2015 IEEE Wireless Telecommunications Symposium (WTS)
New York City, USA
April 15-17, 2015

User-Specific QoS Aware Scheduling and Implementation in 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*—In this paper, we explore user-specific QoS requirements and associated schedulers that are very critical in optimizing the spectral allocation for wireless systems. Two user-specific QoS aware schedulers are proposed that considers the user-specific QoS requirements in the allocation of resources. Depending upon whether improving the MOS (Mean Opinion Score) or both the system capacity and the MOS is the goal, a MOS improvement scheduler or MOS-plus-capacity improvement scheduler is proposed for VoIP applications. Detailed system implementation analysis based upon LTE system specification is performed, and it is shown that very modest modifications to current protocols are needed to support user-specific QoS aware scheduling. System simulations are performed for a set of VoIP users assigned specific QoS target levels in the OPNET Modeler for LTE systems. Simulation results show that appreciable MOS or/and system capacity improvement can be achieved if such user-specific QoS requirements are considered in the proposed user-specific QoS aware schedulers. Also, it is shown that the scheduling period of up to 1000 ms doesn't significantly impair the system performance.**

***Keywords*—AMR, implementation, MAC, MOS, system capacity.**

I. INTRODUCTION

In today's wireless 4G LTE networks, the spectral allocation of resources is either independent of the users' specific perceived QoS (Quality of Service), or at most relies on a set of pre-defined fixed priorities [1], [2]. Although in these standards, the MAC and the PHY layers have an increased role in optimizing the usage of the spectral resources and implementing link quality-aware techniques, optimization is still largely independent of the application context, the users' requirements, and the users' perception of performance degradation. The allocation of resources, does not take into account the QoS required by different applications and their users, beyond simply assigning fixed priorities to traffic classes. Indeed, for a given application type, different users may require different levels of QoS.

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 BER (Bit Error Rate), 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 [3]. 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 talking environments, where some users have a conversation under very noisy conditions, whereas some other users converse under very quiet conditions, thus making users more or less sensitive to packet losses. If the same amount of spectral resource is allocated to users in very noisy and quiet backgrounds, then quite a different user experience will likely be incurred.

The user-specific QoS aware scheduling algorithms were first addressed in [4]–[6], where it was shown that significant MOS improvement or/and system capacity can be achieved if user-specific QoS requirements are considered in the scheduling algorithms. In this paper, the scheduling algorithms are presented briefly, and we will focus on the analysis of detailed implementation complexity for LTE systems where the user-specific QoS related protocol adaptation and scheduling period are addressed.

The rest of the paper is organized as follows. In Section II, user-specific QoS aware schedulers are described. Detailed implementation analysis of user-specific QoS aware schedulers is provided in Section III. Section IV presents the OPNET LTE system simulations. Finally, our conclusions are given in Section V.

II. USER-SPECIFIC QOS SCHEDULER

A. *User-Specific MOS Formulas*

In this paper, user-specific QoS requirements are characterized by their different sensitivities to packet losses. To reflect this different sensitivity, a user-specific packet loss sensitivity factor, α, is introduced to the ITU-T G.107 E-Model equation [7]:

$$R = R_0 - I_d - \alpha \cdot I_{eff} \tag{1}$$

where R_0 is the basic signal-to-noise ratio which has a default value of 93.2 [8], 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 mode) that is used. The parameter I_d is calculated as [8]:

$$I_d = 0.024d + 0.11(d - 177.3)U(d - 177.3) \tag{2}$$

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

2015 IEEE Wireless Telecommunications Symposium (WTS)
New York City, USA
April 15-17, 2015

For AMR codecs, the I_{eff} is given by [7]:

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

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 losses are independent and hence we set $BurstR = 1$. The parameter B_{pl} is the robustness factor which is set to 10 for all AMR codec modes. The parameter I_e is defined for all AMR codec modes in [9], where eight AMR-NB codec modes are defined in LTE [10].

The parameter R is converted to MOS according to (4):

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

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

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 $\{0.8, 1.0, 1.2\}$. Correspondingly, users are classified into 3 categories: users with higher (1.2), normal (1.0), and lower (0.8) sensitivity factors. The higher the value of the sensitivity factor α, the more the user is sensitive to packet loss.

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 [11] and packet loss ratio of 0.05. For a comparison between AMR12.2K mode and $\alpha = 1.0$ with AMR10.2K/7.95K mode and $\alpha = 0.8$, we find the latter case may, under certain conditions, have a higher MOS than the former one. If the scheduler can know, or adaptively learn, each user's specific sensitivity factor, 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.

B. Motivation for MOS Optimization

Figure 2 shows the decreased or increased MOS percentage due to the different sensitivity factors α for different users. The MOS of VoIP users with α of 1.2 is decreased by ˜15%, whereas the MOS of VoIP users with α of 0.8 is increased by ˜15%, when a packet loss ratio of 5% and end-to-end delay of 150 ms are assumed. As the packet loss ratio increases, the MOS will decrease or increase even more.

Therefore, the MOS of VoIP users with α of 1.2 needs to be improved to the corresponding MOS value with α of 1.0, whereas the MOS of VoIP users with α of 0.8 can be decreased to the MOS value of α of 1.0, as depicted in Fig. 2. There are two approaches to decrease the MOS of VoIP users with α of 0.8:

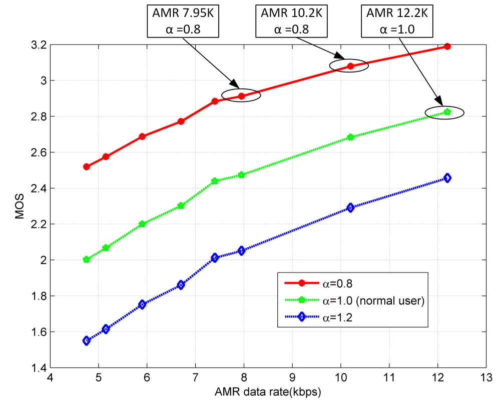

Fig. 1. VoIP MOS as a function of AMR data rate given a packet loss ratio of 0.05 and end-to-end delay of 150 ms.

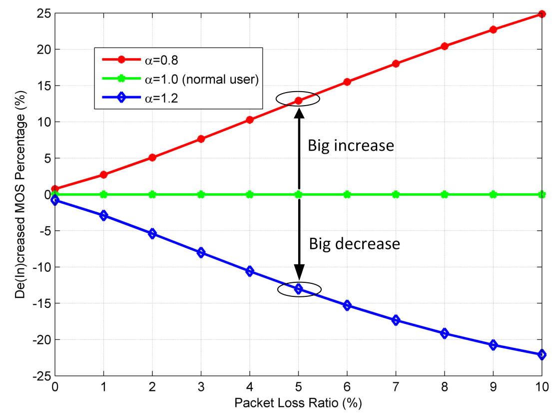

Fig. 2. De(in)creased MOS as a function of packet loss ratio given an end-to-end delay of 150 ms and AMR12.2K.

1) Approach I: This category of users can be deprioritized, that is, given a lower scheduling priority in the MAC scheduler. The MOS optimization only scheduler in the next section uses this method to optimize the MOS.

2) Approach II: The data rate (i.e., AMR mode) of this category of users can be degraded, so that a higher system capacity can be achieved. However, these users are scheduled as normal users in the MAC scheduler. The MOS optimization plus capacity improvement scheduler in the next section uses this method to optimize the MOS.

C. User-Specific QoS Aware Schedulers

1) MOS Optimization only Scheduler: The MOS of users with $\alpha = 1.2$ will be increased, that is, given a higher scheduling priority, whereas users with $\alpha = 1.0$ have a normal scheduling priority, and the MOS of users with $\alpha = 0.8$ are decreased a little and given a lower scheduling priority. This scheduler is denoted as the USQA-M scheduler.

2) MOS Optimization plus Capacity Improvement Scheduler: The MOS of users with $\alpha = 1.2$ will be increased, that is, given a higher scheduling priority, whereas users with $\alpha = 1.0$ have a normal scheduling priority and users with $\alpha = 0.8$ are used to improve the capacity by degrading their AMR codec modes.

2015 IEEE Wireless Telecommunications Symposium (WTS)
New York City, USA
April 15-17, 2015

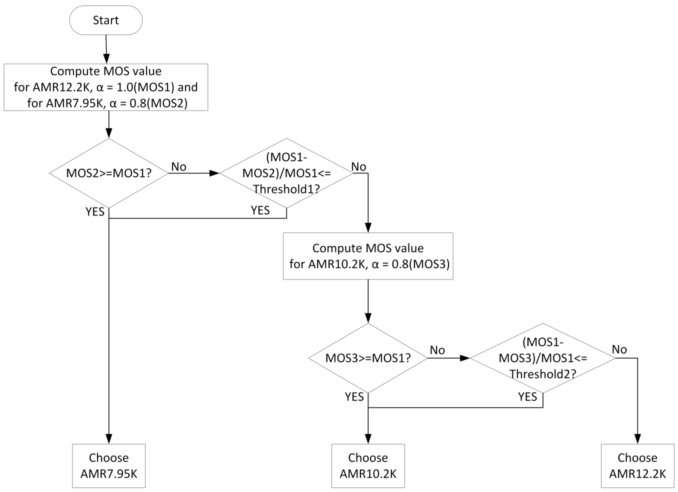

Fig. 3. AMR mode adaptation workflow.

In order to illustrate the main idea of user-specific QoS capacity improvement scheduling, in this paper, we only consider three AMR modes (i.e., AMR 12.2K, 10.2K and 7.95K), and the extension to other AMR modes is straightforward. The workflow of the AMR mode adaptation is shown in Fig. 3. The algorithm starts with the AMR12.2K mode. The thresholds to degrade the AMR mode can be configured to control the desired MOS levels. In this paper, they are set to 0.02, that is, 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 $\alpha = 1.0$. The input to the AMR mode adaptation is the packet loss ratio, while assuming an average end-to-end delay of 150 ms. This scheduler is denoted as the USQA-MC scheduler.

D. User-Specific QoS Aware MAC Scheduling Algorithms

1) Time Domain Scheduler: Users with higher metrics can receive higher scheduling priority in the time domain. The metric for user k is defined as:

$$M_k = TW_k * DOP \tag{5}$$

where for the baseline scheduler, $TW_k = 0.8$ for all users, which means users are not differentiated by their user-specific QoS requirements. For the USQA-M scheduler, $TW_k = 1.0$ for users with $\alpha = 1.2$, $TW_k = 0.7$ for users with $\alpha = 0.8$, whereas for the USQA-MC scheduler, $TW_k = 1.0$ for users with $\alpha = 1.2$ and $TW_k = 0.8$ for other users. In (5) DOP is the packet delay in milliseconds in the MAC buffer. All the users with M_k greater than 20 ms will be selected as the candidate users if the resource is available. The parameter TW_k is configurable based upon different application scenarios.

2) Frequency Domain Scheduler: Each user also has a frequency domain metric for each sub-band and this is sorted for each sub-band among all the scheduled users. Each sub-band is first allocated to the user that has the highest metric, then to the user with the second and third highest metric, and so on until all the resources of this given sub-band are allocated. The metric for user k in each sub-band n is defined by:

$$M_{n,k} = N_k * (MCS_{n,k} - MCS_{wb,k} + FW_{n,k}) \tag{6}$$

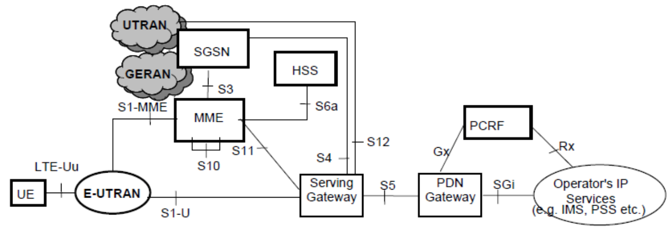

Fig. 4. LTE network architecture.

where N_k is the number of PRBs (Physical Radio Blocks) allocated to user k. $MCS_{n,k}$ and $MCS_{wb,k}$ are the MCS index of user k in sub-band n and wideband respectively. For the baseline scheduler, $FW_{n,k} = 0$ for all users, which means users are not differentiated by their specific QoS requirements. For the USQA-M and USQA-MC schedulers, $FW_{n,k} = 1$ and $FW_{n,k} = -1$ for users with $\alpha = 1.2$ in their best sub-band and other sub-bands respectively, and $FW_{n,k} = 0$ for other users. The parameter $FW_{n,k}$ is configurable.

III. Implementation Analysis

A. LTE Network Architecture

The LTE network architecture (non-roaming) including the network elements and the standardized interfaces [12] is presented in Fig. 4. The LTE network is comprised of the EPC (Evolved Packet Core) and the E-UTRAN (Evolved Universal Terrestrial Radio Access Network). The EPC consists of many logical nodes (S-GW [Serving Gateway], PDN-GW [PDN Gateway], MME [Mobility Management Entity], and PCRF [Policy and Charging Rules Function] etc.), and the E-UTRAN is made up of the eNodeB (evolved NodeB). Each of these network elements is interconnected by means of standardized interfaces (e.g., Rx, Gx, S5, S11, and S1-MME).

B. LTE End-to-end Procedures

LTE end-to-end QoS-related procedures are shown in Fig. 5 [13], [14], and discussed in the following sections. These LTE end-to-end procedures are composed of three major functions: 1) SIP signaling, 2) AF (Application Function, e.g., IMS [IP Multimedia Subsystem]) session establishment/modification, and 3) EPS bearer establishment.

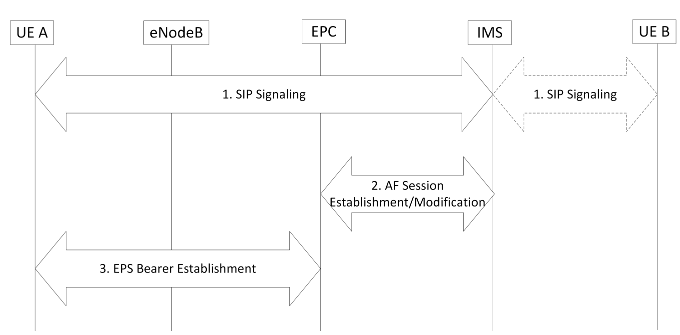

Fig. 5. LTE end-to-end procedures.

C. LTE QoS Related Protocols

From Section III.B, the LTE QoS related protocols are shown as follows:

1) SIP Protocol: The SIP protocol is used to create, modify, and terminate sessions such as Internet multimedia conferences, Internet telephone calls [15]. It uses the SDP (Session Description Protocol) to describe a session.

2) Diameter Base Protocol (Rx and Gx Interfaces): The Diameter base protocol provides an Authentication, Authorization and Accounting (AAA) framework for applications such as network access or IP mobility [14].

3) GTP-C (Control) Protocol (S5 and S11 Interfaces): The control plane of the GPRS Tunneling Protocol (GTP) is responsible for creating, maintaining and deleting tunnels on Sx (e.g., S5, S11) interfaces [16].

4) S1-AP Protocol (S1-MME Interface): The S1-AP protocol provides the signaling service between the E-UTRAN and the EPC [17].

D. LTE QoS Parameters Mapping

The AF can map from SDP within the AF session signaling to Service Information passed to the PCRF over the Rx interface. The PCRF maps messages from the Service Information received over the Rx interface to the Authorized IP QoS parameters that are passed to the PCEF (Policy and Charging Enforcement Function) in the PDN-GW via the Gx interface. The PCEF maps messages from the Authorized IP QoS parameters received from the PCRF to the access specific QoS parameters, which are the QoS parameters that the MAC layer can access [18].

E. User-Specific QoS Parameter Acquisition

There are two methods to acquire the user-specific QoS parameters to be used by the user-specific QoS aware MAC schedulers. The first one is to obtain the user-specific QoS parameters dynamically through the signaling messages (i.e., SIP, Diameter protocol etc.) that are delivered to the MAC layer. The other is to acquire the user-specific QoS parameters through the SPR (Subscriber Profile Repository) database in the PCRF that are delivered to the MAC layer. The difference between these two methods is in how the PCRF obtains the user-specific QoS parameters. After the PCRF acquires the QoS parameters, the subsequent procedures will be the same so that the pertinent QoS parameters are conveyed to the MAC layer.

For the first method, as noted above, the user-specific QoS parameters are obtained by the PCRF through signaling from the SIP and Rx interface protocols.

For the second method, no special SIP signaling is required before the PCRF sends the QoS parameters further to the PCEF through the Gx interface. In most commercial systems, the network operator can obtain the user-specific QoS requirements that are based primarily upon age. When users subscribe to a service from the network operator, they often provide their relevant information such as age, name that can be used by the network to derive the user-specific QoS parameters. To be more specific, when a bearer is to be established or modified, the PCRF inquires of the SPR database about the relevant information of this user. If the relevant information shows that this user is older than a given age (e.g., 55), this user will be considered as a user with a lower sensitivity factor; otherwise, they are regarded as a normal user.

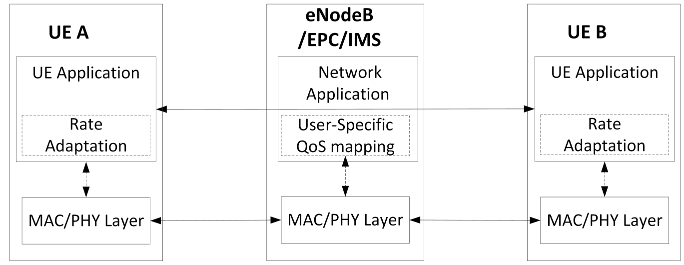

Fig. 6. System architecture.

F. System Architecture

The system architecture and interfaces based on the LTE system are illustrated in Fig. 6, where only relevant modules are shown [19]. In order to implement the user-specific QoS aware schedulers, the AMR mode adaptation algorithm would be implemented in the Rate Adaptation module, whereas the user-specific MAC scheduling algorithms would be implemented in the eNodeB MAC layer. These are software only changes and can be readily accommodated in future versions of LTE [or in future 5G systems].

G. Optimization Process

As an example of the optimization process, when a voice session is to be initialized through the SIP protocol [20], the sender and receiver UE applications will negotiate with each other the application level QoS parameters such as supported AMR codec modes through the IMS [21]. User-specific QoS parameters could also be sent to the IMS by UE.

Next, user-specific QoS requirements will be mapped from the subscription database in the network (e.g., SPR) [2], [18] or user-specific QoS parameters obtained from UE during session initiation, as described in Section III.E. The user-specific QoS parameters shall be delivered to the MAC/PHY Layer in the eNodeB by the EPC/IMS and used to perform the user-specific QoS aware scheduling.

Finally, after the process of rate adaptation, the receiver UE will send the rate control command (e.g., CMR [Codec Mode Request] for VoIP) to the sender UE through the RTP or RTCP protocol [21]–[23] if the data rate is to be changed. Meanwhile the MAC layer in the eNodeB will perform the user-specific QoS aware scheduling.

It should be noted that, due to the changing channel environment for each user and varying network condition, the optimization process should be dynamic and periodic to achieve the maximum system performance gain.

H. Protocol Adaptation

Based upon the analysis above, the following protocol adaptation is proposed to support the user-specific QoS aware scheduling. As described above in Sections III.A-G, the

TABLE I. USER-SPECIFIC SDP MEDIA TYPE DEFINITION

User-specific QoS	Media Type
Audio (low sensitivity factor)	100 (0 + 100)
Audio (normal sensitivity factor)	0
Audio (high sensitivity factor)	200 (0 + 200)

Gx, S5/S11, and S1-MME interface protocols need to be adapted for the second user-specific QoS parameters acquisition method. For the first acquisition method in addition SIP and Rx interface protocol adaptation is needed. Moreover, the RTCP protocol that is used to convey the rate control command also needs to be analyzed to support the user-specific QoS aware scheduling.

1) SIP Protocol: When dynamic user-specific QoS information needs to be conveyed from the UE to the IMS, the SIP protocol [15], [20] needs to be adapted accordingly. The body of a SIP message contains a description of the session, encoded in SDP [24]. An SDP session description consists of a session-level section followed by zero or more media-level sections. Each media-level section starts with an "m=" line. The "m=" line is defined as follows:

m=<media><port><proto><fmt>

<media> is the media type. Currently defined media are "audio", "video", "text", "application", and "message".

So if user-specific QoS requirements need to be conveyed from the UE to the IMS, one method is to implicitly convey the user-specific QoS requirement through the media type field. A user-specific Audio type can be added and defined, e.g., 100 indicates the user-specific Audio media type with a lower sensitivity factor, and 200 indicates the user-specific Audio media type with a higher sensitivity factor as described in Table I. If new media types are defined this way, correspondingly, in the network and the peer UE, the media type field needs to be parsed differently.

2) RTCP Protocol: When an adapted data rate mode needs to be signaled from the receiver UE to the sender UE, it uses the RTCP protocol. The current RTCP protocol supports the rate adaptation signaling, so it can be reused without any modification [21].

3) Rx Interface: The Media-Component-Description AVP (Attribute Value Pair) is conveyed in the Diameter AAR message, and it contains Service Information for a single media component within an AF session [25]. If the user-specific QoS parameters need to be conveyed from the UE, the Media-Component-Description AVP definition needs to be modified. The media-type field in the Media-Component-Description AVP can be used to convey the user-specific QoS requirements as Section 1) above on the SIP protocol describes. Similarly the PCRF needs to parse the media-type field differently according to Table I.

4) Gx Interface: The PCRF may provide authorized QoS information to the PCEF after using the mapping rules to map from the Service Information to the authorized QoS information. The authorized QoS information shall be provisioned within a CCA or RAR Diameter message as QoS-Information AVP. The provisioning of the authorized QoS (which is composed of QCI, ARP and bitrates) is performed from the PCRF to the PCEF [26].

TABLE II. MAPPING FROM USER-SPECIFIC QOS TO QCI

User-specific QoS	QCI
Voice (low sensitivity factor)	101 (1 + 100)
Voice (normal sensitivity factor)	1
Voice (high sensitivity factor)	201 (1 + 200)

In the PCRF, the QCI field needs to be derived based upon the SPR database or the Service Information obtained from the AF through the Rx interface. If the user-specific QoS information is conveyed from the Rx interface, the PCRF can derive the QCI value according to the media type field in the Service Information. If the user-specific QoS information is not conveyed from the Rx interface, the PCRF can use the data from the SPR database to derive the user-specific QoS parameters as shown in Section III.E. Specifically, since the QCI values 0, 10-64, 67-68, and 71-255 are reserved for future use [2], the basic QCI value (i.e., the QCI value derived when no user specific QoS requirements are considered) plus 100 can be used to denote the user-specific QoS with a lower sensitivity factor, while the basic QCI value plus 200 can be used to denote the user-specific QoS with a higher sensitivity factor. The mapping from the user-specific QoS information to the QCI value for VoIP is shown in Table II.

5) S5/S11 Interface: The Create Bearer Request message shall be sent on the S5 interface by the PDN-GW to the S-GW and on the S11 interface by the S-GW to the MME as part of the EPS Bearer establishment procedure [16]. The Bearer Quality of Service (Bearer QoS) is transferred via GTP tunnels through the Create Bearer Request message, where the QCI field has been redefined and added additional user-specific QoS values as described above in Section 4) on the Gx interface, the QCI doesn't need any further modification except different parsing according to Table II in the respective protocols. The PDN-GW and S-GW only needs to forward the Bearer QoS information to the subsequent nodes of S-GW and MME respectively.

6) S1-MME Interface: The E-RAB Setup Request Message is sent by the MME to request the eNodeB to assign resources on Uu and S1 interfaces for one or several E-RABs [17]. The E-RAB Level QoS Parameters are conveyed in the E-RAB Setup Request Message, where the QCI has been redefined and added additional user-specific QoS information as described above in Section 4) on the Gx interface. This message doesn't need any further modification except different parsing according to Table II in the respective protocols. Finally the MAC layer can make user of this user-specific QoS information to perform a more advanced resource scheduling, i.e., user-specific QoS aware scheduling.

I. Scheduling Period

We assume the scheduling period of the rate adaptation algorithms is the frame period of applications, that is, 20 ms for VoIP AMR applications. It is necessary to explore what the system capacity gain will be if the scheduling period is increased in a tradeoff for the reduced complexity. The simulation results in Section IV are shown when the scheduling period is increased from 20 to 2000 ms for VoIP users.

2015 IEEE Wireless Telecommunications Symposium (WTS)
New York City, USA
April 15-17, 2015

TABLE III. SYSTEM SIMULATION CONFIGURATION

Parameter	Assumption
Cellular layout	1 macrocell
Cell radius	1 kilometer
Path loss model	3GPP suburban macrocell
Mobility model	Random Way Point (30/60 km/h)
Carrier frequency	Uplink:1920MHz Downlink:2110MHz
System bandwidth	10MHz
Channel model	ITU Vehicle A
Total BS TX power	40dBm
UE power	23dBm
VoIP codec modes	AMR12.2K, AMR10.2K, AMR7.95K
Number of users	54 VoIP users
Scheduler	Dynamic scheduling USQA-M, USQA-MC scheduler and Baseline scheduler
Other assumptions	Ideal uplink receiver (no block error and packet loss)

TABLE IV. SYSTEM SIMULATION CASES

Cases	Assumption	USQA scheduler
Case 1	54 VoIP users (18 users' α = 0.8, 18 users' α = 1.0, 18 users' α = 1.2), 30 km/h.	USQA-M scheduler
Case 2	Same as Case 1.	USQA-MC scheduler
Case 3	54 VoIP users (α = 0.8), 30 km/h.	USQA-MC scheduler
Case 4	54 VoIP users (α = 0.8), 60 km/h.	USQA-MC scheduler

IV. SYSTEM SIMULATION

A. System Simulation Configuration

The system simulation was run using the OPNET 17.5 Modeler with the LTE modules. The system simulation configuration is partly based upon LTE macrocell system simulation baseline parameters [27] as shown in Table III. the simulation was performed to evaluate the downlink scheduling, with an ideal uplink receiver.

B. System Simulation Cases

Four cases were simulated as described in Table IV. Cases 1 and 2 are used to evaluate the performance of the USQA-M and USQA-MC schedulers respectively, where 54 VoIP users have different sensitivity factors α (18 users' $\alpha = 0.8$, 18 users' $\alpha = 1.0$ and 18 users' $\alpha = 1.2$). Cases 3-4 are used to test the scheduling period for cases of vehicular speeds of 30 km/h and 60 km/h respectively.

C. System Simulation Results

In this paper, the downlink MAC throughput is used to derive the approximate system capacity improvement. System capacity improvement is measured by the increase in the maximum supportable number of users by the system. A rough mapping from the downlink MAC throughput to the system

TABLE V. AVERAGE MOS VALUE

Cases	Scheduler	User category	MOS	MOS improvement
Case 1	USQA-M	User(1.2) User(1.0) User(0.8)	3.86 3.84 3.87	9% -1% -2%
	Baseline	User(1.2) User(1.0) User(0.8)	3.54 3.88 3.95	N/A N/A N/A
Case 2	USQA-MC	User(1.2) User(1.0) User(0.8)	3.75 3.83 3.90	6% -1.3% -1.3%
	Baseline	User(1.2) User(1.0) User(0.8)	3.54 3.88 3.95	N/A N/A N/A

TABLE VI. SYSTEM CAPACITY COMPARISON

Cases	Scheduler	VoIP MAC throughput (Mbps)	Capacity improvement
Case 2	USQA-MC Baseline	1.000 1.045	4.5% N/A

capacity improvement can be done based upon (7).

$$\frac{1/MAC\ throughput\ for\ USQA\text{-}MC}{1/MAC\ throughput\ for\ Baseline} - 1 \qquad (7)$$

The simulation results for MOS value and capacity are shown in Table V and VI respectively.

For Case 1, the average MOS of VoIP users with $\alpha = 1.2$ is increased by ˜9%, whereas the average MOS of VoIP users with $\alpha = 0.8$ is decreased a little.

For Case 2, the system capacity is increased by ˜4.5%, whereas the average MOS of VoIP users with $\alpha = 1.2$ is increased by ˜6%. The MOS gain is not as good as that of Case 1. The reason is that users with $\alpha = 0.8$ are scheduled as normal users in the MAC scheduler so that they have a normal scheduling priority to compete for resources with users with $\alpha = 1.2$. In this case, only 1/3 of the users have sensitivity factors of 0.8. As this ratio increases, the system capacity improvement gain will further be increased, as verified in Cases 3-4 where all users have a sensitivity factor of $\alpha = 0.8$.

Figure 7 show the VoIP capacity as a function of scheduling period from 20 ms to 2000 ms. From Fig. 7, we find that as the scheduling period increases, the performance gain will decrease correspondingly. As the scheduling period increases to 1000 ms, the capacity improvement will fall below 10% for the case of 60 km/h, whereas the capacity improvement is still good for the case of 30 km/h.

V. CONCLUSIONS

In this paper, we introduce the concept of user-specific QoS requirements and demonstrate its importance in spectral allocation. Two user-specific QoS aware schedulers are proposed aimed at improving the MOS or both the MOS and the system capacity in wireless systems for VoIP applications. Detailed system implementation analysis based upon LTE systems is performed to show that very modest modifications on current LTE protocols are needed to support the user-specific QoS

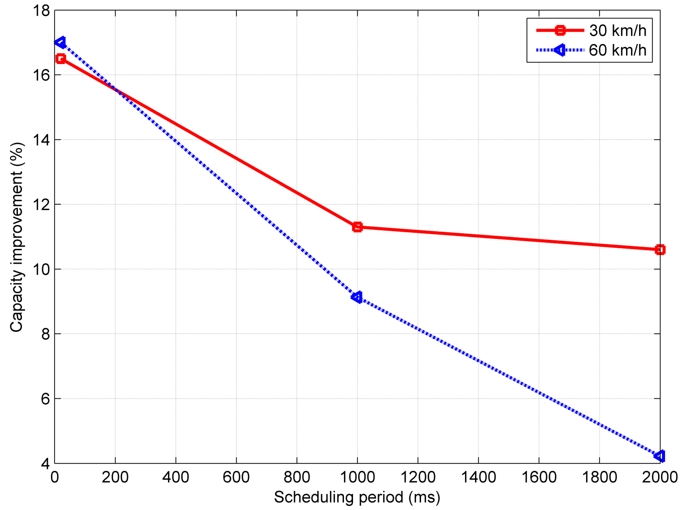

Fig. 7. VoIP Capacity improvement as a function of scheduling period.

aware scheduling. Simulation results demonstrate that appreciable MOS improvement or/and capacity improvement can be observed for these two schedulers, and also show that the scheduling period of rate adaptation algorithms of up to 1000 ms doesn't significantly impact the system performance.

ACKNOWLEDGMENT

This research was supported by NSF Grant 1352883. We would like to thank Drs. Huseyin Arslan and Zygmunt Haas for many stimulating and enlightening discussions during this research. We would like to thank Riverbed for providing its NetModeler Suite of simulation software.

REFERENCES

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

[2] 3GPP, "TS 23.203 - v12.6.0 - digital cellular telecommunications system (phase 2+); universal mobile telecommunications system (UMTS); LTE; policy and charging control architecture."

[3] J. D. Pearson, C. H. Morrell, S. GordonSalant, L. J. Brant, E. J. Metter, L. L. Klein, and J. L. Fozard, "Gender differences in a longitudinal study of ageassociated hearing loss," *J. Acoust. Soc. Am.*, vol. 97, no. 2, pp. 1196–1205, Feb. 1995.

[4] C. He and R. D. Gitlin, "User specific QoS and its application in resources scheduling for wireless system," in *ICA3PP 2014*, pp. 809-821, Aug. 2014.

[5] C. He and R. D. Gitlin, "Application-specific and QoS-aware scheduling for wireless systems," in *IEEE PIMRC 2014*, pp. 1147–1151, Sep. 2014.

[6] C. He, G. E. Arrobo, and R. D. Gitlin, "Improving system capacity based upon user-specific QoS for heterogeneous networks," in *IEEE WCNC 2015*, Mar. 2015.

[7] ITU-T, "ITU-T G.107-the e-model: a computational model for use in transmission planning," Dec. 2011.

[8] 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.

[9] F. Mertz and P. Vary, "Efficient voice communication inWireless packet networks," in *2008 ITG Conference on Voice Communication (SprachKommunikation)*, Oct. 2008, pp. 1–4.

[10] 3GPP, "TS 26.071 - v11.0.0 - mandatory speech CODEC speech processing functions;AMR speech CODEC; general description."

[11] ITU-T, "ITU-t g.114-one-way transmission time," May 2003.

[12] 3GPP, "TS 23.401 - v10.10.0 - LTE; general packet radio service (GPRS) enhancements for evolved universal terrestrial radio access network (e-UTRAN) access."

[13] 3GPP, "TS 24.301 - v12.6.0 - universal mobile telecommunications system (UMTS); LTE; non-access-stratum (NAS) protocol for evolved packet system (EPS); stage 3."

[14] J. Arkko, E. Guttman, P. R. Calhoun, and J. Loughney, "Diameter base protocol." [Online]. Available: https://tools.ietf.org/html/rfc3588

[15] G. Camarillo, M. Handley, J. Peterson, J. Rosenberg, A. Johnston, H. Schulzrinne, and R. Sparks, "SIP: Session initiation protocol." [Online]. Available: http://tools.ietf.org/html/rfc3261

[16] 3GPP, "TS 29.274 - v12.6.0 -universal mobile telecommunications system (UMTS); LTE; 3gpp evolved packet system (EPS); evolved general packet radio service (GPRS) tunnelling protocol for control plane (GTPv2-c); stage 3."

[17] 3GPP, "TS 36.413 - v10.6.0 - LTE; evolved universal terrestrial radio access network (e-UTRAN); s1 application protocol (s1ap)."

[18] 3GPP, "TS 29.213 - v11.8.0 - digital cellular telecommunications system (phase 2+); universal mobile telecommunications system (UMTS); LTE; policy and charging control signalling flows and quality of service (QoS) parameter mapping."

[19] 3GPP, "TS 36.300 - v12.2.0 - evolved universal terrestrial radio access (e-UTRA);overall description; stage 2."

[20] 3GPP, "TS 24.229 - v10.12.0 - digital cellular telecommunications system (phase 2+); universal mobile telecommunications system (UMTS); LTE;IP multimedia call control protocol based on session initiation protocol (SIP) and session description protocol (SDP); stage 3."

[21] 3GPP, "TS 26.114 - v10.3.0 - universal mobile telecommunications system (UMTS); LTE; IP multimedia subsystem (IMS); multimedia telephony; media handling and interaction."

[22] M. Westerlund, A. Lakaniemi, J. Sjoberg, and Q. Xie, "RTP payload format and file storage format for the adaptive multi-rate (AMR) and adaptive multi-rate wideband (AMR-WB) audio codecs." [Online]. Available: https://tools.ietf.org/html/rfc4867

[23] 3GPP, "TS 26.101 - v12.0.0 - digital cellular telecommunications system (phase 2+); universal mobile telecommunications system (UMTS); LTE; mandatory speech codec speech processing functions; adaptive multi-rate (AMR) speech codec frame structure."

[24] M. Handley, C. Perkins, and V. Jacobson, "SDP: Session description protocol." [Online]. Available: http://tools.ietf.org/html/rfc4566

[25] 3GPP, "TS 29.214 - v12.5.0 -universal mobile telecommunications system (UMTS); LTE; policy and charging control over rx reference point."

[26] 3GPP, "TS 29.212 - v12.6.0 -universal mobile telecommunications system (UMTS); LTE; policy and charging control (PCC); reference points."

[27] 3GPP, "TS 25.814 - v7.1.0 - physical layer aspects for evolved universal terrestrial radio access (UTRA)."