Application-Specific and QoS-Aware Scheduling for Wireless Systems

Chao He and Richard D. Gitlin

Department of Electrical Engineering
University of South Florida
Tampa, Florida 33620, USA
Email: chaohe@mail.usf.edu, richgitlin@usf.edu

Abstract—**This paper describes how user-specific QoS requirements are a critical innovation that may be used to improve spectral utilization in wireless systems. Rate adaptation and MAC scheduling algorithms that process user-specific QoS to improve system capacity are proposed and evaluated. Through dynamically adapting the AMR codec mode or video data rate and the MAC scheduling algorithm to the user-specific QoS requirements, system capacity, as measured by the number of supportable users, is maximized, while user satisfaction, as measured by the Mean Opinion Score (MOS), is maintained at an acceptable level. OPNET system simulations were performed for a set of VoIP users and video users that were assigned specific QoS target levels. Simulation results show that significant system capacity improvement and acceptable MOS level can be achieved if such user-specific QoS requirements are considered in the rate adaptation and MAC scheduling algorithms.**

Keywords—**System capacity, rate adaptation, MAC scheduler, user-specific QoS, *MOS*, AMR, VoIP, video.**

I. INTRODUCTION

In today's wireless 4G LTE networks, the spectral allocation of resources is either independent of the application's specific Quality of Service (QoS) requirements and of the users' specific perceived QoS, or at most relies on a set of pre-defined fixed priorities [1], [2]. Indeed, from the user's perspective, the QoS required by different applications can be quite variable. Similarly, for a given application type, different users may require different levels of QoS.

For VoIP applications, as a motivating example, consider the fact that the perceived voice quality of different languages may differ substantially when allocated the same data rate and Bit Error Rate (BER), because of the different spectral content of such languages and because of a particular user's auditory spectral response (with variations typically due to aging), making the user more or less sensitive to a particular type of distortion. Consequently, the same amount of degradation, as experienced by individual applications and their users, may have substantially different perceptual effects. Another example is the varying talk environments, where some users have a conversation under very noisy conditions, while some other users converse under very quiet conditions, thus making users more or less sensitive to packet losses. If the same amount of spectral resources is allocated to users in very noisy and quiet backgrounds, then an unacceptable user experience will likely be incurred. As another example, consider that people from different age groups normally have different sensitivity to high frequency(ies) content, which can be

exploited to maximize the system capacity by reducing the bit rate for users with reduced frequency sensitivity [3]. For video applications, as a user-specific QoS example, compared with young adults, older individuals were less sensitive to spatial form defined by temporal structure[4]. So for many older people a lower video data rate provides the same user experience as the full rate video does for younger people. As another user-specific QoS example, to achieve the same user experience for different video content (e.g. news and sports video), the required video data rate can be quite different. The required data rate of news video can be much less than that of sports video. Hence, the user-specific QoS requirement can be utilized by the scheduler to differentiate the users and better make use of the wireless spectral resources.

In most commercial systems, the user-specific QoS requirements can be obtained by the network operator. When users subscribe to a service from the wireless operator, they often provide their relevant information such as age, name, nationality that can be used to derive user-specific QoS requirements, or they can be obtained dynamically (e.g. noisy or quiet environment for VoIP users) whenever a session is started. For video applications, we can also use deep packet inspection [5] to determine the application-specific requirements.

Furthermore, we observe that some previous studies (e.g. [6], [7]), which use the QoS characteristics of an underlying application (typically expressed as a function of the MOS), allocate average spectral resources to applications, independently of the application's actual specific QoS requirement. Though, in the literature, there are MAC schedulers that take into account instantaneous data rates and user's QoS [8], [9]; to the best of our knowledge, no user-specific QoS requirements have been considered in the MOS functions and in the MAC scheduler. Thus, in such schemes, especially for applications with widely varying QoS requirements (even for the same type of application), either the spectral resources are not efficiently utilized or the MOS is significantly degraded.

In this paper we present a novel rate adaptation algorithm and QoS-aware MAC scheduling algorithm that maximize spectrum utilization and maintain user satisfaction by trading off the spectral resource allocations of connections for the application-level QoS based on the user-specific requirements. We focus on voice applications and video applications in the context of 4G LTE wireless systems, and through dynamically adapting data rate and MAC scheduling algorithms to the user-

specific QoS requirements, system capacity (i.e., the number of supportable users) is maximized, while user satisfaction can be maintained at a comparable level.

The paper is organized as follows. In section II, the existing algorithms of MOS definitions for both VoIP and video applications are described. Our user-specific MOS formulas and novel rate adaptation and MAC scheduling algorithms are described in section III. Section IV presents the OPNET system simulation. Finally, our conclusions are presented in Section V.

II. MOS DEFINITIONS

A. Voice over IP [VoIP] Applications

The Adaptive Multi-Rate (AMR) audio codec is an audio data compression scheme that is used in LTE and is optimized for speech coding. AMR consists of a multi-rate speech codec that encodes speech signals at variable bit rates ranging from 4.75 to 12.2 kbit/s [10].

The E-Model algorithm [11], which is a computational model for objective call quality assessment, is described in the ITU-T G.107 recommendation. The computation of the MOS is defined as follows:

$$R = R_0 - I_d - I_{eff} \qquad (1)$$

where R_0 is the basic signal-to-noise ratio which has a default value of 93.2 [12], [13], I_d represents the impairments due to delay, which is the same for all the codec modes, and I_{eff} represents the effect of packet losses and depends on the codec (e.g. AMR, G.711) that is used. I_d is calculated as:

$$I_d = 0.024d + 0.11(d - 177.3)U(d - 177.3) \qquad (2)$$

where d is the end-to-end delay in milliseconds and U is the unit step function [13].

For AMR codecs, the I_{eff} is given by [11]:

$$I_{eff} = I_e + (95 - I_e)\left(\frac{100P_{pl}}{\frac{100P_{pl}}{BurstR} + B_{pl}}\right) \qquad (3)$$

where P_{pl} represents packet loss ratio, $BurstR$ is the Average length of observed bursts in an arrival sequence to the Average length of bursts expected for the network under "random" loss ratio. In this paper we assume the packet loss is independent and hence we set $BurstR = 1$. B_{pl} is the robustness factor which is set to 10 for all AMR codec modes. I_e is defined for all AMR codec modes in [14], where eight AMR-NB codec modes are defined in LTE [10].

R is converted to MOS according to (4):

$$MOS = \begin{cases} 1, & when\ R < 0 \\ 1 + 0.035R + R(R - 60)(100 - R) \\ 7 \cdot 10^{-6}, & when\ R \in [0, 100] \\ 4.5, & when\ R > 100 \end{cases} \qquad (4)$$

From (1)-(4), the lower the delay, or the lower the packet loss ratio, the higher the MOS value.

B. Video Applications

In this paper, a simplified video MOS model [6], [15] is used, where the distortion, as measured by the MSE (Mean Square Error) is assumed to be composed of two additive components, namely the source distortion (D_S) and the loss distortion (D_L):

$$MSE = D_S + D_L = \eta \cdot R^\xi + \beta \cdot PEP \qquad (5)$$

In (5), η, ξ, and β are model parameters and PEP is the packet loss ratio. For different types of video sources, η, ξ, and β take different values. In this paper, we assume $\eta = 1.76 \cdot 10^5$, $\xi = -0.658$, and $\beta = 1750$ as in [15]. The PSNR is a widely used objective measurement of video quality, and is related to the MSE by:

$$PSNR(dB) = 10 * \log_{10}\frac{255^2}{MSE} \qquad (6)$$

A piecewise linear mapping from the PSNR to MOS is shown in (7):

$$MOS = \begin{cases} 1, & when\ PSNR < 20 \\ 1 + \frac{3.5}{20}(PSNR - 20), & when\ PSNR \in [20, 40] \\ 4.5, & when\ PSNR > 40 \end{cases} \qquad (7)$$

From (5)-(7), the higher the data rate, or the lower the packet loss ratio, the higher the MOS value.

III. PROPOSED RATE ADAPTATION AND MAC SCHEDULING ALGORITHMS

The novelty of the proposed Rate Adaptation and MAC Scheduling Algorithms is that they incorporate user-specific QoS requirements into the scheduling and personalize individual UE scheduling utilizing this user-specific QoS information to improve system performance as described below. The proposed scheduling algorithms are composed of three parts, the AMR mode adaptation algorithm, the video data rate adaptation algorithm and the MAC resource scheduling algorithm.

A. LTE baseline Scheduling

The MAC scheduling comprises two scheduling components [8], [9] that are done sequentially in each scheduling time unit, which is known as Transmission Time Interval (TTI) in LTE (TTI = 1ms). The first component is the time domain scheduler (TDS) and the second is the frequency domain scheduler (FDS). The objective of the time domain scheduler is to choose a subset of users requesting frequency resources, while the objective of frequency domain scheduler is to allocate physical resources for the candidate users provided by the time domain scheduler.

The benchmark for performance comparison is the LTE baseline scheduler that doesn't consider the user-specific QoS requirement. where the time domain and frequency domain schedulers function as follows. It is also easy to extend the approach in this paper to other baseline schedulers to do a fair comparison with and without user-specific QoS requirements.

1) Time Domain Scheduler

Users with higher metrics (e.g., packet delay) can get higher scheduling priority and resources in the time domain. The packet delay metric for user k is defined as:

$$M_k = TW_k * Delay\ of\ Packet \quad\quad (8)$$

where $TW_k = 1$ for all users, which means users are not differentiated by their specific QoS requirements. $Delay\ of\ Packet$ is the packet delay in the MAC buffer.

2) Frequency Domain Scheduler

Each user has a C/I (Carrier-to-Interference) metric for each sub-band and is sorted for each sub-band among all the scheduled users. A max C/I approach is used in the LTE baseline scheduler, where each sub-band is first allocated to the user that has the highest C/I , then to the user with the second and third highest C/I, and so on until all the resources of this given sub-band are allocated. The C/I metric for user k in each sub-band n is defined by:

$$M_{n,k} = FW_{n,k} * SINR_{n,k} \quad\quad (9)$$

where $FW_{n,k} = 1$, which means users are not differentiated by their specific QoS requirements, and $SINR_{n,k}$ is the SINR for user k in sub-band n.

B. UE-Specific MOS formulas

1) UE-Specific VoIP MOS Formula

Here we have assumed that different people have similar sensitivity to the end-to-end delay for VoIP applications, so that only UE specific sensitivity to packet losses is studied. To reflect different users' sensitivity to packet losses, a UE specific sensitivity factor, α, is added to (1). Therefore,

$$R = R_0 - I_d - \alpha \cdot I_{eff} \quad\quad (10)$$

In this paper, without loss of generality and also for simplicity of illustration, the packet-loss sensitivity factor α takes values from the following set $\{0.8, 0.9, 1, 1.1, 1.2\}$. The higher the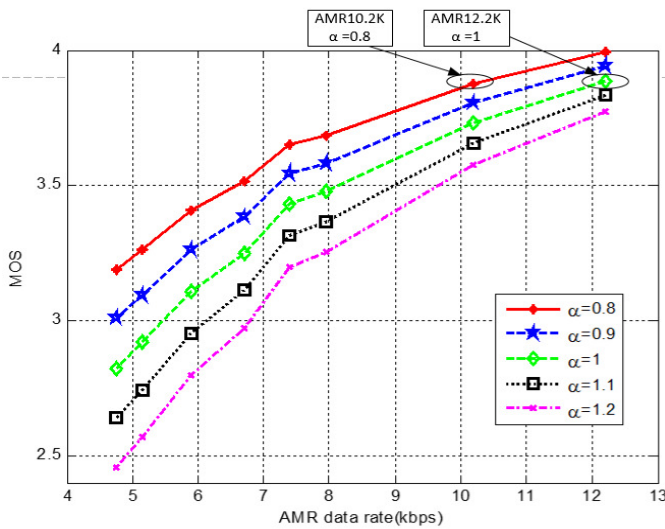 value of the sensitivity factor (α), the user is increasingly sensitive to packet loss. When α takes the value of 1, it is a normal user. When α takes the value greater than 1, the user is more sensitive to packet losses compared with the normal user. When α takes the value less than 1, it is less sensitive to packet losses compared with the normal user.

Figure 1 shows the MOS as a function of different AMR data rates for different sensitivity factors α, given an end-to-end delay of 150 ms and packet loss ratio of 0.01. For a comparison between AMR12.2K mode and α = 1.0 with AMR10.2K mode and α = 0.8, we can find users with AMR10.2K mode and α = 0.8 may have a higher MOS than users with AMR12.2K mode and α = 1.0. If the scheduler can know, or adaptively learn, each user's application specific sensitivity factors, it can degrade the AMR mode for users with a lower sensitivity factor, while maintaining a comparable MOS as that of users with higher AMR mode but a normal sensitivity factor. With this approach, more users can be supported, thus achieving the target of improving system capacity.

2) UE-Specific Video MOS Formula

To reflect user sensitivity to the data rate, a UE specific sensitivity factor γ is added to (5), and it becomes:

$$MSE = D_S + D_L = \gamma \cdot \eta \cdot R^\xi + \beta \cdot PEP \quad\quad (11)$$

When γ takes the value 1, it is a normal user. When γ takes the value greater than 1, the user is more sensitive to the data rate compared with a normal user. When γ takes the value less than 1, it is less sensitive to the data rate compared with a normal user. Fig. 2 shows the MOS as a function of data rate for different sensitivity factors γ under a given packet loss ratio of 0.001. An important observation that can be made from Fig. 2 is that a user with a lower sensitivity factor and a lower data rate can achieve a higher MOS value than that of users with a normal sensitivity factor or higher sensitivity factor and a higher data rate. If the application-aware scheduler knows and makes use of this UE specific sensitivity factor information to optimize the scheduling, it can decrease the data rate for users with a lower sensitivity factor to support more users with an acceptable MOS value. The following sub-section provides
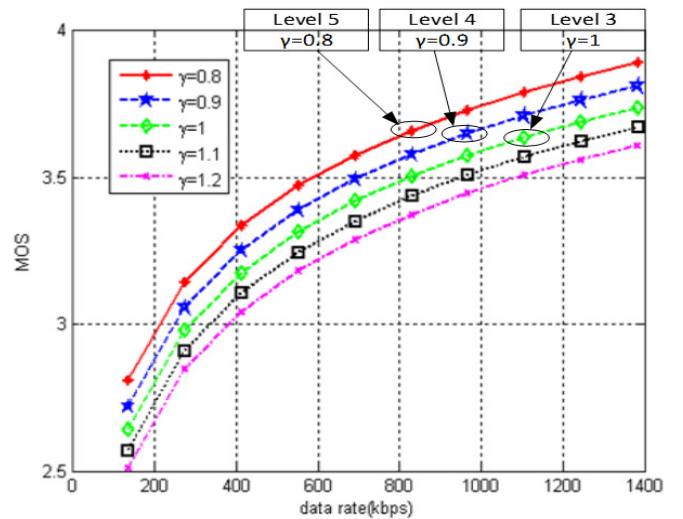

Fig.1. VoIP MOS as a function of AMR data rate given packet loss ratio of 0.01 and end-to-end delay of 150ms

Fig.2. Video MOS as a function of data rate for different sensitivity factors, γ, given a packet loss ratio of 0.001 for video applications.

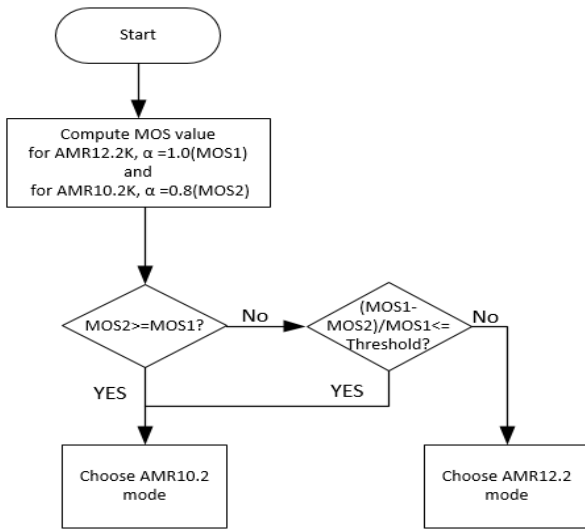Fig.3. AMR mode adaption work flow

further details.

C. Proposed Rate Adaption Algorithms

1) AMR Mode Adaption

In order to illustrate the main idea of user-specific QoS scheduling, in this paper, we only consider two AMR modes (i.e. AMR 12.2K and AMR10.2K), the extension to other AMR modes is straightforward. The workflow of the AMR mode adaption is shown in Fig. 3. The threshold to degrade the AMR mode can be configured to control the desired MOS levels. In this paper, it is set to 0.02. The AMR mode will be degraded if the MOS is decreased by less than 0.02, compared with that of the MOS value for the non-degraded AMR mode, with $\alpha = 1$. The input to the AMR mode adaption is the packet loss ratio fed back from the UEs, while assuming an average end-to-end delay of 150ms.
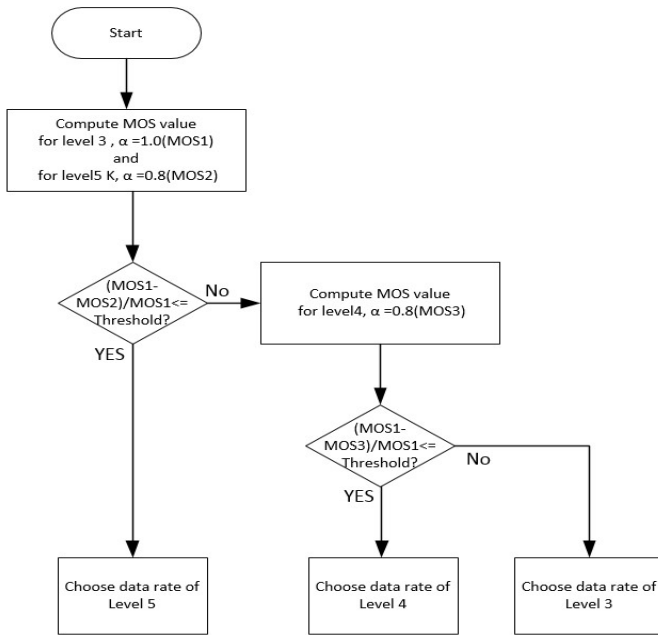

Fig.4. Video data rate adapation workflow

2) Video Data Rate Adaptation

For simplicity 10 levels of data rate are defined in the paper, which loosely correspond to the application requirements, in order to illustrate the main idea of the algorithms. For Level I {1, 2, …, 10}, the corresponding data rate is 135 * 128 pixels * (11-I) frames/s * 8bytes/pixel. The workflow of the video data rate adaptation is shown in Fig. 4. For simplicity of illustration, three levels of data rate are assumed in the video data rate adaptation.

The video data rate level to be selected depends upon the respectively calculated MOS for each level of data rate. Similar to VoIP users, the threshold to degrade the video data rate can be configured to control the desired MOS levels. In this paper, it is set to 0, which means that theoretically the MOS of degraded users will not be decreased.

D. MAC Resource Scheduling Algorithm [3]

1) Time Domain Scheduler

The same metric is applied as the LTE baseline scheduler except that $TW_k = \frac{1}{0.2}$ for VoIP degraded users and $TW_k = 1$ in other cases.

2) Frequency Domain Scheduler

The same metric is applied as the LTE baseline scheduler except that $FW_{n,k} = 10$ for VoIP degraded users for their respective best sub-band and $FW_{n,k} = 1$ in other cases.

IV. SYSTEM SIMULATION

A. System Simulation Configuration

The system simulation was run using the OPNET 17.5 Modeler with the LTE modules. In this paper, a single cell with 50 AMR VoIP users and 10 video users was tested for the downlink scheduling, with an ideal uplink receiver.

B. System Simulation Scenarios

Two scenarios were designed and simulated as described in

TABLE I   SYSTEM SIMULATION CONFIGURATION

Parameter	Assumption
Cellular Layout	1 Cell
Cell Radius	1Kilometer
Path loss model	3GPP suburban Macrocell
Mobility model	Random Way Point (RWP) with speed of 0.1km/h
Carrier Frequency	Uplink:1920MHz Downlink:2110MHz
System Bandwidth	10MHz
Channel model	ITU Pedestrian A
Total BS TX power	40dBm
UE power class	23dBm
VoIP codec modes	AMR12.2,AMR10.2K, and mixed codec modes
Video date rate	1105.920 Kbits/s (Level 3) 967.680 Kbits/s (Level 4) 829.440 Kbits/s (Level 5)
Number of Users	50 VoIP Users, 10 video users
Scheduler	Dynamic scheduling The proposed scheduler and LTE baseline scheduler
Other assumptions	Ideal uplink receiver(no block error and packet loss)

TABLE II.　　SYSTEM SIMULATION SCENARIOS

Scenarios	Assumption
Scenario 1	50 AMR VoIP Users, each user takes a value for the packet-loss sensitivity factor α of 0.8.
Scenario 2	50 AMR VoIP Users, each user takes a value for the packet-loss sensitivity factor α of 0.8. 10 video users, each user takes a value for the data rate sensitivity factor α of 0.8.

Table II. In Scenario 1, 50 users have packet-loss sensitivity factors of 0.8, while in Scenario 2, 50 VoIP users have packet-loss sensitivity factors of 0.8 and 10 video users have data rate sensitivity factors of 0.8.

C.　Simulation Results

The simulation results for capacity improvement are shown in Table III. In this paper, the downlink MAC throughput is used to derive the approximate system capacity improvement. System capacity improvement is measured by the improvement of maximum supportable number of users by the system. A rough mapping from the downlink MAC throughput to the system capacity improvement can be done based upon (12). According to (12), it can be calculated that approximately 11.7% capacity improvement can be achieved for the proposed scheduler for scenario 1, while 12.7 % capacity improvement is achieved for scenario 2. Capacity improvement exceeding 10% is viewed as significant by the cellular industry[16].

$$\text{Capacity impovement} \qquad (12)$$
$$= \frac{1/(\text{MAC throughput for proposed scheduler})}{1/(\text{MAC throughput for baseline scheduler })} - 1$$

The simulation results for the average MOS for both VoIP users and video users (Scenario 2) are shown in Table IV. From the Table IV, we see that only very slight decrease of MOS will be incurred in order to achieve the desired system capacity improvement.

V.　CONCLUSIONS

In this paper, we presented novel rate adaptation and MAC scheduling algorithms that can significantly improve the

TABLE III.　　SYSTEM CAPACITY COMPARISON

Scenarios	Scheduler	MAC throughput(kbps)	Capacity improvement (%)
Scenario 1	Baseline	896.60	0
	Proposed	802.72	11.7%
Scenario 2	Baseline	9,062	0
	Proposed	7,123	12.7%

TABLE IV.　　AVERAGE MOS COMPARISON

Scenarios	Scheduler	Average MOS for VoIP users	Average MOS for video users
Scenario 1	Baseline	4.055	N/A
	Proposed	4.023	N/A
Scenario 2	Baseline	4.04	3.50
	Proposed	3.99	3.47

system capacity by considering user-specific QoS requirements, while maintaining an acceptable MOS level at the same time. The MAC scheduler used for comparsion can be easily extended to other MAC schedulers to perform a comparison with and without user-specific QoS requirements. Initial simulation and research show that user-specific QoS study can be a very promising research field that is quite novel and will open up new approaches to improving the quality, capacity, and performance of wireless systems.

VI.　ACKNOWLEDGMENT

REFERENCES

[1] H. Ekstrom, "QoS control in the 3GPP evolved packet system," *IEEE Communications Magazine*, vol. 47, no. 2, pp. 76–83, 2009.

[2] 3GPP, "TS 23.203 - V10.9.0 - Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control architecture." 3GPP.

[3] C. He and R. D. Gitlin, "User Specific QoS and its Application in Resources Scheduling for Wireless System," to be presented at the 5th International Workshop on Wireless Networks and Multimedia, 2014.

[4] R. Blake, M. Rizzo, and S. McEvoy, "Aging and Perception of Visual Form From Temporal Structure," *Psychol Aging*, vol. 23, no. 1, pp. 181–189, Mar. 2008.

[5] A. Chaudhary and A. Sardana, "Software Based Implementation Methodologies for Deep Packet Inspection," in *2011 International Conference on Information Science and Applications (ICISA)*, 2011, pp. 1–10.

[6] S. Khan, S. Duhovnikov, E. Steinbach, and W. Kellerer, "MOS-Based Multiuser Multiapplication Cross-Layer Optimization for Mobile Multimedia Communication," *Advances in Multimedia*, vol. 2007, pp. 1–11, 2007.

[7] A. Saul and G. Auer, "Multiuser Resource Allocation Maximizing the Perceived Quality," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, pp. 1–15, 2009.

[8] P. Kela, J. Puttonen, N. Kolehmainen, T. Ristaniemi, T. Henttonen, and M. Moisio, "Dynamic packet scheduling performance in UTRA Long Term Evolution downlink," presented at the 3rd International Symposium on Wireless Pervasive Computing, 2008. ISWPC 2008, 2008, pp. 308–313.

[9] K. C. Beh, S. Armour, and A. Doufexi, "Joint Time-Frequency Domain Proportional Fair Scheduler with HARQ for 3GPP LTE Systems," presented at the Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th, 2008, pp. 1–5.

[10] 3GPP, "TS 26.071 - V11.0.0 - Mandatory speech CODEC speech processing functions;AMR speech CODEC; General description." .

[11] Telecommunication standardization sector of ITU, "G.107: The E-model: a computational model for use in transmission planning." International Telecommunication Union, Dec-2011.

[12] J. Fitzpatrick, "An E-Model based adaptation algorithm for AMR voice calls," in *Wireless Days (WD), 2011 IFIP*, 2011, pp. 1–6.

[13] R. G. Cole and J. H. Rosenbluth, "Voice over IP Performance Monitoring," *SIGCOMM Comput. Commun. Rev.*, vol. 31, no. 2, pp. 9–24, Apr. 2001.

[14] F. Mertz and P. Vary, "Efficient Voice Communication inWireless Packet Networks," in *2008 ITG Conference on Voice Communication (SprachKommunikation)*, 2008, pp. 1–4.

[15] D. Jurca and P. Frossard, "Media Flow Rate Allocation in Multipath Networks," *IEEE Transactions on Multimedia*, vol. 9, no. 6, pp. 1227–1240, Oct. 2007.

[16] "Private communication with Chih-Lin I, Chief Scientist, China Mobile Research Institute," Jun. 25, 2014.