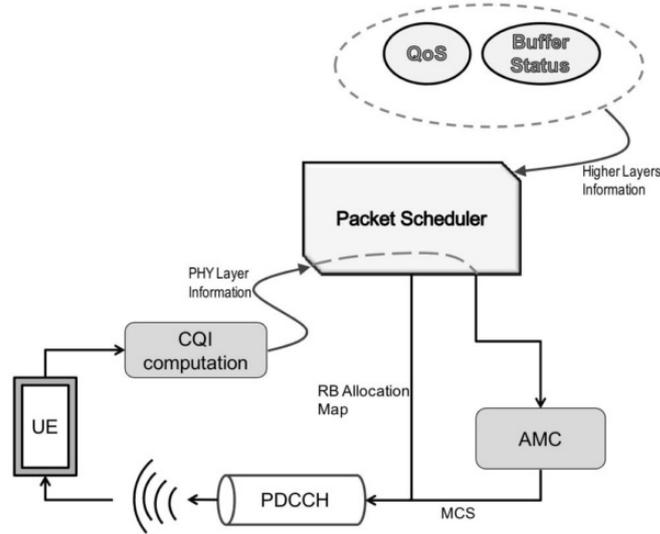


Figure 4.1: Simplified model of a packet scheduler

- Each UE reads the PDCCH payload and in case it has been scheduled, accesses to the proper PDSCH payload.

4.2.2 The correlated scheduling algorithms

The purpose of scheduling algorithm in LTE is to maximum system performance [146]. The chosen UE is suitable for scheduling decisions. In order to make scheduling decisions, there is much information included, such as the number of sessions, their served rates, link states, and the statuses of session queues [146]. The eNB offers a scheduling decision based on the CQI which is sent back to the eNB from the UE. The CQI is then exploited by the scheduler link adaptation module to select an UE with the most suitable modulation scheme and coding rate at the PHY layer with the objective of the spectral efficiency maximization.

We assume that the metric assigned to the stream i on j -th sub-channel is noted by $w_{i,j}$. In order to obtain the metric, the scheduler usually needs to know the average transmission rate (\bar{R}_i) of flow i , and the flow rate available to the UE on the j -th sub-channel. In particular, at each TTI, the estimate \bar{R}_i is given by [17]:

$$\bar{R}_i(k) = 0.8 \times \bar{R}_i(k-1) + 0.2 \times \bar{r}_i(k) \quad (4.1)$$

Where: $\bar{R}_i(k-1)$: The average transmission data rate estimating at the $(k-1)$ -th TTI.
 $\bar{r}_i(k)$: The rate allocated to i -th flow during the k -th TTI.

In the following subsections, we will describe the metric of several well-known scheduling algorithms which are related to our proposed scheduling schemes including: FLS, M-LWDF, and EXP/PF schedulers. We select these scheduling because they perform well and are suitable for real-time services.

4.2.2.1 The FLS scheduler

FLS is a two-level scheduling algorithm which are called upper level and lower level. These levels are distinct and communicated with each other to dynamically allocate the RBs to the

users. At the upper level, a resource allocation method (called FLS) which uses a Discrete-Time (D-T) linear control theory is performed. FLS defines the amount of data that each real-time source should transmit within a single frame to meet its delay constraint. At the lower level, the algorithm uses Proportional Fair (PF) method to allocate RBs to the users at each TTI with considering the bandwidth requirements of FLS to ensure a good level of fairness among multimedia flows. Also at this layer, the scheduler determines number of TTIs/RBs via that each Real-time source will send its packets. In order to calculate the amount of data transmitted, the FLS scheduler uses the following formula:

$$V_i(k) = h_i(k) * q_i(k) \quad (4.2)$$

In which: $V_i(k)$ is the amount of the data transmitted by the flow i in LTE frame k , $*$ operator is the discrete time convolution, $q_i(k)$ is the queue level. It can be said that, $V_i(k)$ is computed by filtering the signal $q_i(k)$ via a time-invariant linear filter with pulse response $h_i(k)$.

4.2.2.2 The M-LWDF scheduler

M-LWDF scheduling algorithm is used to support multiple real-time services in CDMA-HDR systems [147]. For each real-time flow, by considering the maximum time τ_i , the probability is defined as the maximum probability δ_i which is the time of the first packet of the queue exceeds the fixed maximum time $D_{HOL,i}$. In this algorithm, the metrics for real-time and non real-time services are different. In order to offer priority to real-time flows, the metric was given as follows:

$$w_{i,j} = \alpha_i \times D_{HOL,i} \times \frac{r_{i,j}}{\bar{R}_i} \quad (4.3)$$

In which:

- $r_{i,j}$: The rate assigned to i -th flow during the k -th TTI
- \bar{R}_i : The average transmission data rate estimating
- α_i : A factor and is given by: $\alpha_i = -\frac{\log(\delta_i)}{\tau_i}$

4.2.2.3 The EXP/PF scheduler

EXP/PF is a scheduling algorithm which supports multimedia applications in an adaptive modulation and coding and time division multiplexing (AMC/TDM) system [148]. The main purpose of this scheduling algorithm is to enhance the priority for the real-time flows by adding the average fixed maximum time of all active real-time flows. For the real-time services, they receive the increased priorities when their HOL packet delays are approaching the delay deadline. The metric of EXP/PF is calculated as follows:

$$w_{i,j} = \exp\left(\frac{\alpha_i \times D_{HOL,i} - X}{1 + \sqrt{X}}\right) \times \frac{r_{i,j}}{\bar{R}_i} \quad (4.4)$$

Where X is given by: $X = \frac{1}{N_{rt}} \sum \alpha_i \times D_{HOL,i}$, with N_{rt} is the number of active real-time flows in the downlink direction.

The remaining parameters are similar to the descriptions above.

4.3 The calculation of input parameters of the E-model and the WB E-model

4.3.1 An extension of the E-model

We see that, when a voice packet is transmitted over an IP network, it is affected by many network impairments such as PLR, delay, jitter, etc. In the E-model, there is no presence of network jitter. In order to improve user satisfaction, we propose to add the I_j factor to the E-model. In this case, the E-model so-called the extended/enhanced E-model. It's obvious, when add the I_j to the E-model, user perception of the extended E-model will be lower than E-model, and the E-model can be rewritten as the following formula:

$$R = 93.2 - I_d - I_{ef} - I_j \quad (4.5)$$

Equation (4.5) shows that the R-factor depends on end-to-end delay (I_d), total loss probability (I_{ef}), and jitter buffer (I_j). Hence, in order to compute the R-factor, we must count these factors. The I_d is a factor which is affected by end-to-end delay and is calculated according to the equation 3.3. In that equation, D_{e2e} represents end-to-end delay (or mouth-to-ear delay) of speech packet. This can be obtained via some functions of LTE-Sim software. The I_{ef} is determined according to packet loss. In order to compute this factor, we use the equation in [144] as follows:

$$I_{ef} = \lambda_1 + \lambda_2 \times \ln(1 + \lambda_3 \times e_l) \quad (4.6)$$

In which: The λ_1 represents the voice quality impairment factor caused by the encoder, λ_2 and λ_3 represent the effect of loss on voice quality for a given codec. Such that, these factors depend on the voice codec used. In this study, we use LTE-Sim [17] to simulate. This software supports only G.729 codec, thus, for this codec, the factors above has values as follows: $\lambda_1 = 11, \lambda_2 = 40, \lambda_3 = 10$. While e_l is the total loss probability (consisting of network and buffer layout) which has the value in range of 0 to 1. This factor is also obtained from some functions of LTE-Sim software.

The I_j represents the impacts of network jitter to voice quality. It also depends on the voice source codec. In this thesis, we use the method proposed in [100] as follows:

$$I_j = C_1 \times H^2 + C_2 \times H + C_3 + C_4 \times \exp\left(\frac{-T}{K}\right) \quad (4.7)$$

In which:

- C_1, C_2, C_3, C_4 are coefficients. These factors depend on the voice codec, for the G.729 codec, these factors have the values as follows: $C_1 = -15.5, C_2 = 33.5, C_3 = 4.4, C_4 = 13.6$
- K is time instant, $K = 30$
- T is the fixed buffer size of the voice codec. For the G.729 codec, the packet size is 20ms, thus, normally $T = \alpha \times 20$ where $\alpha = 2, 3, 4, 5, 6, etc.$
- H is a factor of Pareto distribution and in range of 0.55 to 0.9. According to [100], the MOS slightly drops when H increases and it does not affect significantly on MOS score, thus, in this study, we select $H = 0.6$ for the simulation.

The final expression of the R-factor when utilizing the G.729 codec is described in Equation (4.8).

$$R = 63.28 - [0.024 \times d + 0.11 \times (d - 177.3) \times H(d - 177.3)] - 40 \times \ln(1 + 10 \times e_l) - 13.6 \times \exp\left(\frac{-4}{3}\right) \quad (4.8)$$

The R-factor is then mapped to the MOS. MOS reflects user perception. It is obvious, the higher MOS, the higher user satisfaction and inverse. However, determining exactly user satisfaction is still a big challenge of mobile network operators.

4.3.2 Calculation of the input parameters of the WB E-model

In order to compute the R_{wb} factor, we have to count the values of $I_{d,wb}$ and $I_{e,eff,wb}$ factors. The $I_{d,wb}$ factor is determined similarly to the case of narrowband audio, i.e. it is also calculated according to the following equation 3.3.

The $I_{e,eff,wb}$ is determined according to packet loss. $I_{e,eff,wb}$ is determined as the equation 3.6.

In that equation, $I_{e,wb}$ is the respective impairment factor without any packet loss. P_{pl} : Packet loss rate. It can be also obtained after finishing the simulation scenario in the LTE-Sim software. B_{pl} : A codec-specific factor which characterizes its robustness against packet loss. The values of $I_{e,wb}$, B_{pl} is represented in Table 4.1 [33].

Table 4.1: Values of $I_{e,wb}$ and B_{pl} according to AMR-WB modes

AMR-WB mode	Bitrate (bps)	$I_{e,wb}$	B_{pl}
0	6.6	39	12.8
1	8.85	25	13.5
2	12.65	11	13
3	14.25	10	14.1
4	15.85	7	13.1
5	18.25	5	12.5
6	19.85	4	12.3
7	23.05	1	13
8	23.85	6	12.2

After calculating the R_{wb} -factor, it is converted to the R-factor ($R = R_{wb}/1.29$), and then is mapped to the MOS via equation (1.3). MOS score performs user perception.

4.4 The proposed scheduling schemes

4.4.1 The proposed scheduling scheme based on user perception for narrowband audio services

In this proposed scheduling scheme, we consider the characteristics of the real-time services such as VoIP. These services are sensitive to packet loss and delay, thus, scheduling process should consider various factors. In the related scheduling algorithms above, the authors almost focused on Head of Line packet delay, virtual token length besides other factors such as α_i , $r_{i,j}$ and \bar{R}_i . MOS is a parameter which represents user perception, thus, it

should appear in the metric of scheduling algorithms. The higher MOS is, the higher user satisfaction is. MOS needs to be automatically calculated at the receiver and is sent to the eNodeB via feedback technique. For the MQS, according to our knowledge, there are no articles which mention about it. We think that, this factor has strong effects on the system performance. In the LTE-Sim [17], this factor is fixed equal to 0. This means the MQS is infinite. Hence the MQS is not considered in the scheduling process. However, in fact, the MQS should be finite. If the MQS value is infinite then the delay will increase and the congestion could be thus increased. Therefore, the MQS should be considered as a necessary factor in the metric priority of scheduling algorithms.

The main idea of the proposed scheduling algorithm is the consideration of user satisfaction (MOS) and the MQS factor (called also $Q_{i,max}$) included into the metric of the scheduling algorithm. This means the higher priority metric ($w_{i,j}$) value, the higher priority for the corresponding UE. The fixed maximum time $D_{HOL,i}$ and the maximum probability δ_i are included in the Equation (4.3) to calculate the factors of I_d, I_{ef} , respectively. The metric in our scheduling scheme for voice services is defined as follows:

$$w_{i,j} = \frac{MOS_i \times (Q_{i,max} - Q_i)}{\tau_i} \times \frac{r_{i,j}}{\bar{R}_i} \quad (4.9)$$

Where:

- $Q_i, \tau_i, r_{i,j}$ and \bar{R}_i have the same significances in the previous formulas.
- $Q_{i,max}$: The MQS of the user i . This value can be obtained in bytes via some functions in LTE-Sim [17].
- MOS_i : User perception, is computed using the extended E-model

For video and non real-time services, we propose to use the method in the M-LWDF scheduler. The $w_{i,j}$ is a priority matrix for each RB_j is assigned to UE_i . It is calculated based on the MOS, the remaining queue size ($Q_{i,max} - Q_i$), the maximum time τ_i and the channel condition. MOS is computed at the receiver and is feedbacked to the eNodeB in order to make the scheduling decision of UE. MOS included in the metric will fully exploit user perception.

The procedures of the proposed scheduling scheme and the scheduling metric ($w_{i,j}$) are described as in Algorithm 3. The implemented model of the proposed scheduling scheme is represented in this case such as in Figure 4.2.

4.4.2 The proposed scheduling scheme for wideband audio services

4.4.2.1 The proposed scheduling scheme based on user satisfaction

This scheduling scheme is the extension of the E-MQS scheduler for wideband audio service so called WE-MQS scheduler. We see that, the lack of LTE-Sim software is that it supports only G.729 codec while VoLTE uses AMR-WB codec to get high user satisfaction. G.729 has only an unique mode which has the bitrate of 8 kbps and the packet size of 32 bytes generated in each 20 ms while AMR-WB has 9 modes. In fact, modes of AMR-WB are changed according to channel condition (i.e. C/I ratio). So user perception is calculated at scheduled instant. In the proposed scheduling scheme, we propose to complement AMR-WB codec into LTE-Sim software by reconfiguring some parameters and modifying essential source codes. With the presence of AMR-WB, we can simulate VoLTE traffic more easily.

Algorithm 3 Proposed scheduling scheme for narrowband audio users in LTE Downlink

- 1: **procedure** SCHEDULING AND RESOURCE ALLOCATION
 - 2: **Step 1:** Identify real-time or any other traffic is queuing in the buffer
 - 3: **Step 2:** Determine the user metrics ($Q_{i,max}$, Q_i , MOS_i) and other parameters. Calculate the scheduling metric ($w_{i,j}$) according to the Equation 4.9
 - 4: **Step 3:** Find the user having the metric that is met the scheduling criteria
 - 5: **Step 4:** Count the set of available RBs to allocate which RBs to the chosen users
 - 6: **Step 5:** Assign the resource block R^* to the user U^* which has the satisfied metric
 - 7: **Step 6:** Schedule the user U^* first
 - 8: **Step 7:** Delete the user U^* and resource block R^* from their corresponding lists
 - 9: **Step 8:** Repeat the steps from 1 to 7 until all users are scheduled.
- 1: **procedure** w_{ij} CALCULATION FOR NARROWBAND AUDIO USERS
 - 2: **Step 1:** Calculate $Q_{i,max} - Q_i$, τ_i , $r_{i,j}$, and \bar{R}_i
 - 3: **Step 2:** Calculate R-factor
 - 4: Calculate I_d based on Equation (3.3)
 - 5: Calculate $I_{e,eff}$ based on the Formula (4.6)
 - 6: Calculate I_j based on the Formula (4.7)
 - 7: Calculate R-factor based on Equation (4.8)
 - 8: **Step 3:** Calculate MOS score using Equation (1.3)
 - 9: **Step 4:** Calculate $w_{i,j}$ based on Formula (4.9)

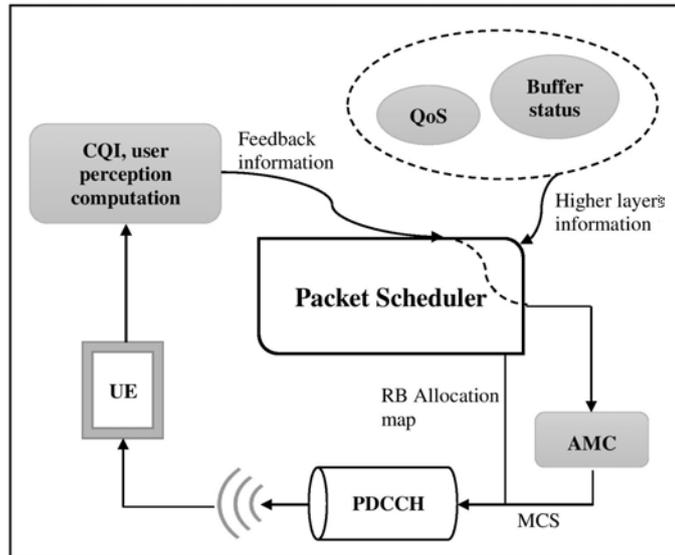


Figure 4.2: Implemented model of the proposed scheduling scheme

The main idea of the proposed scheduling algorithm is the consideration of user satisfaction (MOS) and the MQS factor (called also $Q_{i,max}$) included into the metric of the scheduling algorithm. This means the higher priority metric (w_{ij})value, the higher priority for the UE. The fixed maximum time $D_{HOL,i}$ and the maximum probability δ_i are included in the Equations (3.3) and (3.6) to calculate the factors of $I_{d,wb}$, $I_{e,wb}$, respectively. The metric in the proposed scheduling scheme for voice users is defined as follows:

$$w_{i,j} = \frac{MOS_i \times (Q_{i,max} - Q_i)}{\tau_i} \times \frac{r_{i,j}}{\bar{R}_i} \quad (4.10)$$

Where:

- Q_i , τ_i , $r_{i,j}$ and \bar{R}_i have the same significances in the previous formulas.
- $Q_{i,max}$: The MQS of the user i . This value can be obtained in bytes via some functions in LTE-Sim [17].
- MOS_i : User perception, is calculated using the WB E-model

For video and non real-time services, we propose to use the metric of the M-LWDF scheduler. The $w_{i,j}$ is a priority matrix for each RB_j is assigned to UE_i . It is calculated based on the MOS, the remaining queue size ($Q_{i,max} - Q_i$), the maximum time τ_i and the channel condition. MOS is computed at the receiver and sent back to the eNB to make the scheduling decision of UE. MOS included in the metric will fully exploit user perception.

In fact, the AMR-WB mode is dynamically determined and optimized at the AMR-WB encoder according to channel quality using rate adaptation control algorithm detailed in [16]. The limitation of LTE-Sim is that it supports only G.729 codec for VoIP. Therefore, the proposed scheduler can not get the mode chosen from AMR-WB encoder at Application layer. In order to overcome this problem, we propose a procedure which allows to choose AMR-WB mode from C/I ratio that is available in LTE-Sim. With the proposed procedure, the proposed scheduler can choose dynamically source mode according to channel quality. The threshold values of C/I ratio are chosen according to [149] and [150]. The procedures for updating AMR-WB mode, allocating resource, and calculating the metric of the proposed scheduler are described in Algorithm 4 and Algorithm 5.

The procedure of Update AMR-WB packet size is used to update packet size according to channel condition and is used for all schedulers while the procedure of w_{ij} calculation is used only in the proposed scheduler for calculating the metric.

The implemented model of the proposed scheduling scheme in this case is represented such as in Figure 4.3. This Figure represents the main RRM modules that interact with the downlink packet scheduler. The entire process of a packet scheduler can be divided into a sequence of tasks which are repeated in each TTI [66]. When voice packets arrive at the buffer in the eNB, they are given a time stamp and are queued for transmission. For each packet in the queue, the HOL delay and the QL are estimated. If the HOL packet delay or QL exceeds a specified threshold for the flow, then that packet will be discarded. The packet scheduler decides which users will be served according to a scheduling algorithm based on the metric of corresponding algorithm. In order to have a suitable scheduling scheme, we have to have the trade-off between channel state, QoS requirements and queue status. In the proposed scheduling scheme, when making the decisions, the scheduler takes into account the instantaneous or average channel conditions, HOL packet delay, status of receiving buffer such as QL and MQS or type of service being used, especially user perception.

Algorithm 4 Updating AMR-WB packet size according to the channel conditions

```

1: procedure UPDATING AMR-WB PACKET SIZE FOR THE Config-WB-Code 0
2: Step 1: Calculate  $C/I$  ratio: available in LTE-Sim
3: Step 2: Update packet size of AMR-WB codec
4:   if calculated  $C/I \leq 6.5$  then
5:      $mode = 0$ 
6:     Update  $Packet\_size = 30$ 
7:   else if calculated  $C/I \leq 12.5$  then
8:      $mode = 1$ 
9:     Update  $Packet\_size = 36$ 
10:  else
11:     $mode = 2$ 
12:    Update  $Packet\_size = 45$ 

1: procedure UPDATING AMR-WB PACKET SIZE FOR THE Config-WB-Code 2
2: Step 1: Calculate  $C/I$  ratio: available in LTE-Sim
3: Step 2: Update packet size of AMR-WB codec
4:   if calculated  $C/I \leq 6.5$  then
5:      $mode = 0$ 
6:     Update  $Packet\_size = 30$ 
7:   else if calculated  $C/I \leq 12.5$  then
8:      $mode = 1$ 
9:     Update  $Packet\_size = 36$ 
10:  else if calculated  $C/I \leq 18.5$  then
11:     $mode = 2$ 
12:    Update  $Packet\_size = 45$ 
13:  else
14:     $mode = 4$ 
15:    Update  $Packet\_size = 53$ 

1: procedure UPDATING AMR-WB PACKET SIZE FOR THE Config-WB-Code 4
2: Step 1: Calculate  $C/I$  ratio: available in LTE-Sim
3: Step 2: Update packet size of AMR-WB codec
4:   if calculated  $C/I \leq 6.5$  then
5:      $mode = 0$ 
6:     Update  $Packet\_size = 30$ 
7:   else if calculated  $C/I \leq 12.5$  then
8:      $mode = 1$ 
9:     Update  $Packet\_size = 36$ 
10:  else if calculated  $C/I \leq 18.5$  then
11:     $mode = 2$ 
12:    Update  $Packet\_size = 45$ 
13:  else
14:     $mode = 8$ 
15:    Update  $Packet\_size = 73$ 

```

Algorithm 5 Scheduling scheme for wideband audio users in LTE Downlink

- 1: **procedure** SCHEDULING AND RESOURCE ALLOCATION
 - 2: **Step 1:** Identify real-time or any other traffic is queuing in the buffer
 - 3: **Step 2:** Determine the user metrics ($Q_{i,max}$, Q_i , MOS_i) and other parameters. Calculate the scheduling metric ($w_{i,j}$) according to the Equation 4.10
 - 4: **Step 3:** Find the user having the metric that is met the scheduling criteria
 - 5: **Step 4:** Count the set of available RBs to allocate which RBs to the chosen users
 - 6: **Step 5:** Assign the resource block RB* to the user equipment UE* which has the satisfied metric
 - 7: **Step 6:** Schedule the user equipment UE* first
 - 8: **Step 7:** Delete the user UE* and resource block RB* from their corresponding lists
 - 9: **Step 8:** Repeat the steps from 1 to 7 until all users are scheduled.
-
- 1: **procedure** w_{ij} CALCULATION FOR WIDEBAND AUDIO USERS
 - 2: **Step 1:** Calculate $Q_{i,max} - Q_i$, τ_i , $r_{i,j}$, and \bar{R}_i
 - 3: **Step 2:** Calculate R-factor
 - 4: Calculate $I_{d,wb}$ based on Equation (3.3)
 - 5: Calculate $I_{e,eff,wb}$ based on the chosen *mode* using Formula (3.6).
 - 6: Calculate R_{wb} -factor based on Equation (1.5)
 - 7: **Step 3:** Calculate MOS score using Equation (1.3)
 - 8: **Step 4:** Calculate $w_{i,j}$ based on Formula (4.10)
-

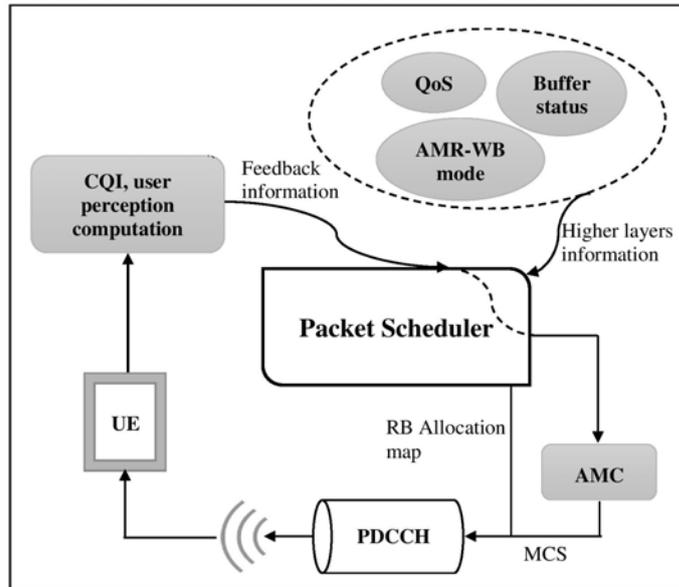


Figure 4.3: Implemented model of the proposed scheduling scheme

4.4.2.2 The proposed scheduling scheme based on user perception and the integration of VoIP-Priority mode

4.4.2.2.1 Dynamic scheduling Since LTE is an All-IP network, packets are usually scheduled using L1/L2 scheduling. Voice packets can also be scheduled using the same method. From viewpoint of radio resource usage, this scheme is most flexible because it permits scheduling the packets when they are needed. However, the limitation of this scheme is that it needs a large amount of signaling. For the fully dynamic scheduling, each UE sends resource request in uplink for each voice packet, the eNB allocates uplink resource for every transmission and retransmission separately by the L1/L2 control signaling. For the dynamic scheduling, the voice packet scheduling enjoys the full diversity of channel in both time and frequency domain with the expense of large control signaling. The dynamic scheduler is used as a base scheduler in LTE network.

4.4.2.2.2 Persistent scheduling In persistent scheduling scheme, the scheduling does not depend on channel condition. It is similar to the resource allocation of a CS network that is made for each call and the allocation is maintained during the duration of the call. This method generates a large amount of control signal. RRC signaling or several types of enhanced L1/LE signaling is used to allocate a sequence of TTI-RU chunks (time/frequency resource) as well as a fixed modulation scheme for a voice user. The allocated resource also consists of resources required for HARQ retransmission. The main issue of this method is the wastage of resource because in wireless environment, it is very difficult to determine exactly the number of retransmission as well as resource needed for a voice packet. This is due to channel condition, interference, noise, etc. So that, it is very hard to allocate a fixed number of resource to each user and it may cause negative effects such as very low usage of capacity.

4.4.2.2.3 Semi-persistent scheduling This is a hybrid method of Dynamic scheduling and Persistent scheduling [151]. VoIP packets uses a small quantity of control signaling to determine the channel quality after every fixed interval and persistently schedules the VoIP packets. This supports best to VoIP traffics due to its controlled dynamic nature and utilization of small control signaling.

4.4.2.2.4 VoIP-Priority mode VoIP-Priority mode was proposed by Sunggu Choi et al. in [145], it allocates RBs for VoIP calls before any other traffic. The limitation of this method is that when VoIP calls density is high, other traffics are not allocated. However, the authors already solved this problem by complementing a duration procedure. It is controlled dynamically to adjust consecutive TTIs according to total drop ratio of the packets measured at the eNodeB. In this mode, RBs are allocated based on Channel Adaptive Fair Queuing [152]. The metric of the scheduler in VoIP priority mode is determined based on QL and SINR and is calculated as the following equation:

$$w_{i,j} = Q_l(i) \times \gamma(i) \quad (4.11)$$

Where (Q_l) is queue length and (γ) is SINR of active VoIP call i . Formula (4.11) indicates that UE which has the longer queue length and the better channel quality then will have the higher priority. It can be said that, the VoIP priority mode is useful when the density of VoIP calls is high. However, the downside is that other traffic to be starved.

Therefore, the duration is deployed. It is controlled dynamically and depends on total of the drop ratio of the packets measured at the eNodeB. A predefined minimum and maximum drop ratio is utilized to adjust VoIP priority duration. Specifically, the maximum count of VoIP priority duration is increased when the drop ratio is less than the minimum threshold, and the maximum count of VoIP priority duration is decreased when the drop ratio exceeds the maximum threshold because in this case there is not enough resources to allocate. If the drop ratio is in range of the min to the max threshold then the duration is kept a constant. For more details of the duration, refer to [145].

4.4.2.2.5 The proposed scheduler based on user perception and the VoIP-Priority mode In this scheduling scheme, we integrate the VoIP-Priority mode into the proposed scheduling scheme such as represented in Figure 4.4. The algorithm priority mode is only enabled when there is VoIP user in the buffer. So the proposed scheduler has the method of resource allocation that is similar to the VoIP-Priority mode but has the different priority metric. This is mentioned in the previous subsections. In order to decrease the negative effects on other traffics, the duration of VoIP-Priority mode is deployed. This procedure allows determining maximum allowable count of consecutive VoIP priority TTIs. The maximum count increases or decreases depending on overall VoIP packet drop ratio which is measured at the eNodeB. By using preset minimum and maximum thresholds for the VoIP packet drop ratio, if the drop ratio is below the minimum threshold, the maximum consecutive count decreases by one because it implies that either there are enough resources or VoIP calls are given sufficient number of PRBs. Otherwise, if the drop ratio is above the maximum threshold, the maximum count is doubled because it means that the PRBs allocated to VoIP calls are not enough. Finally, if the drop rate stays between the minimum and maximum threshold, the current maximum count is kept. This adjustment algorithm is detailed in [145].

The implemented model of the proposed scheduling scheme in this case is represented such as in Figure 4.5.

4.5 Simulation environment and Performance evaluation

4.5.1 Simulation environment

4.5.1.1 Traffic model

In the simulation scenario, the eNB is located at the center of the macrocell using an omni-directional antenna in a 10 MHz bandwidth. Each UE uses a VoIP flow, a Video flow, and a INF-BUF flow at the same time. For the narrowband audio flow, a G.729 voice stream with a bit-rate of 8 kbps was considered. For the wideband audio flow, AMR-WB codec was integrated to the LTE-Sim and was used to simulate. The voice flow is a bursty application that is modelled with an ON/OFF Markov chain [153]. For the video flow, a trace-based application that generates packets based on realistic video trace files with a bit-rate of 242 kbps was used in [154] and it is also available in [17]. In order to obtain a realistic simulation of an H.264 SVC video streaming, we used an encoded video sequence “*foreman.yuv*”, which is publicly available. The LTE propagation loss model is formed by four different models including: Path loss, Multipath, Penetration and Shadowing [155].

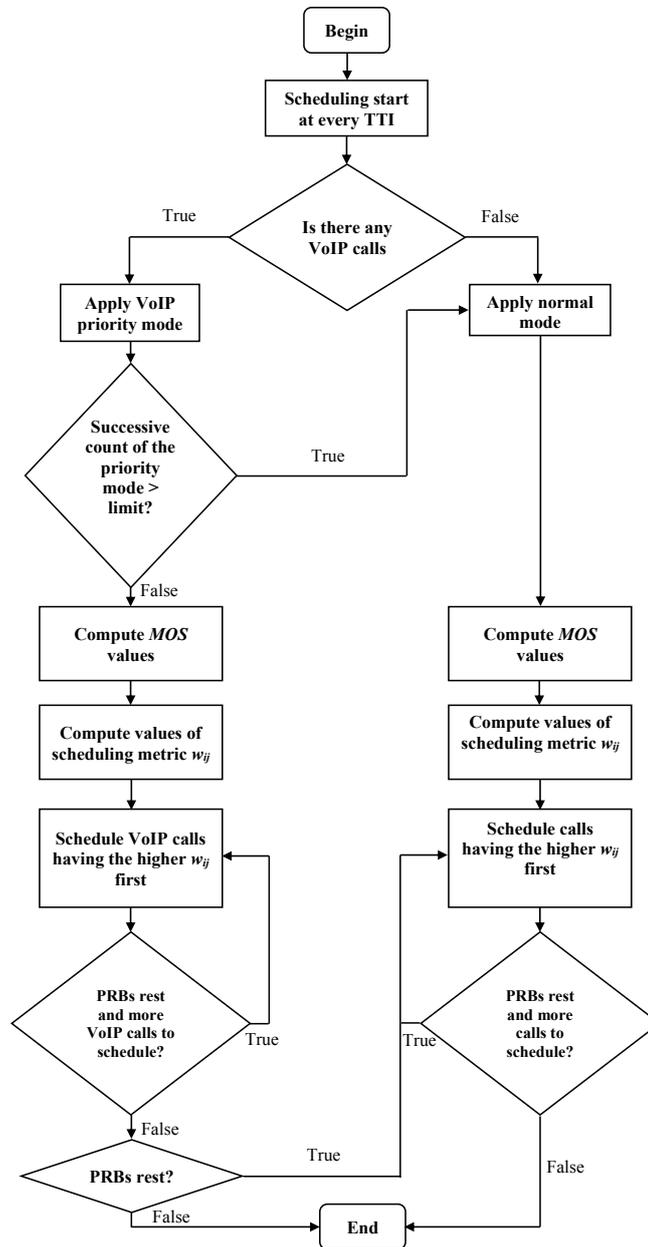


Figure 4.4: The proposed scheduling algorithm

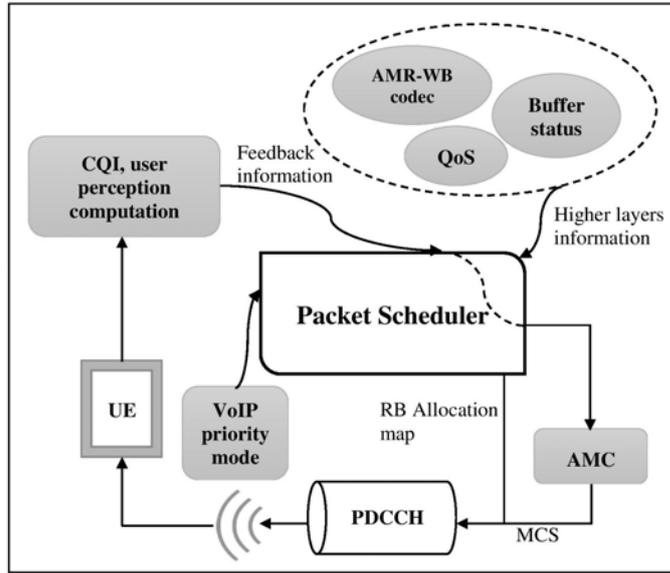


Figure 4.5: Implemented model of the proposed scheduling scheme

- Path loss: $PL = 128.1 + 37.6 \times \log_{10}(d)$, with d is the distance between the UE and the eNB in km
- Multipath: Jakes model
- Penetration loss: 10 dB
- Shadowing: Log-normal distribution with mean 0 dB and standard deviation of 8 dB.

4.5.1.2 Simulation parameters

The simulation process is performed in a single cell with interference with the number of users in the interval [10, 50] which randomly move at a speed of 30 km/h (in city) or at 120 km/h (on highway). There are three simulation scenarios corresponding to the three scheduling schemes. Specifically, simulation scenario 1 corresponds to the scheduling scheme for narrowband audio users, simulation scenario 2 is used for the scheduling scheme for wideband audio users, and simulation scenario 3 belongs to the scheduling scheme for wideband audio users with the integration of VoIP-Priority mode. The other basic parameters used in the simulation scenarios are represented in Table 4.2.

4.5.2 Performance measures

4.5.2.1 End-to-end delay

The end-to-end delay of the system (system delay, D_s) is computed as the average Head of Line (HOL) packet delay for the whole simulation time [156]. Mathematically, it is expressed as follows:

$$D_s = \frac{1}{T} \sum_{t=1}^T \frac{1}{K} \sum_{i=1}^K HOL_i(t) \quad (4.12)$$

Table 4.2: Basic simulation parameters

Simulation Parameters	Simulation scenario 1	Simulation scenario 2	Simulation scenario 3
Simulation duration	100 s	100 s	100 s
Frame structure	FDD	FDD	FDD
Cell radius	1 km	1 km	1 km
Bandwidth	10 MHz	10 MHz	10 MHz
Video bit-rate	242 kbps	242 kbps	242 kbps
VoIP bit-rate	8 kbps	6.6, 8.85, 12.65 kbps	12.65 kbps
User speed	30, 120 km/h	30 km/h	30 km/h
Number of user	10, 20, 30, 40, 50 UEs	10, 20, 30, 40, 50 UEs	10, 20, 30, 40, 50 UEs
Maximum delay	0.1 s	0.1 s	0.1 s
Maximum VoIP packet drop rate			5 %
Minimum VoIP packet drop rate			2 %
MQS	10 ⁵ bytes	10 ⁵ bytes	10 ⁵ bytes
Packet Scheduler	M-LWDF	FLS, M-LWDF, EXP/PF	M-LWDF, WE-MQS, and WE-MQS-VoIP priority
Traffic model	VoIP, Video	VoIP, Video, and INF-BUF	VoIP, Video, and INF-BUF

In which:

- $HOL_i(t)$ is the HOL packet delay of user i at time t
- K is the total number of users
- T is the total simulation time.

4.5.2.2 Packet Loss Rate

The PLR is the ratio of total size of discarded packets resulting in exceeding the time delay to the aggregated packet size of all packets reaching the eNB queue buffer [156]. It is expressed as the following equation.

$$PLR = \frac{\sum_{i=1}^K \sum_{t=1}^T p_discard_i(t)}{\sum_{i=1}^K \sum_{t=1}^T p_size_i(t)} \quad (4.13)$$

Where:

- $p_discard_i(t)$ and $p_size_i(t)$ are the size of discarded packets and the size of all packets that have arrived into the eNB buffer of user i at time t , respectively.
- K and T are similar to the ones in Equation 4.12.

4.5.2.3 Cell throughput

Cell throughput (system throughput, T_s) is defined as the total transmitted packets per second [156]. It is determined as follows.

$$T_s = \sum_{i=1}^K \sum_{t=1}^T p_size_i(t) \quad (4.14)$$

In which:

- $p_size_i(t)$ is the size of transmitted packets of user i at time t .
- K and T are described as same as the ones above.

4.5.2.4 Fairness index

The FI is defined as the difference between the most and the least served users over a given time frame and is determined in service levels [156]. It is mathematically expressed as follows.

$$FI = 1 - \frac{p_size_{max} - p_size_{min}}{\sum_{i=1}^K \sum_{t=1}^T p_size_i(t)} \quad (4.15)$$

Where:

- $(p_size_{max} - p_size_{min})$ are the total size of the transmitted packets of the most and the least served users, respectively.
- $p_size_i(t)$ is the size of all packets that have arrived at the eNB buffer of user i at time t .

4.5.2.5 Spectral efficiency

The spectral efficiency (SE) of the system is the successful usage for whole cell and is defined as follows [157].

$$SE = \frac{1}{B} \sum_{c=1}^C \sum_{n=1}^N \sum_{k=1}^K R_n^m(k, t, c) d_n^m(k, t, c) \quad (4.16)$$

In which:

- B is the system bandwidth.
- C, N, K are the number of carrier, total user, and allocated RBs, respectively.
- $R_n^m(k, t, c)$ is the delivered rate in the past, measured over a fixed window of observation of resource block k of carrier c at time t for user n at rate m .
- $d_n^m(k, t, c)$ is a binary indicator that is set to 1 if the user n at rate m is scheduled on RB k of carrier c at time t and to 0 otherwise.

4.5.3 Performance evaluation

4.5.3.1 Performance evaluation of simulation scenario 1

In order to evaluate the performance of the proposed scheduling scheme, we use LTE-Sim software [17] to simulate the proposed scheduler with the other schedulers including the FLS, M-LWDF and EXP/PF. The performance evaluation is compared in terms of delay, PLR, cell throughput, FI and SE. The analyses of the simulation results are represented in the following subsection.

4.5.3.1.1 Delay of narrowband audio users End-to-end delay (called mouth-to-ear delay) is one of the most important factors when evaluating VoLTE QoS. It is the time required for a packet to be transmitted from source to destination in the network. VoLTE has a very tight delay requirement which should be strictly maintained under limits and has to be carefully monitored. End-to-end delay is measured from input audio signal at transmitting side to output audio signal at receiving side. Normally, end-end-delay includes network delay, encoding delay, decoding delay, compression delay, decompression delay and dejitter_buffer delay. According to recommendations of ITU and 3GPP, one way mouth-to-ear delay should be less than 150 ms. However, a delay budget up to 250 ms is still acceptable if 100 ms extra delay required for packet processing and propagation delay in the congestion core network is considered.

Figure 4.6 illustrates the delay of VoIP flow. It is clear that, the M-LWDF and EXP/PF scheduler has the same lowest delay. The proposed scheduler (E-MQS) has the delay which is not significantly higher than the M-LWDF and EXP/PF schedulers and it outperforms the FLS scheduler. This means FLS scheduler has the highest delay while the proposed scheduler has the second position of the highest delay. The M-LWDF and EXP/PF schedulers are nearly same. The FLS scheduler has the highest delay because it includes two algorithms (at upper and lower levels) thus it need more time to process. However, all the schedulers are much less than allowable threshold.

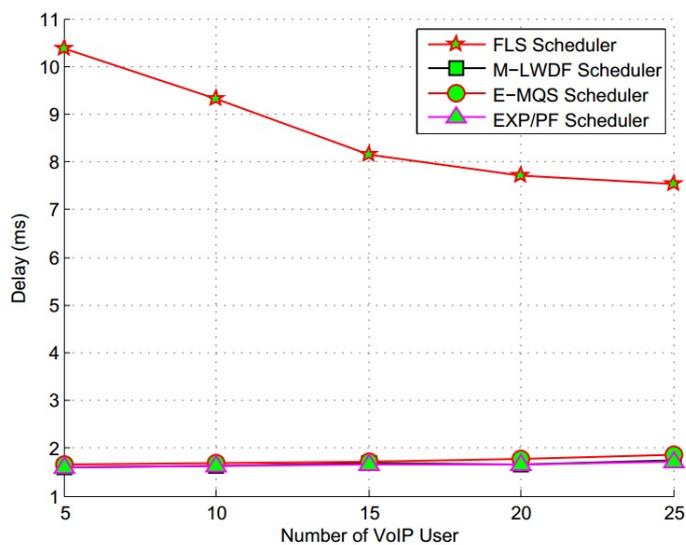


Figure 4.6: Delay vs number of VoIP user

4.5.3.1.2 Packet Loss Rate of narrowband audio users PLR is another important QoS factor of VoLTE to examine and disclose system performance. It shows the failure of one or more transmitted packets to reach their destination across a network. Ideally, in any steady state network, there should be no voice packet loss. However, in fact, voice transmission in wireless networks will include a considerable amount of packet loss and this is why the HARQ technique is used to retransmit lost packets. The failure of voice packets to arrive at the receiver will decrease voice quality and result in poor user perception. However, voice users are still typically satisfied if the PLR is less than 2% based on the

recommendations of the 3GPP.

Figure 4.7 represents the PLR of VoIP flow. When we set the MQS equal to 10^5 bytes (this factor is default set equal to 0 in the LTE-Sim), for the VoIP flow, all schedulers have the decreased PLR when the number of UE increases. Normally, the PLR increases when the number of UE increases, thus, this case is rather special and it represents the unstableness of a real system. As shown in the Figure 4.7, all the schedulers have the PLR which are under 1% while the FLS has the lowest PLR. The E-MQS scheduler has the second position while the M-LWDF has the highest PLR.

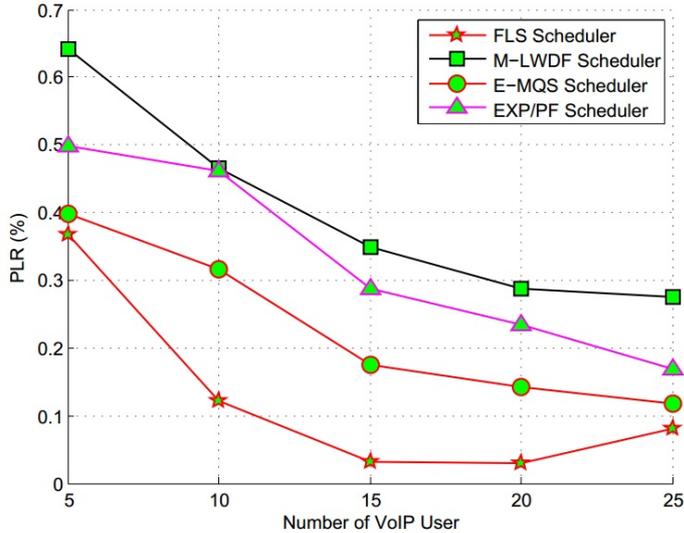


Figure 4.7: PLR vs number of VoIP user

4.5.3.1.3 Cell throughput of narrowband audio users As shown in Figure 4.8, for the VoIP flow, the cell throughput of all the schedulers increases when the number of UE increases. The proposed scheduler always has the cell throughput in top 2 of the highest schedulers for all cases of the number of UEs. This means the proposed scheduler is very suitable for VoIP flow.

4.5.3.1.4 Fairness index of narrowband audio users For the VoIP flow as shown in Figure 4.9, the FIs of all schedulers are not stable when the number of UE increases. However, all these schedulers have the high FIs. The proposed scheduler has the highest FI when the number of UE equals 5, 15, and 20 and obtains the lowest FIs in case of the number of UE equals 10, and 20. In general, the FLS scheduler has the best FI for VoIP flow.

For the FI, it can be conclusive that the FLS scheduler has the best performance. This is due to at the lower layer of this scheduler uses PF algorithm, thus, it ensures the good grade of fairness among multimedia flows. The FI is high which demonstrates that the cell-edge users have been guaranteed the minimum performance. Hence UEs can be served when they move to the edge of the cell.

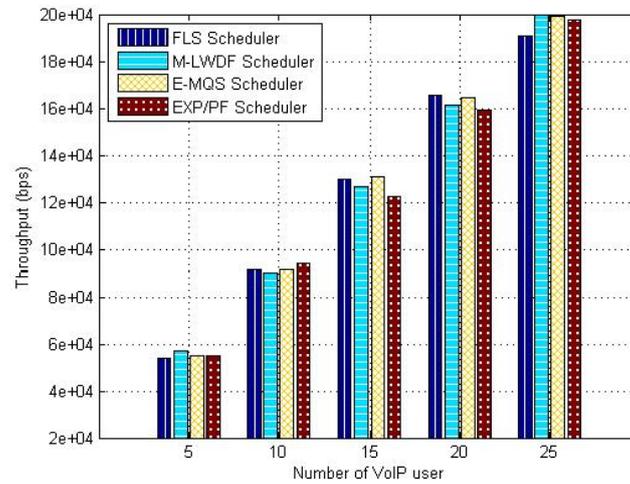


Figure 4.8: Cell throughput vs number of VoIP user

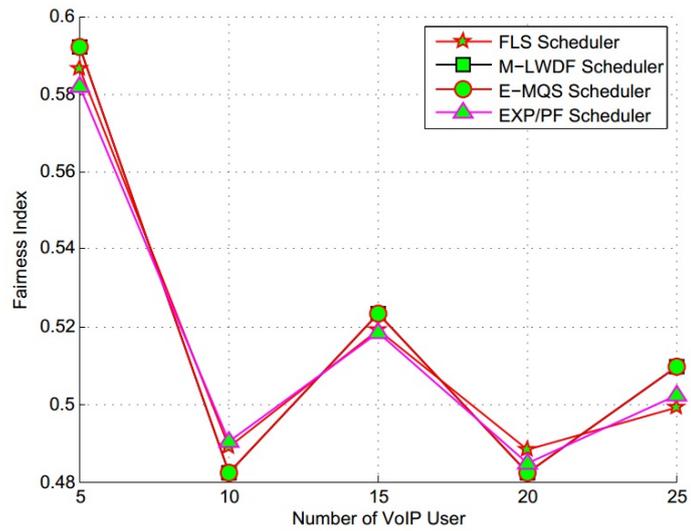


Figure 4.9: FI vs number of VoIP user

4.5.3.1.5 Spectral efficiency of narrowband audio users The successful usage of radio resources is a basic purpose of scheduling algorithms. The SE is seen as the performance measurements for the entire cell. As shown in Figure 4.10, the SE increases when the number of UE increases. In almost cases of the number of UE, the proposed and the FLS schedulers have the same highest SE. However, the difference among schedulers is not significant.

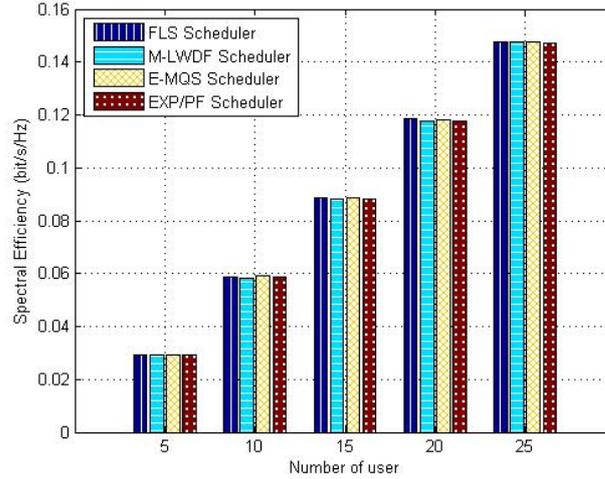


Figure 4.10: Spectral efficiency vs number of user

4.5.3.2 Performance evaluation of simulation scenario 2

In order to evaluate the performance of this proposed scheduling scheme, we still use the LTE-Sim [17] to simulate the proposed scheduler with the other schedulers including the FLS, M-LWDF and EXP/PF. The performance evaluation is compared in terms of delay, PLR, cell throughput, FI and SE. The analysis of the simulation results are represented in the following subsection.

4.5.3.2.1 Delay of wideband audio users Figure 4.11 illustrates the delay of VoIP flow. It is clear that, the proposed scheduler (WE-MQS) has the lowest delay and slightly increases when the number of user (NU) increases, the EXP/PF scheduler keeps the second position of the lowest delay while the FLS scheduler has the highest delay. In this case, the FLS is still rather special and similar to the case of narrowband audio. The M-LWDF has the second highest delay and heavily increases when the NU increases. For the VoIP application, the end-to-end delay should not exceed 150 ms to evaluate voice quality accepted. Although in the simulation scenario, the FLS scheduler has the highest delay but all the schedulers have the delays which are very low, thus, all of them are very suitable for VoIP flow. Compared to the case of narrowband audio, in this case, the M-LWDF and EXP/PF schedulers has not good delay because they strongly increase when the NU increases.

4.5.3.2.2 Packet Loss Rate of wideband audio users PLR shows the failure of one or more transmitted packets to reach their destination across a network. Figure 4.12

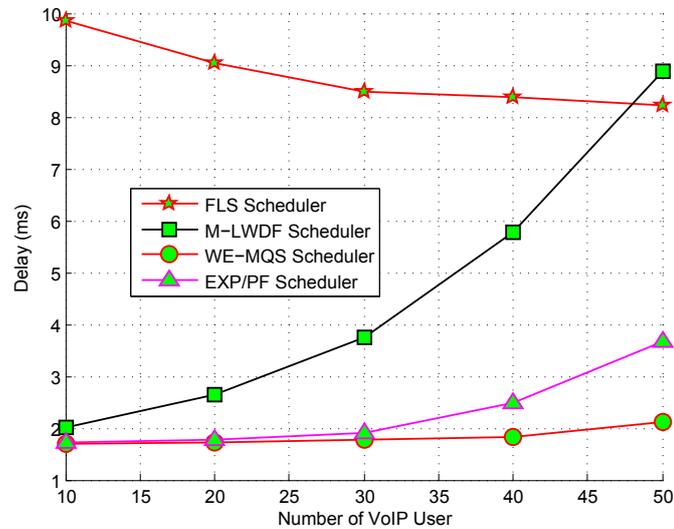


Figure 4.11: Delay vs number of VoIP user

represents the PLR of VoIP flow. When we set the MQS equal to 10^5 bytes (this factor is default set equal to 0 in the LTE-Sim), for the VoIP flow, all schedulers have the PLR decreasing when the number of UE increases. Normally, the PLR increases when the number of UE increases, thus, this case is rather special. Maybe it can be not stable in a real LTE network because the LTE-Sim is rather similar to a real system. As shown in the Figure 4.12, all the schedulers have the PLR which are less than 1%. This is very good for real-time services such as VoIP. It can be said that, the PLR of each of the schedulers is not stable when the NU changes, thus, it's difficult to said that which scheduler has the best PLR because each of them keeps the highest or lowest PLR in a range of the NU. For example, the proposed scheduler has the highest PLR when the NU is in range of 10 to 38, but when the NU is more than 47, the proposed scheduler has the lowest PLR. It is clear that all schedulers have low PLR, thus, they are very suitable for real-time services such as VoIP.

4.5.3.2.3 Cell throughput of wideband audio users Throughput is a measurement of how many units of information that a system can process in a given amount of time. As shown in Figure 4.13, for the VoIP flow, the cell throughput of all the schedulers increases when the NU increases. Although the proposed scheduler does not have the highest cell throughput, but it is always in top 2 or 3 of the highest cell throughput. There is not any scheduler which always has the highest cell throughput when the NU increases. The difference of cell throughput among the schedulers is not considerable.

4.5.3.2.4 Fairness index of wideband audio users FI is a main requirement that should be taken into account to ensure a minimum performance to the edge-cell users. For the VoIP flow as shown in Figure 4.14, the FIs of all schedulers are not stable when the NU increases. Normally, the FI decreases when the NU increases. It is really difficult to say that which scheduler has the highest FI because they are always changed when the NU increases. The proposed scheduler is almost in the middle of the schedulers. It is similar to the cell throughput, for the FI, the difference among the schedulers is not also significant.

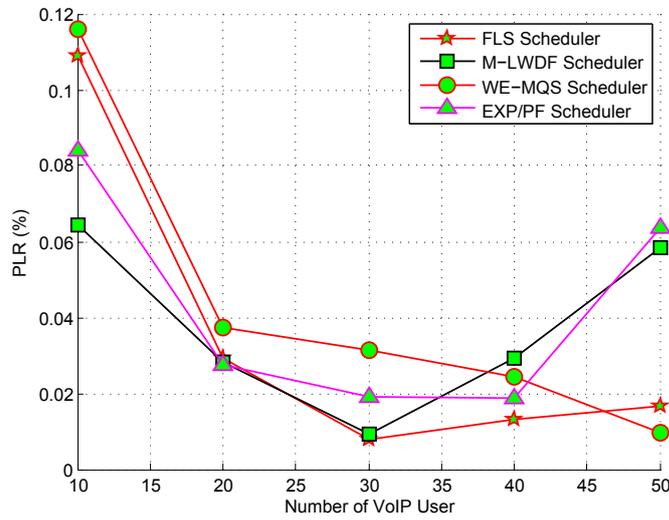


Figure 4.12: Packet loss rate vs number of VoIP user

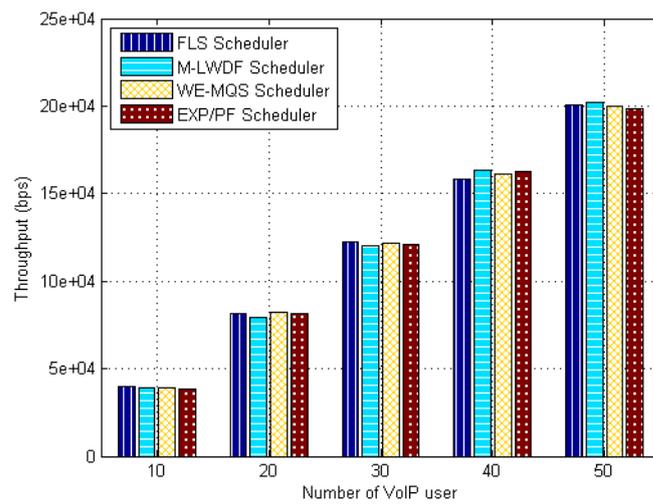


Figure 4.13: Throughput vs number of VoIP user

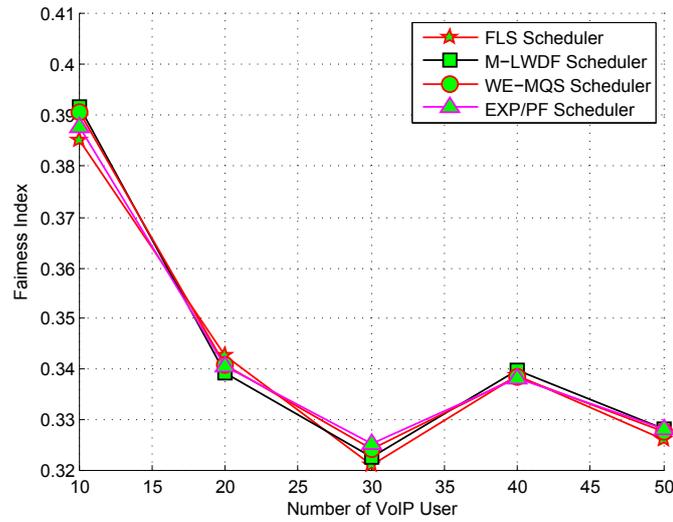


Figure 4.14: Fairness Index vs number of VoIP user

4.5.3.2.5 Spectral efficiency SE is successful usage of radio resources. It is a major principle of the scheduler. SE is regarded as the performance measurements for whole cell. As shown in the Figure 4.15, the M-LWDF and the EXP/PF schedulers always keeps the first and the second positions of the highest SE, respectively. The proposed scheduler always has the second lowest SE while the FLS scheduler has the lowest SE. However, the difference among the schedulers is not also considerable. This means all the schedulers are efficient, thus, they are very suitable for VoIP users.

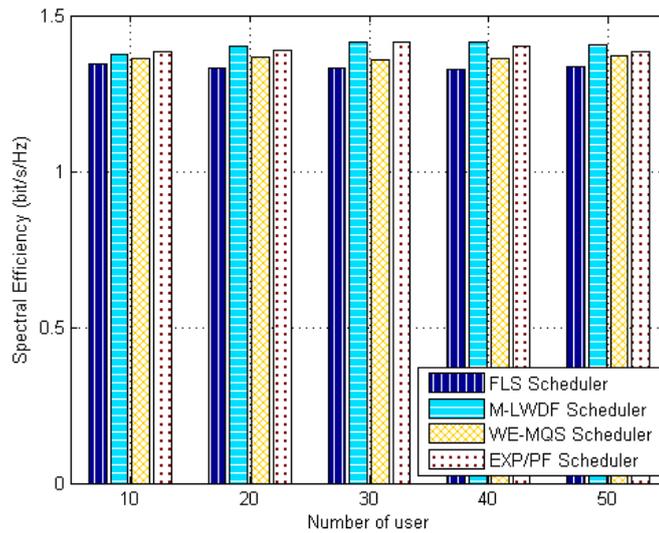


Figure 4.15: Spectral efficiency vs number of user

4.5.3.3 Performance evaluation of simulation scenario 3

For assessing the performance of this proposed scheduler, we compare it to the M-LWDF scheduler, and to the scheduler proposed in [8] (WE-MQS scheduler). We assess the performance in terms of delay, PLR, cell throughput, FI and SE. The analyses of the simulation results are represented in the following subsection.

4.5.3.3.1 Delay vs VoIP user End-to-end delay is the duration required for a packet to be transmitted from source to destination. Figure 4.16 illustrates the delay of VoIP flow. Such as shown in this figure, the priority mode of the proposed scheduler has the lowest delay when compared to the normal mode and the M-LWDF scheduler. Both modes of the proposed scheduler slightly increase when the NU increases while the M-LWDF scheduler strongly increases. In can be said that, when the VoIP priority mode is integrated, the delay decreases significantly.

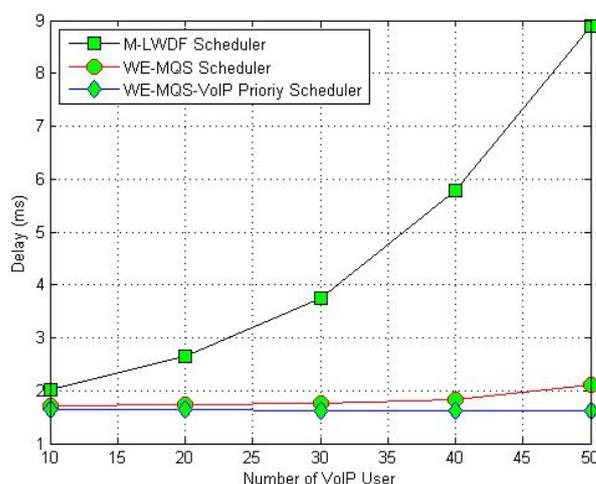


Figure 4.16: Delay vs number of VoIP user

4.5.3.3.2 PLR vs VoIP user PLR represents the failure of one or more packets which are not transmitted successfully to destination. Figure 4.17 represents the PLR of VoLTE traffic. Normally, the PLR increase when the NU increases. In this study, we assess the system performance in a heterogeneous LTE network with mobility. The results in Figure 4.17 are rather special, specifically for the normal mode, the PLR decreases when the NU increases. This may be due to the unstableness in a real system. In general, the priority mode has the lowest PLR in comparison with two remaining others except when the NU equals 30.

For the delay and PLR, it can be concluded that the priority mode is very suitable for VoIP service because it has the best performance.

4.5.3.3.3 Cell throughput Throughput is a measurement of how many units of information a system can process in a given amount of time. As shown in Figure 4.18, for the VoIP flow, the cell throughput of all the schedulers increases when the NU increases.

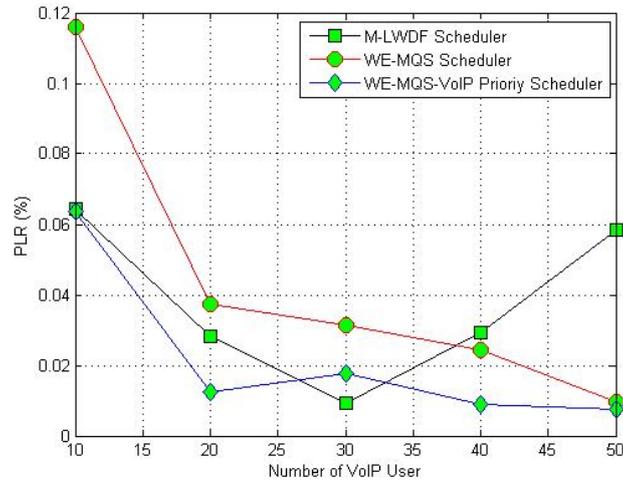


Figure 4.17: Packet loss rate vs number of VoIP user

The priority mode has the best cell throughput when the NU is in range of 10 to 30 and slightly reduces when the NU is more than 30, 40 when compared to the M-LWDF and to the normal mode, respectively.

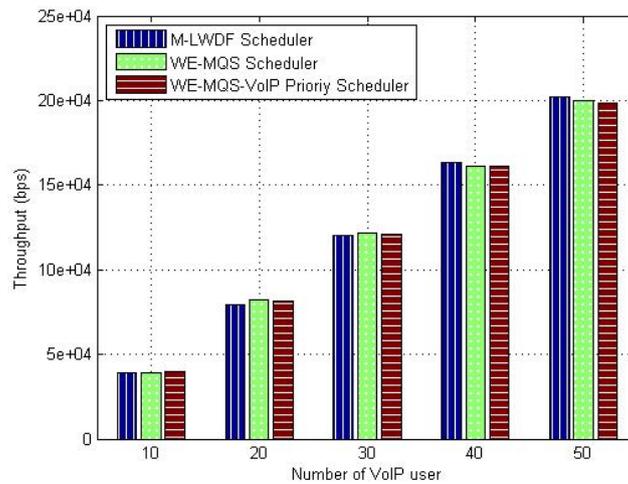


Figure 4.18: Throughput vs number of VoIP user

4.5.3.3.4 Fairness index FI is a main demand that is used to guarantee a minimum performance to edge-cell users. For the VoIP flow as shown in Figure 4.19, the FIs of all the schedulers are not stable when the NU increases. Normally, the FI decreases when the NU increases. The VoIP priority mode has the best FI when the NU is less than 30 and has the lowest FI when the NU is more than 30. However, the difference is not significant. The normal mode is always in the middle of two others.

4.5.3.3.5 Spectral efficiency SE expresses successful usage of radio resources. It is a major principle of the scheduler. SE aims at performance measurements for entire cell. As

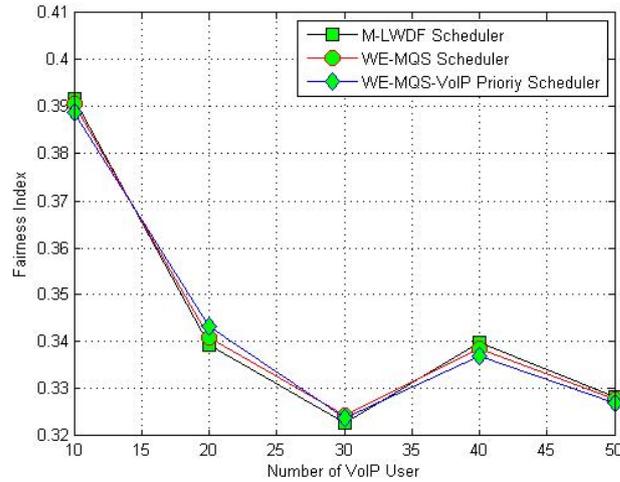


Figure 4.19: Fairness Index vs number of VoIP user

shown in the Figure 4.20, the normal mode almost has the lowest SE for all of the NU while the VoIP priority mode has the same SE when compared to the M-LWDF scheduler when the NU is less than 40. When the NU is more than 40, the M-LWDF scheduler has the higher SE in comparison with the VoIP priority mode. However, this increase is not very considerable.

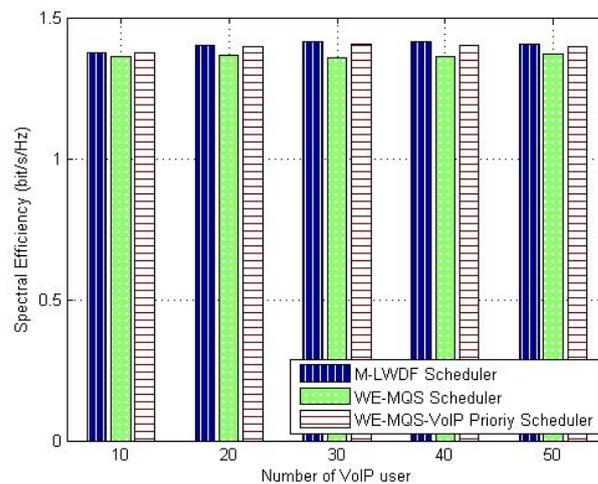


Figure 4.20: Spectral efficiency vs number of VoIP user

4.5.4 Comparison of three proposed scheduling schemes

In this subsection, we describe the comparison of three scheduling schemes proposed in this Chapter. The characteristics of them are listed in Table 4.3. It is clear that, all the schedulers have characteristics of real-time voice services such as VoLTE. Hence, all of them are suitable for real-time voice services where E-MQS is suitable for narrowband audio

services, WE-MQS is suitable for wideband audio services, and when there needs to be an enhancement for VoIP service, WE-MQS-VoIP Priority is the perfect choice.

Table 4.3: Comparison of three proposed schedulers

Scheduler	Threshold		Channel conditions	Buffer status	User perception		Source codec		VoIP priority mode
	Delay	PLR			Extended E-model	WB E-model	G.729	AMR-WB	
E-MQS	x	x	x	x	x		x		
WE-MQS	x	x	x	x		x		x	
WE-MQS-VoIP Priority	x	x	x	x		x		x	x

4.6 Summary

In this Chapter, we propose new Channel-, QoS- and QoE-Aware scheduling schemes for downlink direction in LTE network. The main idea in the proposed schedulers is the consideration of user perception and the MQS factors into the priority metrics in the proposed schedulers. The metric is determined based on the user perception (MOS), the remaining queue size, the fixed maximum time, and the channel conditions. We develop the proposed scheduler for both cases of narrowband and wideband audio users. Besides, in order to enhance VoIP users, the VoIP-Priority mode is integrated into the proposed scheduler. For the narrowband audio users (i.e. simulation scenario 1), the simulation results show that the proposed scheduler not only meets QoS requirements for voice services but also outperforms the FLS, M-LWDF, EXP/PF schedulers in terms of delay, cell throughput, FI and SE. For the PLR, the proposed scheduler has the performance not as well as the FLS scheduler. For the FI, the proposed scheduler does not have the stable performance, it can obtain either the highest or the lowest FI values. It can be conclusive that when considering the MOS and the MQS as factors for the metric in the proposed scheduler, the system performance has been improved significantly. Through all simulation results, it can be seen that, the FLS has the best performance, and the proposed scheduler keeps the second best position in scheduling for VoIP flow.

For the wideband audio users (i.e. simulation scenario 2), the simulation results show that the proposed scheduler not only meets QoS requirements for voice services but also outperforms the FLS, M-LWDF, EXP/PF schedulers in terms of delay and PLR when the NU is more than 47. It is clear that, the FLS, M-LWDF, and EXP/PF are well-known schedulers in wireless networks, and all of them are very suitable for real-time services such as VoIP. The advantage of the proposed scheduler is that it takes the user satisfaction and the remaining queue size into account. In addition, the proposed scheduler integrates the AMR-WB codec which is mandatory for VoLTE to get high user perception. This overcomes the limitation of the LTE-Sim that it supports only G.729 codec for VoIP application. When integrating the VoIP-Priority mode (i.e. simulation scenario 3), the simulation results show that the proposed scheduler meets QoS requirements for voice services. In addition, it overcomes the M-LWDF scheduler and the normal mode for delay and PLR. With the integration of the VoIP priority mode, the proposed scheduler is enabled, thus, it has the lowest delay and PLR. For the throughput, FI, and SE, in general, it nearly has the same performance when compared to the M-LWDF scheduler. The advantage of the proposed scheduler is that it takes the user satisfaction and the remaining queue size into account

and the presence of the VoIP-Priority mode, thus, it enables the priority for VoIP packets than other traffics.

It can be concluded that when considering user satisfaction and the MQS as essential factors for the priority metrics in the proposed schedulers and the integration of the VoIP priority mode, the system performance is improved considerably. Through all simulation results, it can be seen that the proposed schedulers have the best performance for VoIP users. Therefore, the proposed schedulers are very suitable and efficient for voice services in LTE downlink direction. The WE-MQS-VoIP Priority scheduler has the best performance and is most suitable for real-time voice services such as VoLTE..

Chapter 5

The proposed object non-intrusive models for predicting voice quality in LTE networks

Contents

5.1	General idea	98
5.2	The proposed object non-intrusive models for predicting voice quality in LTE networks	98
5.2.1	The implemented model for predicting narrowband audio quality	98
5.2.2	The implemented model for predicting wideband audio quality	99
5.3	Simulation environment and Performance evaluation	101
5.3.1	Simulation environment	101
5.3.1.1	Traffic model	101
5.3.1.2	Simulation parameters	101
5.3.2	Performance evaluation of the proposed model for predicting narrowband audio quality	101
5.3.2.1	Effects of jitter buffer on narrowband audio quality	102
5.3.2.2	Effects of delay on narrowband audio quality	102
5.3.2.3	Effects of PLR on narrowband audio quality	105
5.3.3	Performance evaluation of the proposed model for predicting wideband audio quality	105
5.3.3.1	Effects of AMR-WB modes on wideband audio quality	107
5.3.3.2	Effects of delay on wideband audio quality	108
5.3.3.3	Effects of PLR on wideband audio quality	108
5.4	Summary	111

This Chapter presents the proposed object non-intrusive models for predicting voice quality for both narrowband and wideband flows in LTE network. The main idea is represented in Section 5.1. Section 5.2 represents the proposed object non-intrusive models for predicting voice quality. Simulation environment and Performance of the proposed models are described in Section 5.3. Summary of the proposals in this Chapter are represented in Section 5.4.

5.1 General idea

As mentioned in Chapter 1, the intrusive models (e.g. PESQ/PESQ-WB) allow measuring exactly user perception than the non-intrusive ones. However, they are not suitable for real-time services such as VoIP or VoLTE because they require to refer to the original signals. Otherwise, the non-intrusive methods are very suitable for predicting user satisfaction. For this type, RNN-based is also a suitable choice. However, in this Chapter, we present another non-intrusive method. According to our knowledge, at present, there have not been any model for this type. As mentioned before, the E-model or the WB E-model are computational models that allow predicting user perception without the reference to the original signals. And we see that, the lack of the E-model and WB E-model is how to determine its input parameters. Therefore, we need a method to determine the input factors of them. The calculation for inputs of the extended E-model as well as WB E-model are described in Chapter 4. In order to determine them, in this Chapter, we propose to use the LTE-Sim [17] software to calculate delay and PLR for voice users. The LTE-Sim is a famous framework which allow simulating entire LTE network, thus, it is rather similar to a real system. Outputs of the LTE-Sim are delay and PLR which are essential input parameters of the E-model or the WB E-model. In addition, we also take effects of network jitter into account by considering jitter parameter (I_j) as an input parameter of the E-model. For the wideband audio flow, due to the LTE-Sim only support G.729 codec (i.e. used for narrowband audio), we propose to complement AMR-WB codec into LTE-Sim. This supports the LTE-Sim to simulate wideband audio traffics. And in order to obtain more real results, we simulate voice service with mobility in LTE heterogeneous network. It can be said that, the main idea in this Chapter is proceeded from the ideas in Chapter 4 in which the extended E-model and WB E-model are used to predict voice quality over LTE network for narrowband and wideband audio users, respectively. We see that, MNOs have to always monitor user perception to adjust network impairments instantly. Since the extended E-model as well as WB E-model do not require original signal to refer, they are very suitable for measuring quality of live voice services such as VoLTE.

5.2 The proposed object non-intrusive models for predicting voice quality in LTE networks

5.2.1 The implemented model for predicting narrowband audilo quality

VoLTE is a real-time service, and is fully deployed over an IP network, thus, ensuring VoLTE quality is a big challenge. There are very little methods that allow monitoring and predicting VoLTE quality and most of them are subjective ones which have to refer to original signal, thus, they are not suitable for real-time services such as VoIP. In this section, we propose a new object non-intrusive voice quality assessment method which is based on the LTE-Sim software [17] and the extension of the E-model [31]. The LTE-Sim is a software which allow simulating VoIP flow that is rather similar to a real flow. Therefore, we see that, the simulation results are rather exact. The E-model is a computational model which allows predicting voice quality when it is transmitted from source to destination. However, in this model, there is not presence of network jitter. For VoIP flow, network jitter affects significantly on voice quality, thus, we propose to add the effects of network jitter to E-model via I_j factor. When the I_j is added to the E-model, the value of R factor

will be decreased, it leads to the lower user perception when compared to the standard E-model. The proposed model is represented on Figure 5.1.

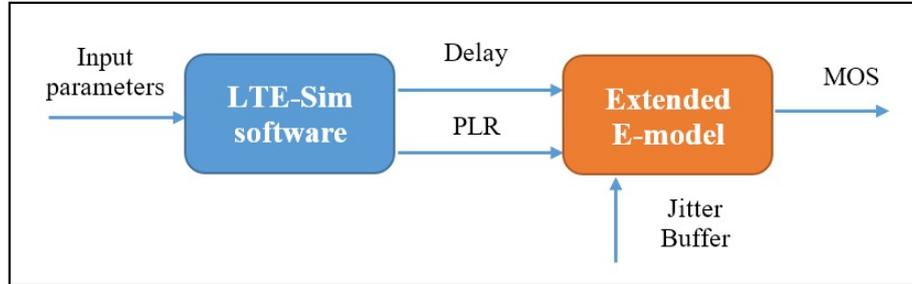


Figure 5.1: The proposed model for predicting narrowband audio quality

The principle of the proposed model as follows: Input parameters of the simulation scenarios are first fed to the LTE-Sim software. After finishing the simulation, we receive simulation results. In these results, delay and PLR are chosen according to the number of user. These factors combining with jitter buffer (JB) are the input parameters of the extended E-model. The output of E-model is R factor, and then it is mapped to MOS score. This MOS score performs user satisfaction. The steps of the proposed model are described as follows:

Algorithm 6 Predicting narrowband audio quality in LTE network

Step 1: Setting input parameters of the simulation scenario

Step 2: Simulating the scenario in the LTE-Sim software

Step 3: Reading the delay, PLR results from the output files of Step 2, and selecting values of JB

Step 4: Feeding the input parameters taken from Step 3 to the extended E-model which is programmed in C-language as follows:

- Calculating R-factor of the extended E-model (refer to Chapter 4)
 - R is then mapped to MOS score (refer to Chapter 1)
-

LTE is a packet-switched network, thus, in order to simulate it as a real system, we select input data flows including a VoIP, a Video and a non real-time flow. In addition, the mobility is also included in our scenario, specifically, we select user speed is 30 km/h (in city) and 120 km/h (on highway). We select the M-LWDF scheduler which is very suitable for real-time services. Besides predicting user perception, we assess the effects of delay, PLR on user satisfaction according to the number of user.

5.2.2 The implemented model for predicting wideband audio quality

It's similar to predicting narrowband audio quality, to measure quality for wideband audio, we combine the LTE-Sim and the WB E-model. However, as mentioned above, the LTE-Sim does not support for simulating wideband audio flows, thus, we propose to add AMR-WB codec to the LTE-Sim. It's also as same as E-model, the WB E-model is a computational model which allows predicting voice quality when it is transmitted from source

to destination. Each bitrate of AMR-WB codec corresponds to a mode. The modes of AMR-WB are configured into 3 configurations such as described in Chapter 1. Each is adapted according to channel quality. Therefore, in order to add this codec to LTE-Sim, we build procedures which allow changing instantly the AMR-WB mode according to channel condition via Signal-to-interference (C/I) ratio. The C/I ratio is calculated at receiver and is sent back to the eNodeB. The changes among the AMR-WB modes depends on C/I thresholds. The proposed procedures for changing the AMR-WB modes (or changing packet size of AMR-WB codec) for the configuration A, B, and C are represented in Chapter 4. The values of C/I thresholds are referred to [149] and [150]. Besides, we also have several essential modifications in the LTE-Sim to suit for the simulation scenario.

The proposed model in this case is represented on Figure 5.2.

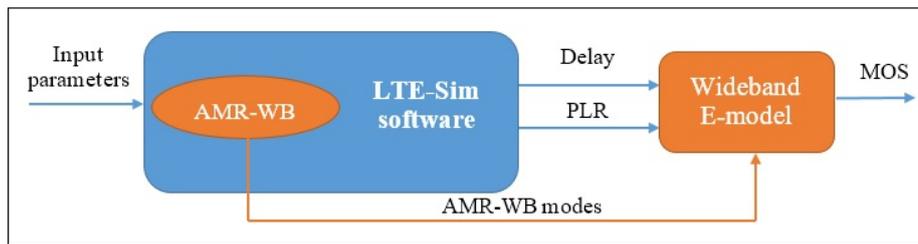


Figure 5.2: The proposed model for predicting wideband audio quality

The principle of the proposed model as follows: Input parameters of the simulation scenarios are firstly fed to the LTE-Sim software. After finishing the simulation, we receive simulation results. In these results, delay and PLR are collected according to the number of user. These factors are the input parameters of the WB E-model. The output of WB E-model is R -factor, and then it is mapped to MOS score. This MOS score performs the user satisfaction. The steps of the proposed model are similar to the narrowband audio flow as follows.

Algorithm 7 Predicting wideband audio quality in LTE network

Step 1: Setting input parameters of the simulation scenario

Step 2: Complementing the procedure of changing AMR-WB packet size according to the C/I ratio into LTE-Sim

Step 3: Simulating the simulation scenario in the LTE-Sim software

Step 4: Reading the delay, PLR results from the output files of Step 3, and selecting values of AMR-WB modes

Step 5: Feeding the input parameters taken from Step 4 to the WB E-model which is programmed in C-language as follows:

- Calculating R_{wb} -factor of the WB E-model (refer to Chapter 4)
 - R_{wb} is then converted to R -factor before being mapped to MOS (refer to Chapters 1 and 4)
-

In order to simulate VoLTE as in a real LTE network, we evaluate the system performance in a heterogeneous network. Specifically, each user utilizes a VoIP, a Video and a non real-time flow at an instant. In addition, the mobility is also included in the simulation

scenario, specifically, we select user speed is 30 km/h. We choose the FLS [87], M-LWDF [147], and EXP/PF [148] schedulers which are very suitable for real-time services. Besides predicting user perception, we assess the effects of delay, PLR on user satisfaction according to the number of user.

5.3 Simulation environment and Performance evaluation

5.3.1 Simulation environment

5.3.1.1 Traffic model

In the simulation scenario, the eNB is located at the center of the macrocell using an omni-directional antenna in a 10 MHz bandwidth. Each UE uses a VoIP flow, a Video flow, and a INF-BUF flow at the same time. For the narrowband audio flow, a G.729 voice stream with a bitrate of 8 kbps was considered. For the wideband audio flow, AMR-WB codec was integrated to the LTE-Sim and was used to simulate. The voice flow is a bursty application that is modelled with an ON/OFF Markov chain [153]. For the video flow, a trace-based application that generates packets based on realistic video trace files with a bitrate of 242 kbps was used in [154] and it is also available in [17]. In order to obtain a realistic simulation of an H.264 SVC video streaming, we use an encoded video sequence “*foreman.yuv*”, which is publicly available. The LTE propagation loss model is formed by four different models including: Path loss, Multipath, Penetration and Shadowing [155] as follows.

- Path loss: $PL = 128.1 + 37.6 \times \log_{10}(d)$, with d is the distance between the UE and the eNB in km
- Multipath: Jakes model
- Penetration loss: 10 dB
- Shadowing: Log-normal distribution with mean 0 dB and standard deviation of 8 dB.

5.3.1.2 Simulation parameters

The simulation process is performed in a single cell with interference with the number of users in the interval [10, 50] which randomly move at a speed of 30 and 120 km/h. There are two simulation scenarios corresponding narrowband and wideband voice traffic flows called Simulation scenario 1 and Simulation scenario 2 used in this Chapter. The other basic parameters used in the simulations are represented in Table 5.1.

5.3.2 Performance evaluation of the proposed model for predicting narrowband audio quality

In order to simulate the traffic model in LTE-Sim, we use Modified Largest Weighted Delay First (M-LWDF) [147] scheduler with mobility in LTE heterogeneous network. The M-LWDF scheduler is very suitable for real-time services such as VoIP, Video, Gaming, etc. The analyses of the simulation results are represented in the following subsections.

Table 5.1: Basic simulation parameters

Simulation Parameters	Simulation scenario 1	Simulation scenario 2
Simulation duration	100 s	100 s
Frame structure	FDD	FDD
Cell radius	1 km	1 km
Bandwidth	10 MHz	10 MHz
Video bit-rate	242 kbps	242 kbps
Voice source codec	G.729	AMR-WB
VoIP bit-rate	8 kbps	6.6, 8.85, 12.65 kbps
User speed	30, 120 km/h	30 km/h
Number of user	10, 20, 30, 40, 50 UEs	10, 20, 30, 40, 50 UEs
Maximum delay	0.1 s	0.1 s
Packet Scheduler	M-LWDF	FLS, M-LWDF, EXP/PF
Traffic model	VoIP, Video, and INF-BUF	VoIP, Video, and INF-BUF

5.3.2.1 Effects of jitter buffer on narrowband audio quality

In order to evaluate the effects of JB on voice quality, we set values of this factor of 40, 60, 80, 100, and 120. Figure 5.3 shows the relationship of the number of user (NU) vs user perception (MOS score) when user speed is 30 km/h. It's clear that, in this case, the E-model has the highest MOS score while the proposed model has the lower MOS. This is obvious because the proposed model considers effect of network jitter on the E-model. All cases have the slightly degraded MOS when the NU increases. This figure also shows that the higher JB, the higher MOS. However, in fact, JB should not be too high. As shown on this figure, when JB is more than 80, the MOS increases not significantly. This is rather clear when JB corresponds to 100 and 120. It can be seen that, when JB is more than 100, its effect on user perception is not significant.

Figure 5.4 shows the effects of JB when the speed is 120 km/h. It's clear that, when the speed is high, the user satisfaction is decreased clearly. All cases have the MOS score heavily decreasing when NU increasing. It's obvious, but the reduction is rather different from the case of the speed of 30 km/h. It's similar to the previous case, the higher JB, the higher user perception. The JB is directly proportional to user satisfaction while the speed is inverse.

5.3.2.2 Effects of delay on narrowband audio quality

Figure 5.5 shows the relationship between delay vs user perception according to the NU. It's clear that, the delay increases when the NU increases. This leads to lower MOS score. The E-model has the highest MOS while the Extended E-model has the lower one. It can be seen that, the higher NU or the higher delay is, the lower MOS score is. This principle is suitable for both the E-model and the Extended E-model for all cases of JB. The detailed results of this case are represented in Table 5.2.

It's similar to the case of delay for user speed of 30 km/h, the effects of delay on user satisfaction when user speed of 120 km/h is shown in Figure 5.6. The principle is similar to the case of user speed of 30 km/h, however, the MOS score is lower and strongly decreases when the NU or delay increases. This is right in both the E-model and the extended E-model. The detailed results are described in Table 5.2.

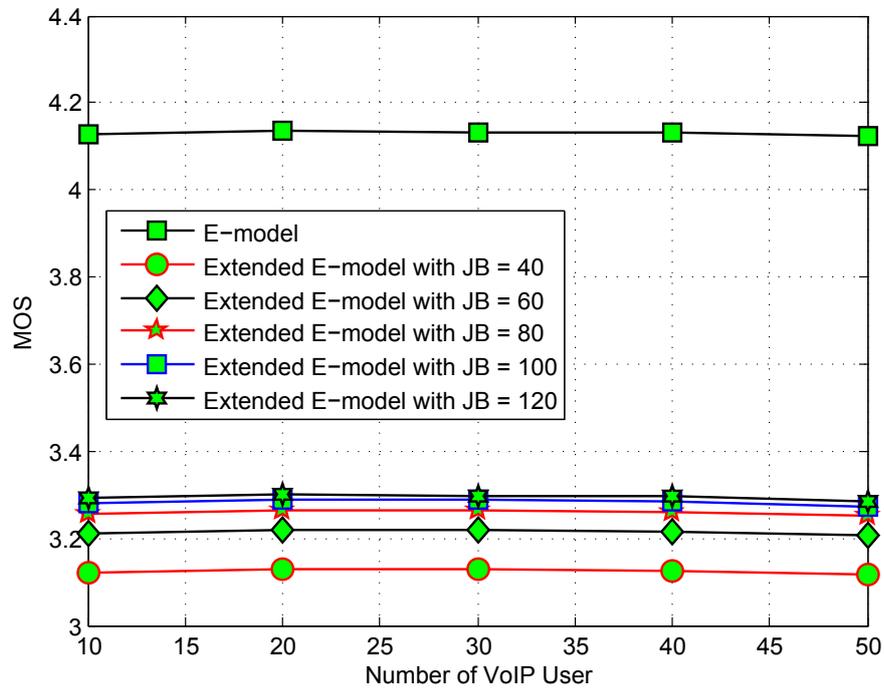


Figure 5.3: MOS vs number of VoIP user at the speed of 30 km/h

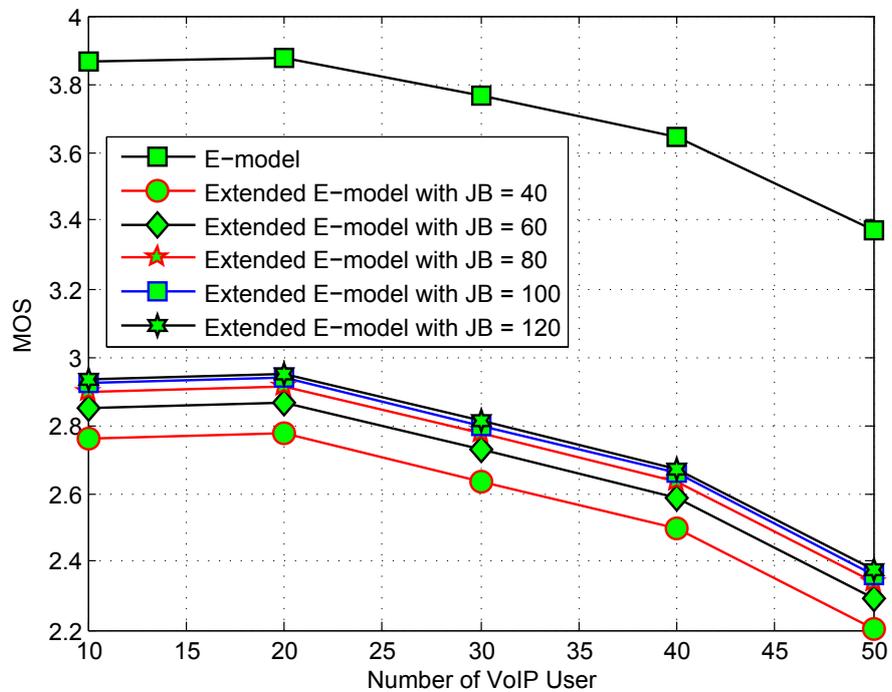


Figure 5.4: MOS vs number of VoIP user at the speed of 120 km/h

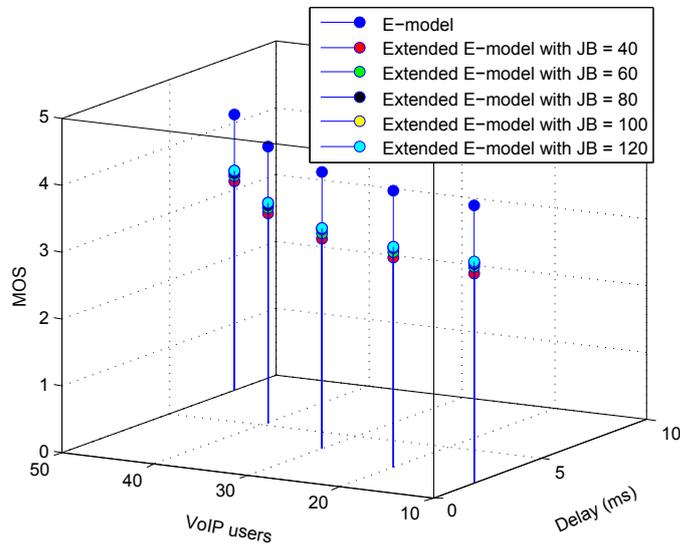


Figure 5.5: Effects of Delay on voice quality at the speed of 30 km/h

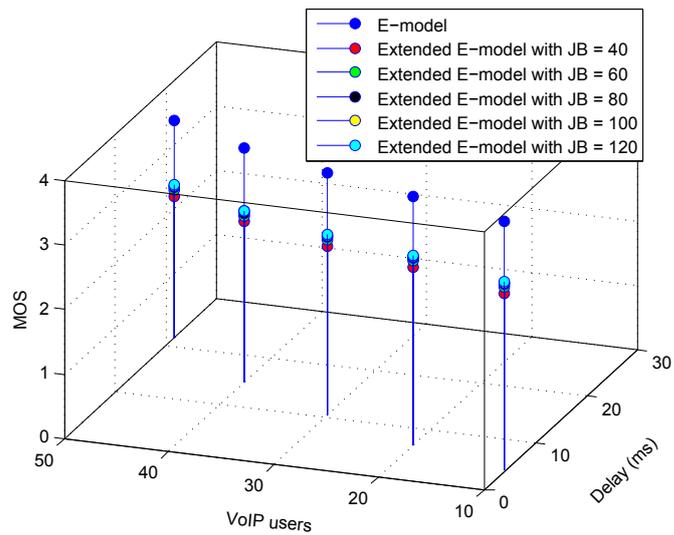


Figure 5.6: Effects of Delay on voice quality at the speed of 120 km/h

5.3.2.3 Effects of PLR on narrowband audio quality

Figure 5.7 illustrates the effects of delay on user perception according to NU. It's clear that, MOS increases when the NU or PLR increases. This is suitable for both E-model and extended E-model. Normally, PLR increases when NU increases, but there is a special case for the NU equals 10. This may be due to the system is always unstable in fact. For both the E-model and the Extended E-model, the higher NU or the higher PLR is, the lower MOS score is. The E-model has the higher MOS when compared to the extended E-model. For the extended E-model, rule for the higher JB, the higher MOS is still met. For the detailed results, refer to Table 5.2.

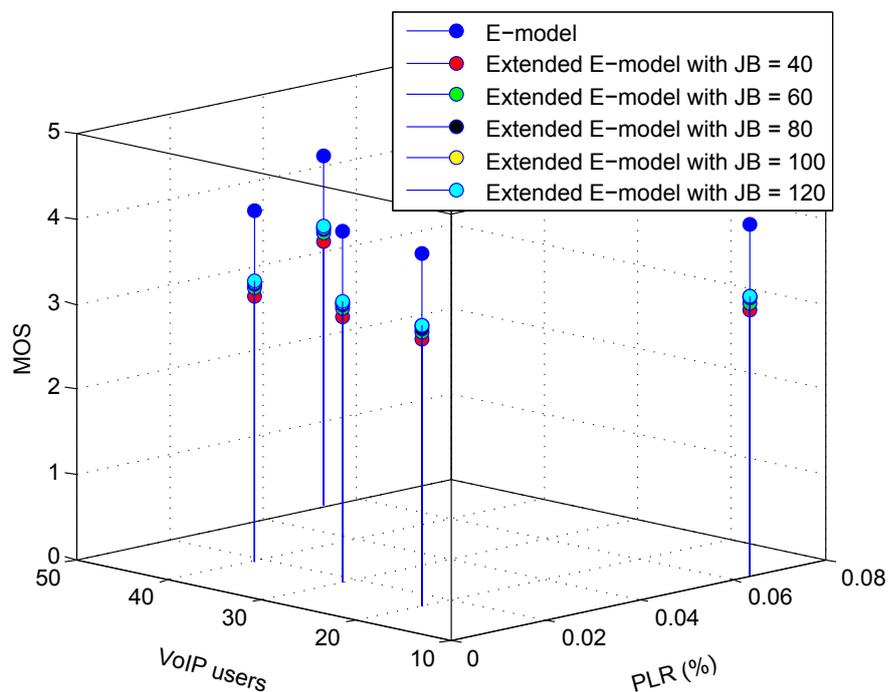


Figure 5.7: Effects of PLR on voice quality at the speed of 30 km/h

In case of user speed equals 120 km/h, the rule for the case of user speed of 30 km/h is still met. The difference here is that in this case, user perception decreases in comparison with the case of user speed of 30 km/h and it strongly decreases when the NU or PLR increases. This is fully suitable. The detailed results of this case are represented in Table 5.2.

5.3.3 Performance evaluation of the proposed model for predicting wide-band audio quality

The analyses of the simulation results are represented in the following subsections.

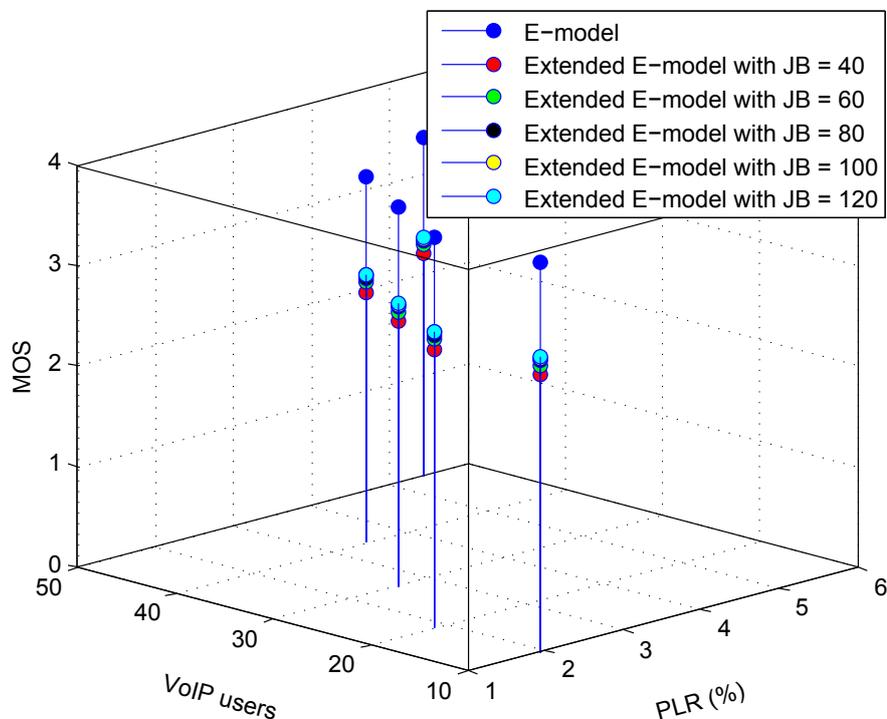


Figure 5.8: Effects of PLR on voice quality at the speed of 120 km/h

Table 5.2: The detailed results of Figures of 5.5 - 5.8

NU	Delay (ms)	PLR (%)	MOS					
			E-model	Extended E-model				
				JB=40	JB=60	JB=80	JB=100	JB=120
User speed: 30 km/h								
10	1.94	0.064	4.1287	3.1204	3.21	3.2556	3.2789	3.2909
20	2.47	0.014	4.1351	3.13	3.2195	3.2651	3.2884	3.3003
30	3.45	0.017	4.1339	3.1282	3.2177	3.2633	3.2886	3.2985
40	5.26	0.018	4.1322	3.1257	3.2153	2.2609	3.2842	3.2961
50	8.03	0.053	4.1252	3.1151	3.2048	3.2504	3.2738	3.2857
User speed: 120 km/h								
10	4.04	1.943	3.8643	2.7591	2.851	2.8981	2.9223	2.9347
20	6.64	1.833	3.8772	2.7753	2.8672	2.9143	2.9385	2.9508
30	10.36	2.618	3.7632	2.6352	2.7271	2.7743	2.7986	2.811
40	14.74	3.448	3.6434	2.496	2.5874	2.6345	2.6587	2.6711
50	21.54	5.432	3.371	2.2044	2.293	2.3389	2.3626	2.3748

5.3.3.1 Effects of AMR-WB modes on wideband audio quality

As mentioned above, in this Chapter, we only simulate the proposed model for wideband audio flow with the configuration A of the AMR-WB codec. This configuration includes three modes of 0, 1, and 2. Figure 5.9 represents the effects of AMR-WB mode 0 on voice quality. It is clear that, the voice quality slightly decreases when the number of VoIP user (NU) increases excepting the FLS scheduler. The FLS scheduler has the best quality compared to the other schedulers, even, it does not decrease when the NU raises. This is rather special. It is due to the unstableness in fact. The MOS score of the FLS scheduler can decrease when the NU is more than 50. The M-LWDF scheduler has the worst quality and it seems to heavily decreases when the NU increases.

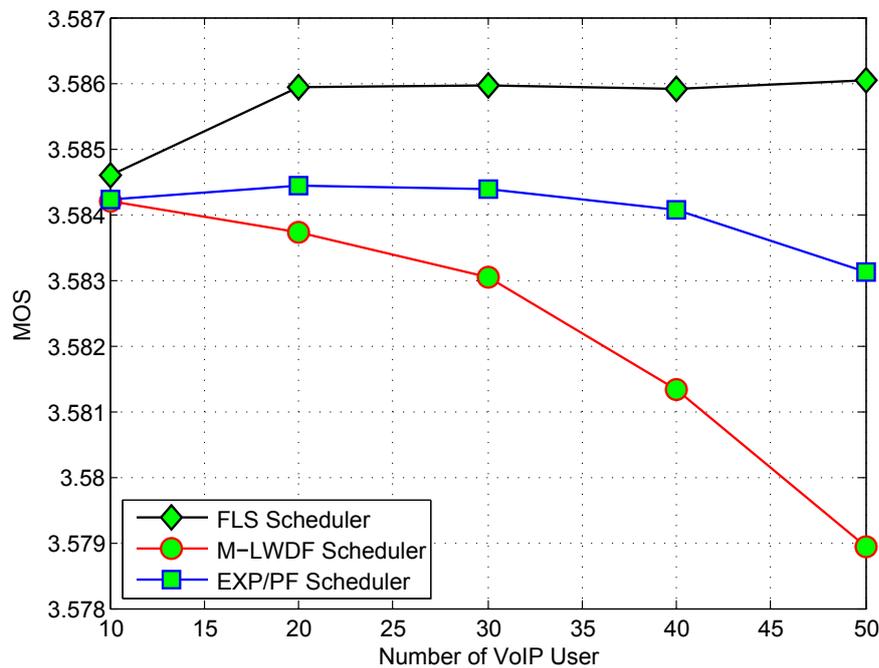


Figure 5.9: MOS vs number of VoIP user with AMR-WB mode 0

For the effects of AMR-WB mode 1 on user perception such as illustrated in Figure 5.10, it is easy to see that user satisfaction increases when the mode of AMR-WB increases. This is fully suitable and logical. It is similar to the case of AMR-WB mode 0, the FLS scheduler still keeps the first position of highest MOS while the M-LWDF scheduler continues to keep the first position of the lowest MOS and the EXP/PF scheduler is in the middle. However, in this case, the FLS scheduler seems to slightly decrease when the NU increase. This is rather different from the case of AMR-WB mode 0. M-LWDF scheduler still trends to heavily decrease MOS score when the NU increases.

For AMR-WB anchor bitrate (i.e. AMR-WB mode 2) such as represented in Figure 5.11, it is also similar to the AMR-WB mode 0 and 1, in this case, the FLS scheduler still has the highest MOS score while the M-LWDF continues to have the lowest MOS score. It is really the difference among the schedulers is not too much, the most clear difference is the trends of them. For example, the FLS scheduler trends to slightly decrease, the EXP/PF scheduler

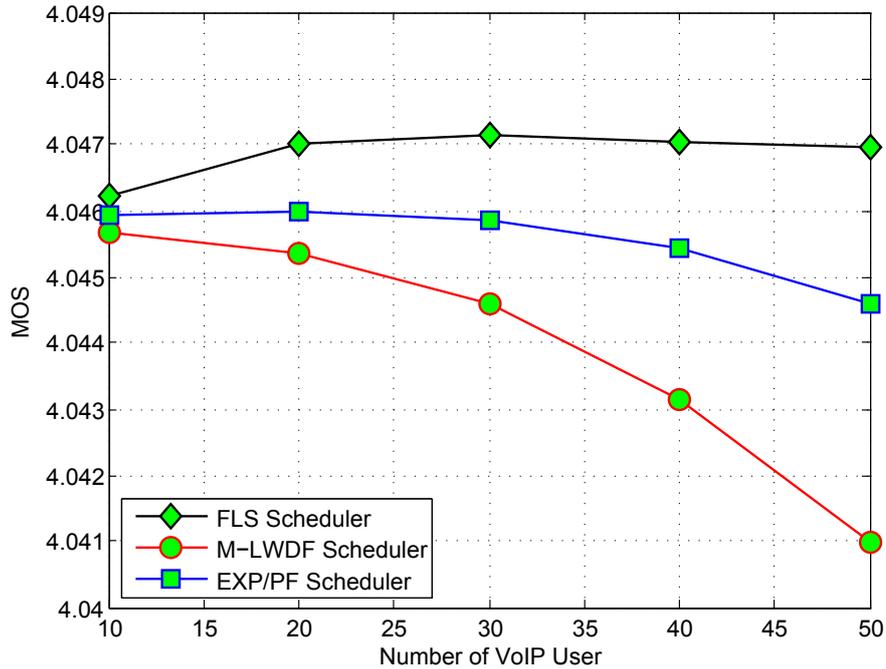


Figure 5.10: MOS vs number of VoIP user with AMR-WB mode 1

trends to averagely reduce while the M-LWDF scheduler trends to heavily decrease when the NU increases.

In order to evaluate the effects of AMR-WB modes on voice quality, we calculate the average MOS score for all modes such as represented in Figure 5.12. It is clear that, in general, the average voice quality decreases when the NU increases. The principles of the schedulers for AMR-WB mode 0, 1, and 2 are still suitable for this case. It can be concluded that FLS scheduler always has the best voice quality, M-LWDF scheduler has the worst one while EXP/PF is in the middle of two remaining schedulers.

5.3.3.2 Effects of delay on wideband audio quality

In order to assess simultaneously the effects of delay and the NU on voice quality, we present the relationship between the delay, the NU vs the MOS score such as shown in Figure 5.13. Both the M-LWDF and the EXP/PF schedulers have the average delay increasing when the NU increases, thus, MOS score decreases. This means the higher NU and the higher delay, the lower user perception. For the FLS scheduler, the delay decreases when the NU increases, thus, the MOS score slightly increases. This is rather special. Maybe it reflects the unstableness in a real system because the LTE-Sim is rather similar to a real system.

5.3.3.3 Effects of PLR on wideband audio quality

For the effects of PLR and the NU on voice quality such as shown in Figure 5.14. All the schedulers have the stable PLR when the NU increases. However, for the average MOS, the FLS scheduler has the best performance, the M-LWDF scheduler has the worst performance

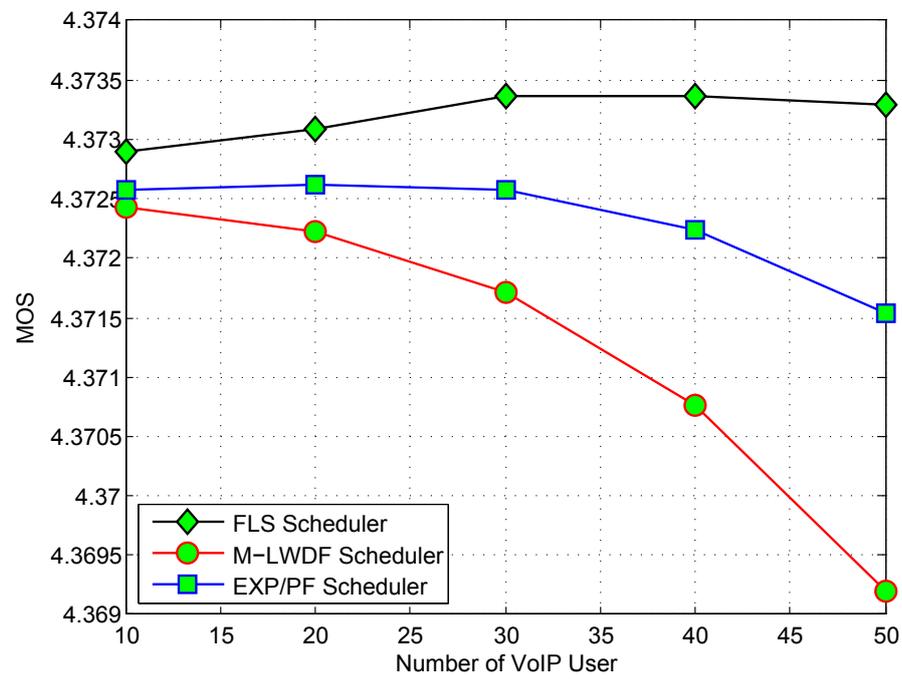


Figure 5.11: MOS vs number of VoIP user with AMR-WB mode 2

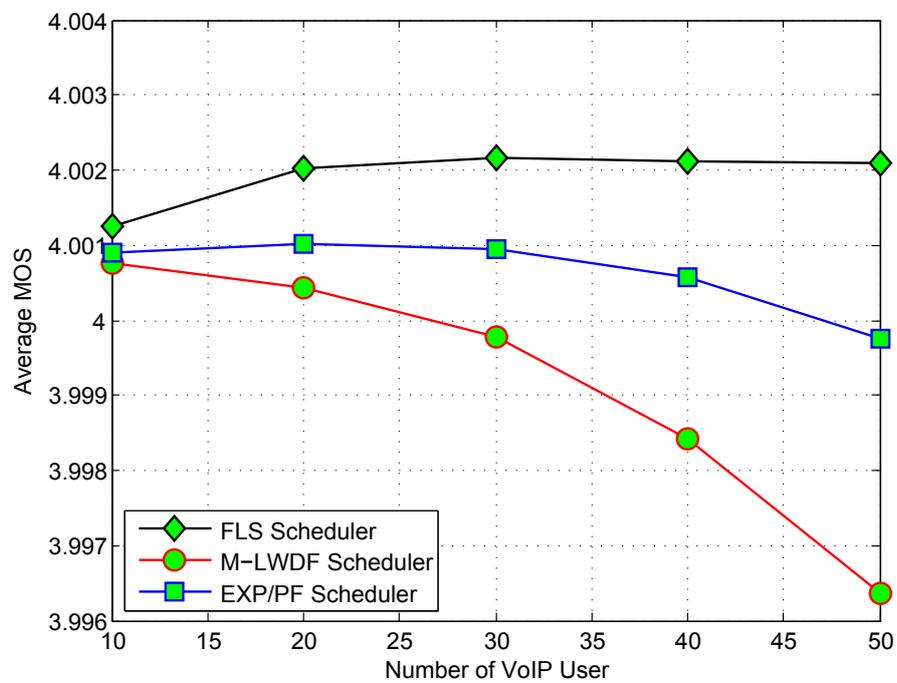


Figure 5.12: Average MOS vs the number of VoIP user

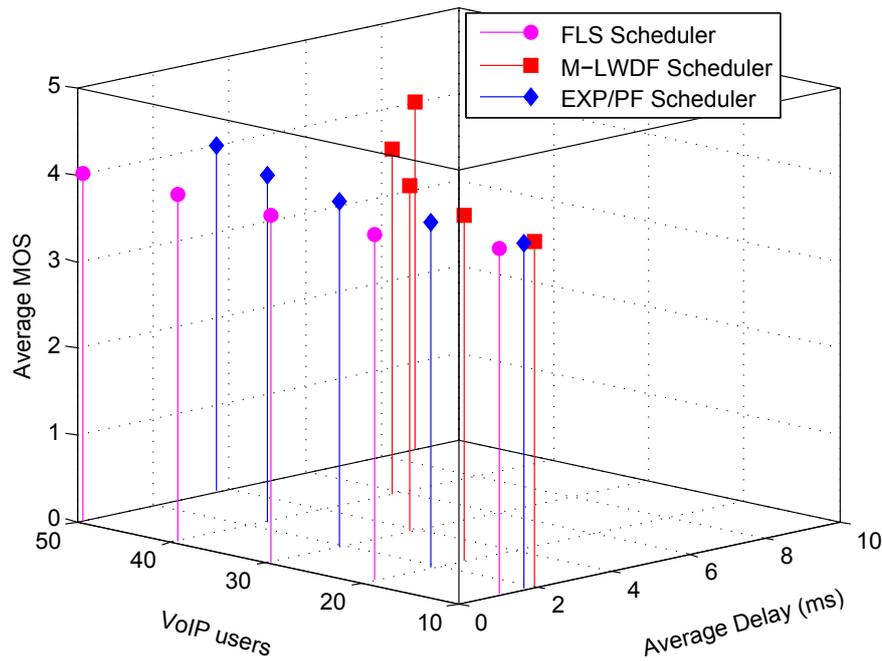


Figure 5.13: Effects of Delay and the number of VoIP user on voice quality

while the EXP/PF is in the middle of two remaining others.

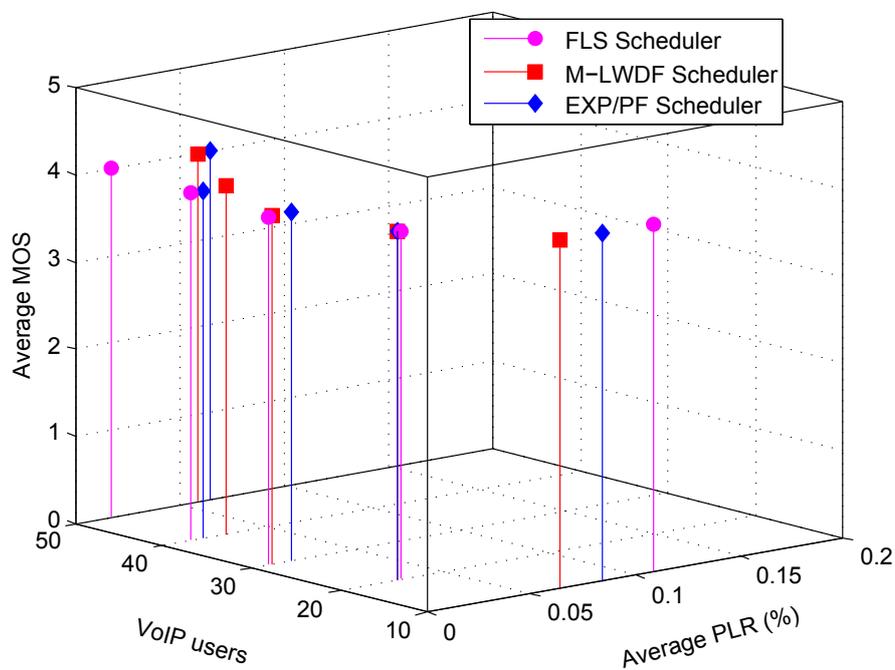


Figure 5.14: Effects of PLR and the number of VoIP user on voice quality

The details of the simulation results of effects of average delay and average PLR according

to the NU are shown in Table 5.3.

Table 5.3: The detailed results of Figures of 5.13 - 5.14

NU	FLS			M-LWDF			EXP/PF		
	Avg. Delay (ms)	Avg. PLR (%)	AVG. MOS	Avg. Delay (ms)	Avg. PLR (%)	AVG. MOS	Avg. Delay (ms)	Avg. PLR (%)	AVG. MOS
10	9.51	0.13	4.0012	1.95	0.066	4.0008	1.75	0.128	4.0009
20	8.83	0.012	4.002	2.5	0.058	4.0004	1.76	0.041	4.001
30	8.34	0.011	4.0022	3.43	0.008	3.9998	1.85	0.028	4.0009
40	8.04	0.015	4.0021	5.29	0.041	3.9984	2.3	0.002	4.0006
50	8.24	0.004	4.0021	8.02	0.048	3.9964	3.24	0.049	3.9997

5.4 Summary

In this Chapter, we propose new non-intrusive assessment models of voice quality over LTE network. The proposed method is the combination of the LTE-Sim framework and the extension of the E-model/the WB E-model for measuring narrowband/wideband audio quality, respectively. For the narrowband audio quality, the simulation results show that, the proposed model has the lower user perception in comparison with the E-model. This is because in the proposed model, we take effects of network jitter into account. The simulation results also show that the effect of JB on user perception is very significant. It's clear that user perception decreases when the user speed increases. For the case of user speed of 30 km/h, the user satisfaction slightly reduces when NU raises while it heavily decreases in case of the user speed of 120 km/h. For both cases of the user speed, the user perception increases when the JB increases. However, when JB is more than 100, user satisfaction increases not significantly. For the delay and PLR, both of them increase when NU increases and meet the principle: the higher delay or PLR, the lower user satisfaction.

For the wideband audio quality, we propose to complement the AMR-WB codec into the LTE-Sim software. The proposed model in this case can predict VoLTE quality not as exactly as intrusive methods (e.g. PESQ-WB) but it does not require original signal to refer, thus, it is very suitable for real-time services such as VoLTE. The advantages of the proposed model is that it is simple, calculate quickly, not much time consumption, and can be applied to many different simulation scenarios which can be configured in the LTE-Sim. It can be used well for purposes of transmission planning as well as voice quality assessment in laboratory for academic community and researchers. The simulation results also show that user perception is rather good when the MOS scores of all the cases are more than 3.5 which is a threshold for almost VoIP users accepted. The simulation results show that the FLS scheduler has the best user perception, the M-LWDF scheduler has the worst one while the EXP/PF scheduler is in the middle of two remaining ones. However, it can be seen that, the differences of user satisfaction among them are not significant. The user perception of those schedulers increases when the mode of AMR-WB codec increases. This is too logical because of the higher AMR-WB mode is, the higher MOS score is. The simulation results also show that the AMR-WB codec has much better voice quality than the G.729 codec. This is obvious because the AMR-WB codec has wider band than the G.729codec. This means the AMR-WB is very suitable for VoLTE service for providing HD (High-definition) video and voice services.

Chapter 6

Conclusions and Perspectives

Contents

6.1	Conclusions	113
6.2	Perspectives	117

6.1 Conclusions

In this thesis, we have studied issues that affects the voice transmission quality in LTE networks and have proposed several new solutions for enhancing LTE channel coding, improving downlink scheduling schemes, and predicting voice quality in LTE networks. LTE is designed for a high data rate and low latency network. LTE is based on an All-IP network, thus, voice service is also transmitted in an IP network. This is very different from a traditional mobile network such as 2G/3G in which voice service is communicated in a circuit-switched network. Since voice service in LTE is transmitted in a heterogeneous network which is based on IMS, the deployment of voice service is very complex, especially for guaranteeing QoS requirements. As mentioned in the Introduction, the challenges of VoLTE service deployment fall in three categories: Technology, Implementation and User satisfaction whereas the ones of VoLTE quality including: Call setup failures, end-to-end QoS, and cell-edge issues. In order to overcome these challenges, there needs to be good solutions of architecture, service, hardware, software of both the network infrastructure as well as user equipment. It is really ideal if there are overall solutions for deploying VoLTE service. This is really a big challenge. We see that, factors that affect strongly on VoLTE quality are the channel coding, scheduling schemes. Besides, in order to ensure user satisfaction, there needs to be models that allow predicting voice quality of VoLTE calls. Hence, in order to ensure VoLTE quality, the new solutions need to be proposed to solve the above challenges.

The proposed solutions in this thesis mainly focus on solving the issues of the group of User satisfaction challenges. Those soft solutions do not require any additional hardware component for eNB or UE. Therefore, they can be applied easily to improve the system performance without any change in hardware. The proposed solutions focus on some parts in LTE systems including channel coding, downlink scheduling schemes, and voice quality prediction. In order to enhance and improve voice transmission quality over LTE network, we need to be many different solutions in both hardware and software. In this thesis, the

proposed solutions are for channel coding and MAC scheduler which are very important elements and integral to LTE network. And to adjust network impairments at these layers, MNOs have to always monitor user perception instantly. For this reason, the new object non-intrusive models for measuring voice quality over LTE are proposed. Through the thesis, user perception which is determined by using the extended E-model as well as the WB E-model. They play an important role in our proposals. The contributions in this thesis are summarized as follows.

As the first contributions, we propose solutions for enhancing channel codec in LTE network such as described in Chapter 3. In particular, at sender side, we first propose an adaptive algorithm of joint source-channel code rate for minimizing the redundant bits generated by channel coding with a slight reduction of voice quality. For the deployment of this goal, we investigate the characteristics of VoLTE service such as source codec, radio protocol stack and Turbo codes which is used for coding data channels in LTE systems. The proposed algorithm is the adaptation of joint source-channel code rate. As described in the previous chapters, source code rate and channel code rate are adapted dynamically according to channel conditions. The selection of source code rate (AMR-WB) or channel code rate (Turbo code) strongly depends on channel quality that is usually performed as C/I ratio and is sent to the eNB by the feedback technique. The aim of this proposal is to jointly combine source code rate with channel code rate. We want to find out the trade-off between voice quality with minimizing the redundant bits generated by channel coding. This has to pay for a slight reduction of user perception. The proposed algorithm is programmed in the Matlab software with the static assumptions of LTE systems. The theoretical results show that with an adaptation of joint source-channel code rate, the redundant bits generated by channel coding can be reduced up to 50% with a slight reduction of voice quality is 1%. This solution can be combined with the selection of MCS to decide the final channel rate which is chosen from a pre-defined table in traditional manner. The efficiency of the proposed algorithm allow reducing required bandwidth of the system. This means more voice users can be served at the same time. The advantages of this solution is that it has simple operations, thus, it does not affect significantly on the system performance. In addition, the proposed solution can be set up in the system without any change in hardware. As mentioned before, when voice packets are coded by channel coder, they have to be decoded by channel decoder. For the decoding algorithms, the error correction function play a very important role because it decides the decoding quality performed in terms of BER performance. This means the lower BER is, the higher voice quality is. Decoding algorithms used in data channels in LTE are MAP-based algorithms including Log-MAP and Max-Log-MAP in which the Log-MAP has the best BER performance but the highest computational complexity whereas the Max-Log-MAP has the worst BER performance but the simplest computational complexity. For this reason, almost the previous proposed solutions focused on improving the Log-MAP algorithm to obtain a new algorithm having the trade-off between the BER performance and the computational complexity. Our second proposed solution is also focused on this goal. Specifically, we proposed an alternative error correction function used in the Log-MAP. In this function, we explore the understanding of polynomial regression function to find out the approximated function. The goal of this proposal is to replace the error correction function used in the Log-MAP algorithm with a new one which has the closest BER performance to the Log-MAP but the simpler computational complexity. Besides, we also investigate the performance of several featured Log-MAP based algorithms consisting of the Constant-Log-MAP, Linear Log-MAP, Non-linear Log-MAP, Max-Log-MAP and the Log-MAP. The

simulation model is implemented in Matlab software with some assumptions of LTE network parameters. We simulate and evaluate the Turbo coding and decoding with the conditions of BPSK modulation and AWGN channel. The numerical results show that the proposed decoding scheme outperforms all the others in terms of BER performance in the low SNR. This means it has the BER performance is closest to the one of the Log-MAP algorithm. The proposed error correction function does not require the lookup-table, it only includes the simple mathematical operations such as multiple and addition, thus, it has the low computational complexity. It also does not require any additional hardware components.

In case of enhancing LTE channel coding, the proposed solutions have the good performance and the low computational complexity without any change required for the system. However, the limitation is that it not simulated in a real environment such as the LTE-Sim, NS2/3, or OPNET softwares, etc. Therefore, they have not proved the high applicability in real LTE systems.

For the second contributions, we present the solutions for improving voice quality by enhancing downlink scheduling schemes such as represented in Chapter 4. The main idea is proceeded from the usage of E-model to measure voice quality such as described in Chapter 3. In these solutions, we propose to use the MQS and user perception which is determined by using the E-model and the WB E-model as the essential factors of the priority metric in LTE downlink packet schedulers. The proposed scheduling schemes are adapted to both narrowband and wideband audio services. Since all of them take into account the MQS and user perception for making the scheduling decision, they do not only meet QoS requirements for a real-time service such as VoLTE but also improve significantly voice quality. In particular, we first proposed a new LTE downlink scheduling scheme called E-MQS for narrowband audio flows. This scheme is designed for narrowband audio users. As mentioned above, the proposed scheduling scheme is based on the MQS and user perception. We proposed to use the MQS and user satisfaction as the essential factors of the priority metric. As explained in Chapter 4, in fact, the MQS is finite because if it is infinitive, the congestion becomes seriously and more packets drop due to the voice packets are timed out or exceed the delay threshold before they are scheduled. For this reason, the MQS should be present in the priority metric. In addition, in order to improve voice quality, more factors need to be complemented into the scheduling criterion. For this reason, user perception is also added to the priority metric. In order to determine user satisfaction, we use the E-model, and to enhance it, we complement the jitter factor into the E-model. In order to evaluate the performance of the proposed scheduling scheme, we compare it to several featured schedulers including M-LWDF, FLS, and EXP/PF in terms of end-to-end delay, PLR, cell throughput, FI and SE. The numerical results show that the proposed scheduler has better performance of delay, cell throughput, FI, and SE in comparison with M-LWDF, FLS as well as EXP/PF schedulers.

Second, we adapt the proposed scheduling scheme (E-MQS) to wideband audio users so called WE-MQS scheduler. As mentioned previously, most all our proposals are simulated in the LTE-Sim software, however, it only support for simulating narrowband audio flows using G.729 codec, thus, to simulate wideband audio flows, we proposed to add AMR-WB codec to the LTE-Sim. In addition, this codec is also mandatory for VoLTE. AMR-WB codec has 9 different bitrate, thus, we developed procedures that allow adapting its bitrate according to channel conditions. In order to predict user perception for wideband audio users, we have to use the WB E-model instead of the E-model. We also investigate the priority metrics of several schedulers such as FLS, EXP/PF, M-LWDF and compare their performance to the

proposed scheduler in terms of end-to-end delay, PLR, cell throughput, FI, and SE. The theoretical results prove that the proposed scheduler has better performance of delay, cell throughput, FI, and SE in comparison with M-LWDF, FLS as well as EXP/PF schedulers.

Third, due to VoLTE is deployed in an All-IP network, it needs to have the highest priority, especially when VoIP density is high. For this reason, we propose to integrate the VoIP-Priority mode into WE-MQS scheduler so called WE-MQS-VoIP Priority scheduler. In order to do that, we must modify the approach of radio resource allocation as described in Chapter 4. The VoIP-Priority mode is only adapted when there is VoIP user in the buffer. In addition, to decrease the negative effects of VoIP-Priority mode, the procedure of the duration of VoIP-Priority mode is used.

In order to assess the performance of the proposed scheduling scheme, we compare its performance to the M-LWDF and WE-MQS schedulers in terms of end-to-end delay, PLR, cell throughput, FI, and SE. The simulation results show that the proposed scheduler overcomes the other schedulers in terms of delay, PLR, cell throughput, FI and SE. It can be concluded that, WE-MQS-VoIP Priority scheduler is the best and the most suitable for VoLTE service.

For the last contributions, in order to ensure voice users always satisfy, VoLTE quality has to be evaluated automatically and instantly such as represented in Chapter 5. This allows adjusting network impairments for guaranteeing QoS requires for VoLTE services. The proposed models are based on the combination between the LTE-Sim software and the E-model or the WB E-model for predicting voice quality of narrowband audio or wideband audio users, respectively. This is executed based on the proposals in Chapter 4. The principles of the proposed models are carried out in two steps: (1) We use the LTE-Sim software to simulate VoLTE services, then we collect the average end-to-end delay and the average PLR according to the number of voice users from the outputs of the simulations. (2) We use these parameters as the essential inputs of the E-model or WB E-model. The output of extended E-model or WB E-model is user perception. In order validate the proposed models, we use the well-known schedulers such as M-LWDF, FLS, and EXP/PF with the mobility in LTE heterogeneous network. The numerical results show that in case of narrowband audio users, when taking into account the jitter factor, the extended E-model has the lower MOS score than the E-model with the same scheduler of M-LWDF. This means the extended E-model displays the result that is closer to human perception. In case of wideband audio flows, the performance is evaluated for the M-LWDF, FLS, and EXP/PF scheduler. For the same simulation scenario, the FLS always has the highest MOS score whereas the EXP/PF scheduler has the lowest MOS score. This means the M-LWDF scheduler is in the middle of the FLS and EXP/PF schedulers. This is fully suitable when referring to the simulation results shown in Chapter 4, the FLS scheduler always has the lowest delay and PLR while the M-LWDF scheduler has the higher delay and PLR than the EXP/PF scheduler for voice users. It is clear that, the end to end delay and PLR are two factors that affect directly the R_{wb} -factor in the WB E-model. It can be seen that both the proposed models based on the object non-intrusive method. This means they do not refer to original signal, thus, they are very suitable for real-time services such as VoLTE. They have the simple operations so they can be applied to measure instantly quality for live voice stream. Besides, they also do not require any change of the hardware system. They also very suitable for researchers in the laboratories for estimating voice quality in LTE systems with the different simulation scenarios configured in the LTE-Sim software.

It can be seen that, through all the proposed solutions in this thesis, user perception

which is determined using the E-model or the WB E-model is the backbone. All the proposed solutions focus on enhancing, improving and predicting voice transmission quality in LTE networks. The latter is proceeded from the former and they support each other. Although, the E-model was initially designed for purposes of transmission planning. However, the simulation results show that if we can determine exactly the input parameters, it can be applied in many different cases to predict user perception and it can be applied in the different layers in LTE systems for optimization issues.

6.2 Perspectives

In this thesis, we have presented the new solutions which aim at solving the problems that affect voice transmission quality in LTE downlink networks as shown in Chapters 3, 4, and 5. For each issue, we have proposed the new methods and have evaluated the performance of them. We have shown the advantages and disadvantages of them and the applicability in LTE systems. Besides the obtained results, the proposed solutions still have limitations. This brings new directions to optimize them for improving voice transmission quality in LTE networks.

1. In case of the first contributions, there are some limitations such as the simulations are assessed with static assumptions in the Matlab software. The performance evaluation in real simulation environments such as the LTE-Sim, OPNET, or NS2/3 is necessary. In addition, the combination of the algorithm of joint source-channel code rate adaptation and the improved Log-MAP algorithm based on the polynomial regression function in the same simulation scenario is essential to demonstrate the comprehensive efficiency of them on the enhancement of LTE channel coding. The algorithm of dynamic source-channel code rate adaptation is also studied according to the direction of minimization of the required bandwidth. The evaluation of the computational complexity of the improved Log-MAP algorithm based on the polynomial regression function needs to be considered.
2. In case of the second contributions, the proposed scheduling schemes need to be evaluated in many different simulation scenarios, and with more voice source codecs. In order to predict user perception more exactly, more network impairments need to be complemented into the E-model as well as the WB E-model. The proposed schedulers are also modified to adapt to LTE uplink direction.
3. In case of the last contributions, in order to measure voice quality more exactly, more network impairments are also added to the E-model as well as the WB E-model. The proposed models are also verified in many different simulation scenarios. Develop the proposed models as a monitoring tool at Application layer to predict instantly live voice stream quality.
4. An overall framework of the proposed solutions for enhancing voice transmission in LTE networks also needs to be considered.
5. Finally, upgrade all the proposed solutions to the context of next generation such as LTE-Advanced, 5G, and they need to be evaluated in LTE femtocell network, etc.

Bibliography

- [1] T. G. mobile Suppliers Association (GSA), “Evolution to lte report: 4g market and technology update,” Tech. Rep., 4 2015. [Online]. Available: http://www.gsacom.com/gsm_3g/info_papers.php4
- [2] G. Hosangadi, A. Rao, R. Bartlett, and M. Schlesener, “Carrier-grade voice over lte: Advantages and challenges,” in *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*. IEEE, 2014, pp. 1–5.
- [3] L. Gupta, “Voice over lte: Status and migration trends.”
- [4] D.-H. Nguyen and H. Nguyen, “An improved log-map algorithm based on polynomial regression function for lte turbo decoding,” in *2015 IEEE International Conference on Communication Workshop (ICCW)*. IEEE, 2015, pp. 2163–2167.
- [5] —, “A dynamic rate adaptation algorithm using wb e-model for voice traffic over lte network,” in *2016 Wireless Days (WD)*. IEEE, March 2016, pp. 1–6.
- [6] D.-H. Nguyen, H. Nguyen, and E. Renault, “A new channel- and qos-aware scheduling scheme for real-time services in lte network,” *International Journal of Applied Information Systems*, vol. 11, no. 4, pp. 1–8, Sep 2016. [Online]. Available: <http://www.ijais.org/archives/volume11/number4/933-2016451601>
- [7] H. N. Duy-Huy Nguyen and E. Renault, “E-mqs - a new downlink scheduler for real-time flows in lte network,” in *Vehicular Technology Conference (VTC Spring), 2016 IEEE 84th*. IEEE, September 2016.
- [8] —, “We-mqs: A new lte downlink scheduling scheme for voice services based on user perception,” *International Journal of Computer Applications*, vol. 142, no. 10, pp. 28–36, May 2016. [Online]. Available: <http://www.ijcaonline.org/archives/volume142/number10/24934-2016909948>
- [9] D.-H. Nguyen, H. Nguyen, and E. Renault, “We-mqs-voip priority: An enhanced lte downlink scheduler for voice services with the integration of voip priority mode,” *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 7, no. 7, pp. 560–567, Jul 2016. [Online]. Available: <http://thesai.org/Publications/ViewPaper?Volume=7&Issue=7&Code=IJACSA&SerialNo=76#sthash.IGIZTvc.dpuf>

- [10] H. N. Duy-Huy Nguyen and E. Renault, "A new non-intrusive assessment method for volte quality based on extended e-model," in *Mobile, Secure, and Programmable Networking: Second International Conference, MSPN 2016, Paris, France, June 01-03, 2016, Selected Papers*. Springer, 2016.
- [11] D.-H. Nguyen, H. Nguyen, and E. Renault, "A new non-intrusive model for measuring volte quality based on wideband e-model," *Communications on Applied Electronics*, vol. 5, no. 8, pp. 1–8, Aug 2016. [Online]. Available: <http://www.caeaccess.org/archives/volume5/number8/636-2016652333>
- [12] N. Saed, T. H. Liew, K. Wee, and S. Y. Ooi, "Video and voice transmission over lte networks," in *Control System, Computing and Engineering (ICCSCE), 2013 IEEE International Conference on*. IEEE, 2013, pp. 259–263.
- [13] H. Ekstrom, "Qos control in the 3gpp evolved packet system," *Communications Magazine, IEEE*, vol. 47, no. 2, pp. 76–83, 2009.
- [14] I. Rec, "G. 722.2 (2003) wideband coding of speech at around 16kbit/s using adaptive multi-rate wideband (amr-wb)," *International telecommunication union, Geneva, Switzerland*, 2003.
- [15] C. Cox, *An introduction to LTE: LTE, LTE-advanced, SAE and 4G mobile communications*. John Wiley & Sons, 2012.
- [16] 3GPP, "Inband Tandem Free Operation (TFO) of speech codecs; Service description; Stage 3," 3rd Generation Partnership Project (3GPP), TS 28.062, 12 2009. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/28062.htm>
- [17] G. Piro, L. A. Grieco, G. Boggia, F. Capozzi, and P. Camarda, "Simulating lte cellular systems: an open-source framework," *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 2, pp. 498–513, 2011.
- [18] 3GPP, "Feasibility study for Orthogonal Frequency Division Multiplexing (OFDM) for UTRAN enhancement," 3rd Generation Partnership Project (3GPP), TR 25.892, 06 2004. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/25892.htm>
- [19] —, "Physical layer aspect for evolved Universal Terrestrial Radio Access (UTRA)," 3rd Generation Partnership Project (3GPP), TR 25.814, 10 2006. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/25814.htm>
- [20] B. T. Scheme, "Lte: the evolution of mobile broadband," *IEEE Communications magazine*, vol. 45, 2009.
- [21] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures," 3rd Generation Partnership Project (3GPP), TS 36.213, 10 2010. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/36213.htm>
- [22] S.-E. Elayoubi, O. B. Haddada, and B. Fourestie, "Performance evaluation of frequency planning schemes in ofdma-based networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, pp. 1623–1633, 2008.

- [23] D. Martín-Sacristán, J. Cabrejas, D. Calabuig, and J. F. Monserrat, “Mac layer performance of different channel estimation techniques in utran lte downlink,” in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*. IEEE, 2009, pp. 1–5.
- [24] 3GPP, “Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification,” 3rd Generation Partnership Project (3GPP), TS 36.321, 06 2010. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/36321.htm>
- [25] —, “Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding,” 3rd Generation Partnership Project (3GPP), TS 36.212, Sep. 2011. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/36212.htm>
- [26] C. Berrou and A. Glavieux, “Near optimum error correcting coding and decoding: Turbo-codes,” *Communications, IEEE Transactions on*, vol. 44, no. 10, pp. 1261–1271, 1996.
- [27] 3GPP, “Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures,” 3rd Generation Partnership Project (3GPP), TS 36.213, 10 2010. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/36213.htm>
- [28] A. Ghosh and R. Ratasuk, *Essentials of LTE and LTE-A*. Cambridge University Press, 2011.
- [29] C. Studer, C. Benkeser, S. Belfanti, and Q. Huang, “Design and implementation of a parallel turbo-decoder asic for 3gpp-lte,” *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 8–17, 2011.
- [30] S. Zarei, “Channel coding and link adaptation,” *Ausgewhlte Kapitel der Nachrichtentechnik, WS*, vol. 2010, pp. 1–14, 2009.
- [31] ITU-T, “Itu-t recommendation g.107: The e-model, a computational model for use in transmission planning,” International Telecommunication Union, Tech. Rep., Dec. 1998.
- [32] I. Rec, “G. 107.1, wideband e-model,” *Int. Telecomm. Union, Geneva*, 2011.
- [33] S. Moller, A. Raake, N. Kitawaki, A. Takahashi, and M. Waltermann, “Impairment factor framework for wide-band speech codecs,” *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 14, no. 6, pp. 1969–1976, 2006.
- [34] C.-L. Chi and C.-H. Kuo, “Quadratic permutation polynomial interleaver for lte turbo coding,” in *Information Security and Intelligence Control (ISIC), 2012 International Conference on*. IEEE, 2012, pp. 313–316.
- [35] A. Nimbalkar, Y. Blankenship, B. Classon, and T. K. Blankenship, “Arp and qpp interleavers for lte turbo coding,” in *2008 IEEE Wireless Communications and Networking Conference*. IEEE, 2008, pp. 1032–1037.
- [36] J.-F. Cheng, A. Nimbalkar, Y. Blankenship, B. Classon, and T. K. Blankenship, “Analysis of circular buffer rate matching for lte turbo code,” in *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*. IEEE, 2008, pp. 1–5.

- [37] T. Breddermann and P. Vary, "Rate-compatible insertion convolutional turbo codes: Analysis and application to lte," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1356–1366, 2014.
- [38] C. Ma and P. Lin, "Efficient implementation of quadratic permutation polynomial interleaver in turbo codes," in *Wireless Communications & Signal Processing, 2009. WCSP 2009. International Conference on*. IEEE, 2009, pp. 1–5.
- [39] A. Ardakani, M. Mahdavi, and M. Shabany, "An efficient vlsi architecture of qpp interleaver/deinterleaver for lte turbo coding," in *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*. IEEE, 2013, pp. 797–800.
- [40] L. Yu, X. Wang, and J. Liu, "An improved rate matching algorithm for 3gpp lte turbo code," in *Communications and Mobile Computing (CMC), 2011 Third International Conference on*. IEEE, 2011, pp. 345–348.
- [41] A. A. Purwita, A. Setio, and T. Adiono, "Optimized 8-level turbo encoder algorithm and vlsi architecture for lte," in *Electrical Engineering and Informatics (ICEEI), 2011 International Conference on*. IEEE, 2011, pp. 1–6.
- [42] A. Akmalkhodzhaev, "Joint decoding of turbo code and amr-wb vocoder in 3gpp lte system," in *2014 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. IEEE, 2014, pp. 402–406.
- [43] T. Adiono *et al.*, "Radix-4 max-log-map parallel turbo decoder architecture with a new cache memory data flow for lte," in *Intelligent Signal Processing and Communications Systems (ISPACS), 2012 International Symposium on*. IEEE, 2012, pp. 792–797.
- [44] K. Zhang, R. Yao, F. Gao, L. Wang, and J. Zhang, "Integrated implementation of fixed-point turbo decoders for lte system," in *2013 22nd Wireless and Optical Communication Conference*. IEEE, 2013, pp. 187–192.
- [45] L. Li, R. G. Maunder, B. M. Al-Hashimi, and L. Hanzo, "A low-complexity turbo decoder architecture for energy-efficient wireless sensor networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 1, pp. 14–22, 2013.
- [46] R. Shrestha and R. P. Paily, "High-throughput turbo decoder with parallel architecture for lte wireless communication standards," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 9, pp. 2699–2710, 2014.
- [47] C.-C. Wong, Y.-Y. Lee, and H.-C. Chang, "A 188-size 2.1 mm² reconfigurable turbo decoder chip with parallel architecture for 3gpp lte system," in *2009 Symposium on VLSI Circuits*. IEEE, 2009, pp. 288–289.
- [48] J.-H. Kim and I.-C. Park, "A unified parallel radix-4 turbo decoder for mobile wimax and 3gpp-lte," in *2009 IEEE Custom Integrated Circuits Conference*. IEEE, 2009, pp. 487–490.
- [49] M. May, T. Inseher, N. Wehn, and W. Raab, "A 150mbit/s 3gpp lte turbo code decoder," in *Proceedings of the Conference on Design, Automation and Test in Europe*. European Design and Automation Association, 2010, pp. 1420–1425.

- [50] T. Ilseher, M. May, and N. Wehn, "A multi-mode 3gpp-lte/hsdpa turbo decoder," in *Communication Systems (ICCS), 2010 IEEE International Conference on*. IEEE, 2010, pp. 336–340.
- [51] Y. Sun and J. R. Cavallaro, "Efficient hardware implementation of a highly-parallel 3gpp lte/lte-advance turbo decoder," *INTEGRATION, the VLSI journal*, vol. 44, no. 4, pp. 305–315, 2011.
- [52] D. Wu, R. Asghar, Y. Huang, and D. Liu, "Implementation of a high-speed parallel turbo decoder for 3gpp lte terminals," in *2009 IEEE 8th International Conference on ASIC*. IEEE, 2009, pp. 481–484.
- [53] C.-C. Wong and H.-C. Chang, "Reconfigurable turbo decoder with parallel architecture for 3gpp lte system," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 57, no. 7, pp. 566–570, 2010.
- [54] T. Ilseher, F. Kienle, C. Weis, and N. Wehn, "A 2.15 gbit/s turbo code decoder for lte advanced base station applications," in *Turbo Codes and Iterative Information Processing (ISTC), 2012 7th International Symposium on*. IEEE, 2012, pp. 21–25.
- [55] W. Huahua and L. Wenwen, "Analysis of turbo decoding algorithm in lte system," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*. IEEE, 2012, pp. 1741–1744.
- [56] J.-H. Kim and I.-C. Park, "Bit-level extrinsic information exchange method for double-binary turbo codes," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 56, no. 1, pp. 81–85, 2009.
- [57] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal map decoding algorithms operating in the log domain," in *Communications, 1995. ICC'95 Seattle, 'Gateway to Globalization', 1995 IEEE International Conference on*, vol. 2. IEEE, 1995, pp. 1009–1013.
- [58] J. Hagenauer and P. Hoeher, "A viterbi algorithm with soft-decision outputs and its applications," in *Global Telecommunications Conference and Exhibition 'Communications Technology for the 1990s and Beyond' (GLOBECOM), 1989. IEEE*. IEEE, 1989, pp. 1680–1686.
- [59] H. Wang, H. Yang, and D. Yang, "Improved log-map decoding algorithm for turbo-like codes," *Communications Letters, IEEE*, vol. 10, no. 3, pp. 186–188, 2006.
- [60] W. J. Gross and P. G. Gulak, "Simplified map algorithm suitable for implementation of turbo decoders," *Electronics Letters*, vol. 34, no. 16, pp. 1577–1578, 1998.
- [61] S. Papaharalabos, P. Sweeney, and B. Evans, "Constant log-map decoding algorithm for duo-binary turbo codes," *Electronics Letters*, vol. 42, no. 12, pp. 709–710, 2006.
- [62] J.-F. Cheng and T. Ottosson, "Linearly approximated log-map algorithms for turbo decoding," in *Vehicular Technology Conference Proceedings, 2000. VTC 2000-Spring Tokyo. 2000 IEEE 51st*, vol. 3. IEEE, 2000, pp. 2252–2256.

- [63] S. Talakoub, L. Sabeti, B. Shahrrava, and M. Ahmadi, "An improved max-log-map algorithm for turbo decoding and turbo equalization," *Instrumentation and Measurement, IEEE Transactions on*, vol. 56, no. 3, pp. 1058–1063, 2007.
- [64] S. Ali and M. Zeeshan, "A utility based resource allocation scheme with delay scheduler for lte service-class support," in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. IEEE, 2012, pp. 1450–1455.
- [65] M. Alasti, B. Neekzad, J. Hui, and R. Vannithamby, "Quality of service in wimax and lte networks [topics in wireless communications]," *Communications Magazine, IEEE*, vol. 48, no. 5, pp. 104–111, 2010.
- [66] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in lte cellular networks: Key design issues and a survey," *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 2, pp. 678–700, 2013.
- [67] D. Singh and P. Singh, "Radio resource scheduling in 3gpp lte: a review," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 4, no. 6, pp. 2405–2411, 2013.
- [68] A. S. Tanenbaum and H. Bos, *Modern operating systems*. Prentice Hall Press, 2014.
- [69] S. Hussain, "Dynamic radio resource management in 3gpp lte," Ph.D. dissertation, Blekinge Institute of Technology, 2009.
- [70] P. Kela, J. Puttonen, N. Kolehmainen, T. Ristaniemi, T. Henttonen, and M. Moisio, "Dynamic packet scheduling performance in ultra long term evolution downlink," in *Wireless Pervasive Computing, 2008. ISWPC 2008. 3rd International Symposium on*. IEEE, 2008, pp. 308–313.
- [71] A. Demers, S. Keshav, and S. Shenker, "Analysis and simulation of a fair queueing algorithm," in *ACM SIGCOMM Computer Communication Review*, vol. 19, no. 4. ACM, 1989, pp. 1–12.
- [72] A. L. Stolyar and K. Ramanan, "Largest weighted delay first scheduling: Large deviations and optimality," *Annals of Applied Probability*, pp. 1–48, 2001.
- [73] R. K. Almatarneh, M. H. Ahmed, and O. A. Dobre, "Performance analysis of proportional fair scheduling in ofdma wireless systems," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*. IEEE, 2010, pp. 1–5.
- [74] R. Kwan, C. Leung, and J. Zhang, "Proportional fair multiuser scheduling in lte," *IEEE Signal Processing Letters*, vol. 16, no. 6, pp. 461–464, 2009.
- [75] K. W. Choi, W. S. Jeon, and D. G. Jeong, "Resource allocation in ofdma wireless communications systems supporting multimedia services," *IEEE/ACM transactions on networking*, vol. 17, no. 3, pp. 926–935, 2009.
- [76] A. Pokhariyal, K. I. Pedersen, G. Monghal, I. Z. Kovacs, C. Rosa, T. E. Kolding, and P. E. Mogensen, "Harq aware frequency domain packet scheduler with different degrees of fairness for the utran long term evolution," in *2007 IEEE 65th Vehicular Technology Conference-VTC2007-Spring*. IEEE, 2007, pp. 2761–2765.

- [77] J. Huang and Z. Niu, "Buffer-aware and traffic-dependent packet scheduling in wireless ofdm networks," in *2007 IEEE Wireless Communications and Networking Conference*. IEEE, 2007, pp. 1554–1558.
- [78] G. Monghal, K. I. Pedersen, I. Z. Kovacs, and P. E. Mogensen, "Qos oriented time and frequency domain packet schedulers for the utran long term evolution," in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*. IEEE, 2008, pp. 2532–2536.
- [79] D. N. Skoutas and A. N. Rouskas, "Scheduling with qos provisioning in mobile broadband wireless systems," in *Wireless Conference (EW), 2010 European*. IEEE, 2010, pp. 422–428.
- [80] Y. Zaki, T. Weerawardane, C. Gorg, and A. Timm-Giel, "Multi-qos-aware fair scheduling for lte," in *Vehicular technology conference (VTC spring), 2011 IEEE 73rd*. IEEE, 2011, pp. 1–5.
- [81] H. A. M. Ramli, R. Basukala, K. Sandrasegaran, and R. Patachaianand, "Performance of well known packet scheduling algorithms in the downlink 3gpp lte system," in *Communications (MICC), 2009 IEEE 9th Malaysia International Conference on*. IEEE, 2009, pp. 815–820.
- [82] R. Basukala, H. M. Ramli, and K. Sandrasegaran, "Performance analysis of exp/pf and m-lwdf in downlink 3gpp lte system," in *2009 First Asian Himalayas International Conference on Internet*. IEEE, 2009, pp. 1–5.
- [83] B. Sadiq, R. Madan, and A. Sampath, "Downlink scheduling for multiclass traffic in lte," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, p. 1, 2009.
- [84] P. Chandur, R. Karthik, and K. M. Sivalingam, "Performance evaluation of scheduling algorithms for mobile wimax networks," in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*. IEEE, 2012, pp. 764–769.
- [85] M. Iturralde, A. Wei, T. A. Yahiya, and A.-L. Beylot, "Resource allocation for real time services using cooperative game theory and a virtual token mechanism in lte networks," in *2012 IEEE Consumer Communications and Networking Conference (CCNC)*. IEEE, 2012, pp. 879–883.
- [86] B. Sadiq, S. J. Baek, and G. De Veciana, "Delay-optimal opportunistic scheduling and approximations: The log rule," *IEEE/ACM Transactions on Networking (TON)*, vol. 19, no. 2, pp. 405–418, 2011.
- [87] G. Piro, L. A. Grieco, G. Boggia, R. Fortuna, and P. Camarda, "Two-level downlink scheduling for real-time multimedia services in lte networks," *IEEE Transactions on Multimedia*, vol. 13, no. 5, pp. 1052–1065, 2011.
- [88] G. Monghal, D. Laselva, P.-H. Michaelsen, and J. Wigard, "Dynamic packet scheduling for traffic mixes of best effort and voip users in e-utran downlink," in *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*. IEEE, 2010, pp. 1–5.

- [89] Y. Fan, P. Lundén, M. Kuusela, and M. Valkama, “Efficient semi-persistent scheduling for voip on eutra downlink,” in *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th.* IEEE, 2008, pp. 1–5.
- [90] Y.-S. Kim, “An efficient scheduling scheme to enhance the capacity of voip services in evolved ultra uplink,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, no. 1, p. 1, 2008.
- [91] D. Sabella, M. Caretti, and R. Fantini, “Energy efficiency evaluation of state of the art packet scheduling algorithms for lte,” in *Wireless Conference 2011-Sustainable Wireless Technologies (European Wireless), 11th European.* VDE, 2011, pp. 1–4.
- [92] S. Videv and H. Haas, “Energy-efficient scheduling and bandwidth-energy efficiency trade-off with low load,” in *2011 IEEE International Conference on Communications (ICC).* IEEE, 2011, pp. 1–5.
- [93] ITU-T, “Itu-t recommendation p.800: Methods for subjective determination of transmission quality,” International Telecommunication Union, Tech. Rep., Aug. 1996.
- [94] —, “Itu-t recommendation p.862: Perceptual evaluation of speech quality (pesq): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” International Telecommunication Union, Tech. Rep., Feb. 2001.
- [95] —, “Itu-t recommendation p.862.2: Wideband extension to recommendation p.862 for the assessment of wideband telephone networks and speech codecs,” International Telecommunication Union, Tech. Rep., Nov. 2007.
- [96] —, “Itu-t recommendation g.107.1: Wideband e-model,” International Telecommunication Union, Tech. Rep., Jun. 2015.
- [97] K. Radhakrishnan, H. Larijani, and T. Buggy, “A non-intrusive method to assess voice quality over internet,” in *Performance Evaluation of Computer and Telecommunication Systems (SPECTS), 2010 International Symposium on.* IEEE, 2010, pp. 380–386.
- [98] H. Larijani and K. Radhakrishnan, “Voice quality in voip networks based on random neural networks,” in *Networks (ICN), 2010 Ninth International Conference on.* IEEE, 2010, pp. 89–92.
- [99] T. Ghalut and H. Larijani, “Non-intrusive method for video quality prediction over lte using random neural networks (rnn),” in *Communication Systems, Networks & Digital Signal Processing (CSNDSP), 2014 9th International Symposium on.* IEEE, 2014, pp. 519–524.
- [100] L. Ding and R. A. Goubran, “Speech quality prediction in voip using the extended e-model,” in *Global Telecommunications Conference, 2003. GLOBECOM’03. IEEE*, vol. 7. IEEE, 2003, pp. 3974–3978.
- [101] A. Meddahi, H. Afif, and D. Zeghlache, ““ packet-e-model”: e-model for wireless voip quality evaluation,” in *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*, vol. 3. IEEE, 2003, pp. 2421–2425.

- [102] L. Carvalho, E. Mota, R. Aguiar, A. F. Lima, A. Barreto *et al.*, “An e-model implementation for speech quality evaluation in voip systems,” in *null*. IEEE, 2005, pp. 933–938.
- [103] A. Takahashi, A. Kurashima, and H. Yoshino, “Objective assessment methodology for estimating conversational quality in voip,” *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 14, no. 6, pp. 1984–1993, 2006.
- [104] H. Assem, D. Malone, J. Dunne, and P. O’Sullivan, “Monitoring voip call quality using improved simplified e-model,” in *Computing, networking and communications (ICNC), 2013 international conference on*. IEEE, 2013, pp. 927–931.
- [105] ITU-T, “Itu-t recommendation g.107: The e-model, a computational model for use in transmission planning,” International Telecommunication Union, Tech. Rep., Jun. 2015.
- [106] K. C. Mishra and P. C. Das, “An overview of quality of service measurement and optimization for voice over internet protocol implementation,” *International Journal of Advances in Engineering & Technology*, vol. 8, no. 1, p. 2053, 2015.
- [107] M. Abareghi, M. M. Homayounpour, M. Dehghan, and A. Davoodi, “Improved itu-p. 563 non-intrusive speech quality assessment method for covering voip conditions,” in *Advanced Communication Technology, 2008. ICACT 2008. 10th International Conference on*, vol. 1. IEEE, 2008, pp. 354–357.
- [108] D.-S. Kim and A. Tarraf, “Anique+: A new american national standard for non-intrusive estimation of narrowband speech quality,” *Bell Labs Technical Journal*, vol. 12, no. 1, pp. 221–236, 2007.
- [109] ITU-T, “Itu-t recommendation p.564: Conformance testing for voice over ip transmission quality assessment models,” International Telecommunication Union, Tech. Rep., Nov. 2007.
- [110] A. D. Clark, P. D. F. Iee *et al.*, “Modeling the effects of burst packet loss and recency on subjective voice quality,” 2001.
- [111] S. Broom and M. Hollier, “Speech quality measurement tools for dynamic network management,” *Measurement of Speech and Audio Quality in Networks (MESAQIN)*, 2003.
- [112] T. H. Falk and W.-Y. Chan, “Nonintrusive speech quality estimation using gaussian mixture models,” *IEEE Signal Processing Letters*, vol. 13, no. 2, pp. 108–111, 2006.
- [113] J. Wang, J. Luo, S. Zhao, and J. Kuang, “Non-intrusive objective speech quality measurement based on gmm and svr for narrowband and wideband speech,” in *Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference on*. IEEE, 2008, pp. 193–198.
- [114] J. Wang, Y. Zhang, Y. Song, S. Zhao, and J. Kuang, “An improved non-intrusive objective speech quality evaluation based on fgmm and fnn,” in *Image and Signal Processing (CISP), 2010 3rd International Congress on*, vol. 7. IEEE, 2010, pp. 3495–3499.

- [115] L. Sun and E. C. Ifeachor, "Perceived speech quality prediction for voice over ip-based networks," in *Communications, 2002. ICC 2002. IEEE International Conference on*, vol. 4. IEEE, 2002, pp. 2573–2577.
- [116] S. Mohamed, G. Rubino, and M. Varela, "Performance evaluation of real-time speech through a packet network: a random neural networks-based approach," *Performance evaluation*, vol. 57, no. 2, pp. 141–161, 2004.
- [117] —, "A method for quantitative evaluation of audio quality over packet networks and its comparison with existing techniques," *Measurement of speech and audio quality in networks (MESAQIN)*, no. 2, p. 0, 2004.
- [118] A. P. C. da Silva, M. Varela, E. d. S. e Silva, R. M. Leão, and G. Rubino, "Quality assessment of interactive voice applications," *Computer Networks*, vol. 52, no. 6, pp. 1179–1192, 2008.
- [119] W. Cherif, A. Ksentini, D. Négru, and M. Sidibe, "A_psq: Pesq-like non-intrusive tool for qoe prediction in voip services," in *2012 IEEE International Conference on Communications (ICC)*. IEEE, 2012, pp. 2124–2128.
- [120] T. Hößfeld, P. Tran-Gia, and M. Fiedler, "Quantification of quality of experience for edge-based applications," in *Managing Traffic Performance in Converged Networks*. Springer, 2007, pp. 361–373.
- [121] T. Hößfeld, D. Hock, P. Tran-Gia, K. Tutschku, and M. Fiedler, "Testing the iqx hypothesis for exponential interdependency between qos and qoe of voice codecs ilbc and g. 711," in *18th ITC Specialist Seminar Quality of Experience, Karlskrona, Sweden*, 2008, pp. 105–114.
- [122] Y. Jung and C. Manzano, "Burst packet loss and enhanced packet loss-based quality model for mobile voice-over internet protocol applications," *IET Communications*, vol. 8, no. 1, pp. 41–49, 2014.
- [123] S. Jelassi, H. Youssef, C. Hoene, and G. Pujolle, "Voicing-aware parametric speech quality models over voip networks," in *2009 Global Information Infrastructure Symposium*. IEEE, 2009, pp. 1–8.
- [124] L. Ding, Z. Lin, A. Radwan, M. S. El-Hennawey, and R. A. Goubran, "Non-intrusive single-ended speech quality assessment in voip," *Speech Communication*, vol. 49, no. 6, pp. 477–489, 2007.
- [125] T. H. Falk and W.-Y. Chan, "Hybrid signal-and-link-parametric speech quality measurement for voip communications," *IEEE transactions on audio, speech, and language processing*, vol. 16, no. 8, pp. 1579–1589, 2008.
- [126] J. Matta, C. Pépin, K. Lashkari, and R. Jain, "A source and channel rate adaptation algorithm for amr in voip using the emodel," in *Proceedings of the 13th international workshop on Network and operating systems support for digital audio and video*. ACM, 2003, pp. 92–99.

- [127] J. Makinen and J. Vainio, "Source signal based rate adaptation for gsm amr speech codec," in *Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on*, vol. 2. IEEE, 2004, pp. 308–313.
- [128] Y. Huang, J. Korhonen, and Y. Wang, "Optimization of source and channel coding for voice over ip," in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*. IEEE, 2005, pp. 4–pp.
- [129] F. Sabrina and J.-M. Valin, "Priority based dynamic rate control for voip traffic," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*. IEEE, 2009, pp. 1–8.
- [130] M. Chen and M. N. Murthi, "Optimized unequal error protection for voice over ip," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*, vol. 5. IEEE, 2004, pp. V–865.
- [131] A. Khalifeh and H. Yousefi'zadeh, "Optimal rate adaptation for voip over wireless tandem links," in *MILITARY COMMUNICATIONS CONFERENCE, 2010-MILCOM 2010*. IEEE, 2010, pp. 287–292.
- [132] M. J. Karam and F. A. Tobagi, "Analysis of delay and delay jitter of voice traffic in the internet," *Computer Networks*, vol. 40, no. 6, pp. 711–726, 2002.
- [133] S. A. Ahson and M. Ilyas, *VoIP handbook: applications, technologies, reliability, and security*. CRC Press, 2008.
- [134] B. Sklar, *Digital communications*. Prentice Hall NJ, 2001, vol. 2.
- [135] E. Piri, J. Pinola, F. Fitzek, and K. Pentikousis, "Rohc and aggregated voip over fixed wimax: an empirical evaluation," in *Computers and Communications, 2008. ISCC 2008. IEEE Symposium on*. IEEE, 2008, pp. 1141–1146.
- [136] Y. Xiang, J. Luo, and C. Hartmann, "Inter-cell interference mitigation through flexible resource reuse in ofdma based communication networks," in *European wireless*, 2007, pp. 1–4.
- [137] R. M. Heiberger and E. Neuwirth, *R through Excel: A spreadsheet interface for statistics, data analysis, and graphics*. Wiley Online Library, 2009.
- [138] B. R. Hargreaves and T. P. McWilliams, "Polynomial trendline function flaws in microsoft excel," *Computational Statistics & Data Analysis*, vol. 54, no. 4, pp. 1190–1196, 2010.
- [139] L. L. Lim and D. W. G. Lim, "Hybrid log-map algorithm for turbo decoding over awgn channel," in *ICWMC 2011, The Seventh International Conference on Wireless and Mobile Communications*, 2011, pp. 211–214.
- [140] M.-A. Khalighi, "Effect of mismatched snr on the performance of log-map turbo detector," *Vehicular Technology, IEEE Transactions on*, vol. 52, no. 5, pp. 1386–1397, 2003.

- [141] A. Alfayly, I. Mkwawa, L. Sun, and E. Ifeachor, "Qoe-based performance evaluation of scheduling algorithms over lte," in *Globecom Workshops (GC Wkshps), 2012 IEEE*. IEEE, 2012, pp. 1362–1366.
- [142] A. Alfayly, I.-H. Mkwawa, L. Sun, and E. Ifeachor, "Qoe-driven lte downlink scheduling for voip application," in *Consumer Communications and Networking Conference (CCNC), 2015 12th Annual IEEE*. IEEE, 2015, pp. 603–604.
- [143] M. Nasimi, M. Kousha, and F. Hashim, "Qoe-oriented cross-layer downlink scheduling for heterogeneous traffics in lte networks," in *Communications (MICC), 2013 IEEE Malaysia International Conference on*. IEEE, 2013, pp. 292–297.
- [144] M. S. Mushtaq, B. Augustin, and A. Mellouk, "Qoe-based lte downlink scheduler for voip," in *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*. IEEE, 2014, pp. 2190–2195.
- [145] S. Choi, K. Jun, Y. Shin, S. Kang, and B. Choi, "Mac scheduling scheme for voip traffic service in 3g lte," in *Vehicular Technology Conference, 2007. VTC-2007 Fall. 2007 IEEE 66th*. IEEE, 2007, pp. 1441–1445.
- [146] D. Chiranjeevi, V. Mahender, B. Kiran, and B. R. Kumar, "Channel aware scheduling algorithms for 3gpp lte downlink."
- [147] P. Ameigeiras, J. Wigard, and P. Mogensen, "Performance of the m-lwdf scheduling algorithm for streaming services in hsdpa," in *Vehicular technology conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, vol. 2. IEEE, 2004, pp. 999–1003.
- [148] J.-H. Rhee, J. M. Holtzman, and D.-K. Kim, "Scheduling of real/non-real time services: adaptive exp/pf algorithm," in *Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual*, vol. 1. IEEE, 2003, pp. 462–466.
- [149] A. Technologies, Available:http://rfmw.em.keysight.com/rfcomms/refdocs/gsm/default.htm#gprsla_amr_bse_config.html#CBDDEBCD, May 2010.
- [150] —, Available:http://rfmw.em.keysight.com/rfcomms/refdocs/gsm/gprsla_wb_amr_bse_config.html, May 2010.
- [151] S. Saha and R. Quazi, "Priority-coupling-a semi-persistent mac scheduling scheme for voip traffic on 3g lte," in *Telecommunications, 2009. ConTEL 2009. 10th International Conference on*. IEEE, 2009, pp. 325–329.
- [152] L. Wang, Y.-K. Kwok, W.-C. Lau, and V. K. Lau, "Channel adaptive fair queueing for scheduling integrated voice and data services in multicode cdma systems," *Computer Communications*, vol. 27, no. 9, pp. 809–820, 2004.
- [153] C.-N. Chuah and R. H. Katz, "Characterizing packet audio streams from internet multimedia applications," in *Communications, 2002. ICC 2002. IEEE International Conference on*, vol. 2. IEEE, 2002, pp. 1199–1203.
- [154] M. Reisslein, L. Karam, and P. Seeling, "H. 264/AVC and SVC Video Trace Library: A Quick Reference Guide <http://trace.eas.asu.edu>," 2009.

-
- [155] 3GPP, “Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN),” 3rd Generation Partnership Project (3GPP), TR 25.913, 12 2009. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/25913.htm>
- [156] K. Sandrasegaran, H. A. M. Ramli, and R. Basukala, “Delay-prioritized scheduling (dps) for real time traffic in 3gpp lte system,” in *2010 IEEE Wireless Communication and Networking Conference*. IEEE, 2010, pp. 1–6.
- [157] F. Haider, E. Hepsaydir, and N. Binucci, “Performance analysis of lte-advanced networks in different spectrum bands,” in *Wireless Advanced (WiAd), 2011*. IEEE, 2011, pp. 230–234.

Appendix A

Effects of the MQS on system performance

The following simulation results were published in [6].

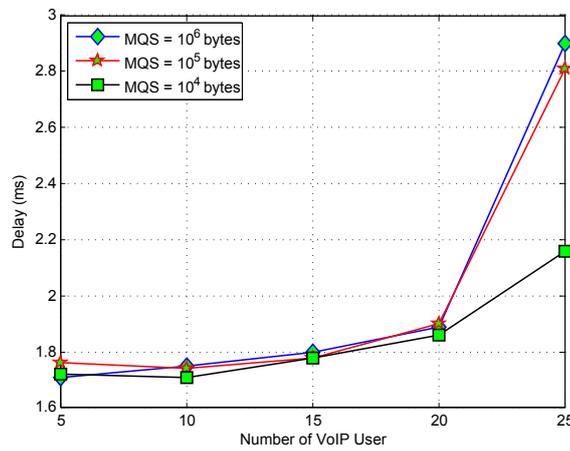


Figure A.1: The effect of the MQS on delay of VoIP flow

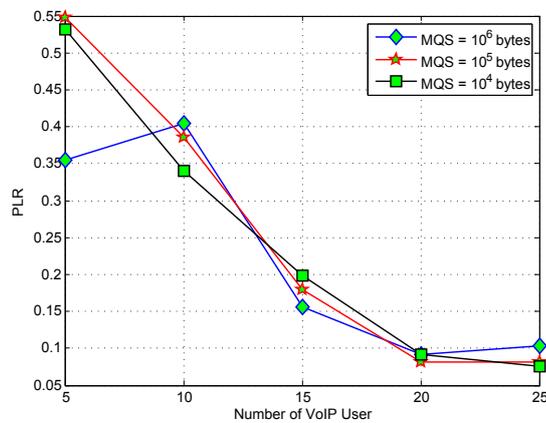


Figure A.2: The effect of the MQS on PLR of VoIP flow

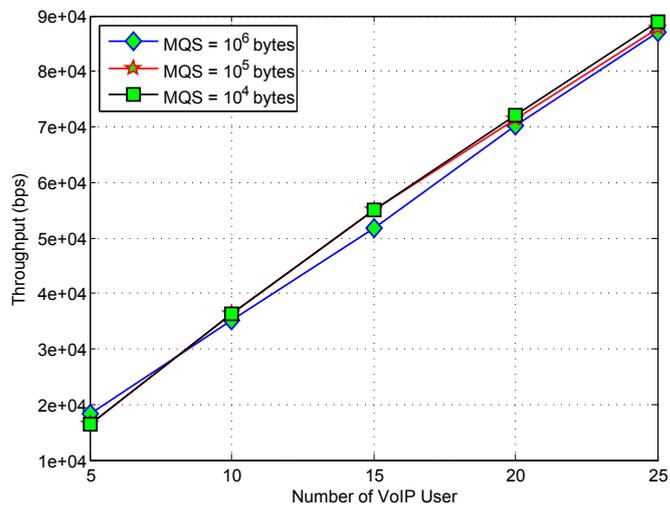


Figure A.3: The effect of the MQS on cell throughput of VoIP flow

Appendix B

Performance evaluation of the E-MQS scheduler in LTE heterogeneous with mobility

In the simulation scenario, each user equipment uses a VoIP flow, a video flow, and a “best effort” flow at the same time, and randomly moves at a speed of 120 km/h.

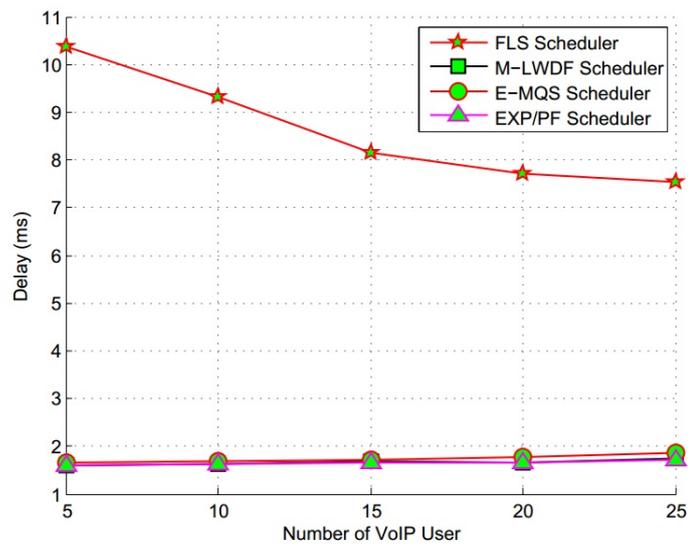


Figure B.1: Delay vs VoIP users at user speed of 120 km/h

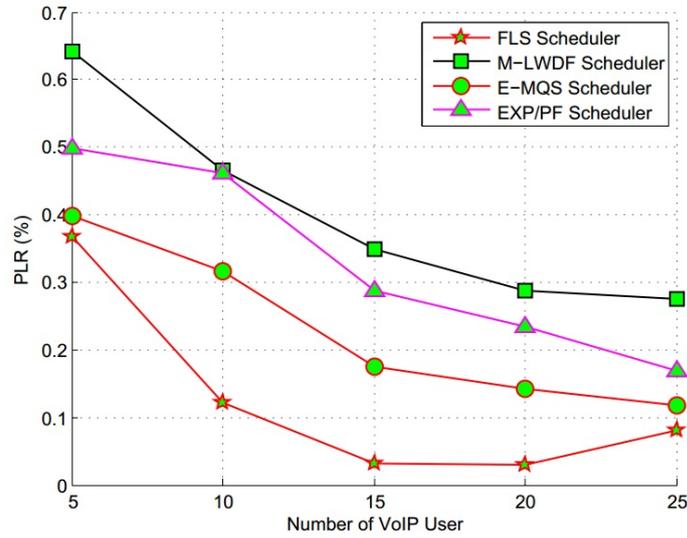


Figure B.2: PLR vs VoIP users at user speed of 120 km/h

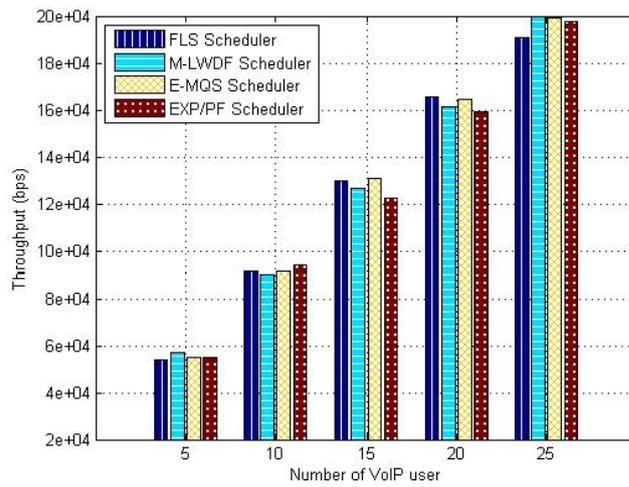


Figure B.3: Cell throughput vs VoIP users at user speed of 120 km/h

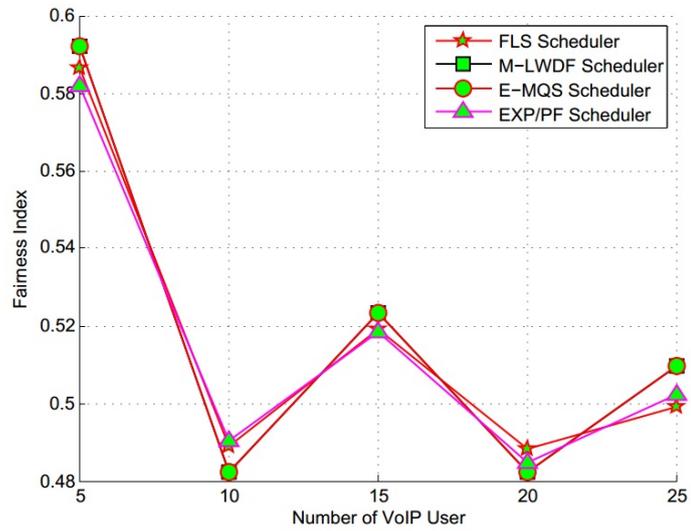


Figure B.4: FI vs VoIP users at user speed of 120 km/h

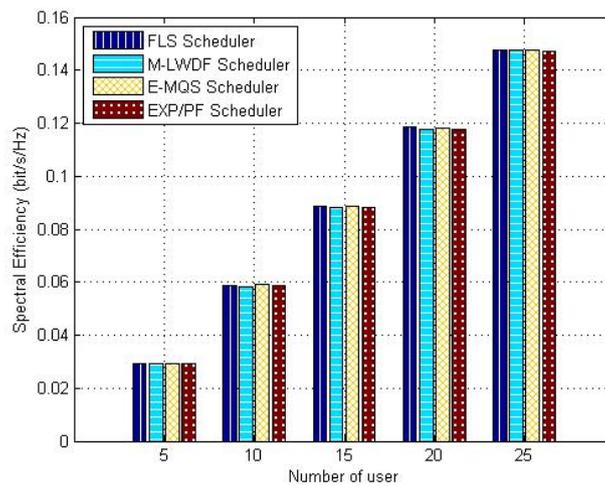


Figure B.5: SE vs the number of user at user speed of 120 km/h

