

# WE-MQS: A new LTE Downlink Scheduling Scheme for Voice Services based on User Perception

Duy-Huy Nguyen  
SAMOVAR, Télécom  
SudParis, CNRS, Université  
Paris-Saclay  
9 rue Charles Fourier – 91011  
Evry Cedex

Hang Nguyen  
SAMOVAR, Télécom  
SudParis, CNRS, Université  
Paris-Saclay  
9 rue Charles Fourier – 91011  
Evry Cedex

Éric Renault  
SAMOVAR, Télécom  
SudParis, CNRS, Université  
Paris-Saclay  
9 rue Charles Fourier – 91011  
Evry Cedex

## ABSTRACT

The Long Term Evolution (LTE) is a high data rates and fully All-IP network. Scheduling and resource allocation are very important tasks in LTE network. So that, the optimization of the real-time services transmission is very essential. However, there are many challenges when deploying the live multimedia services such as VoIP, Video, etc. This paper proposes a new scheduling scheme which based on the Wideband (WB) E-model, Channel- and QoS-Aware (known as WE-MQS scheduler) for voice traffic in LTE downlink direction. The voice traffic flow is very sensitive to network impairments such as delay, packet loss, jitter, etc. The proposed scheduling scheme is based on the extension of the WB E-model and the consideration of Maximum Queue Size (MQS) as a factor for the metric. Since this scheduling scheme considers Mean Opinion Score (MOS) values, thus, it gets higher user perception. The simulation results show that the proposed scheme has the performance which not only satisfies QoS requirements of voice services but also outperforms well-known schedulers such as Frame Level Scheduler (FLS), Modified Largest Weighted Delay First (M-LWDF) and Exponential/Proportional Fair (EXP/PF) schedulers in terms of delay for all the number of user (NU) and Packet Loss Rate (PLR) when the NU is more than 47. For the cell throughput, Fairness Index (FI), and Spectral Efficiency, the proposed scheduler is always in the middle of the remaining schedulers. The performance evaluation is compared in terms of Delay, PLR, Throughput, FI, and FI for FLS, M-LWDF, EXP/PF schedulers and the proposed one.

## Keywords

Scheduling scheme, user perception, LTE, WB E-model, VoLTE, WB-MQS.

## 1. INTRODUCTION

In the digital device market, there are various smart mobile devices such as iPhone, iPad, Android, etc. which has enough powerful to support a wide range of multimedia communications such as VoIP, Video, Gaming, etc. This means there need be a high data rate network to well support these services. LTE is developed by the Third Generation Partnership Project (3GPP) [1]. It 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 Quality of Service (QoS) for various multimedia applications. In order to support inter-symbol interference and selecting fading, LTE uses Orthogonal Frequency Division Multiple Access (OFDMA) in the downlink. Basic components of LTE network include a powerful eNodeB (eNB) station and several User equipments (UEs) in addition to a gateway [2]. The eNB station 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 Radio Resource Management (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 [3].

LTE architecture defines a complete table with packet delay budget and packet loss rates for performing different service class in the network [1] which some or all may be executed by service provider. This table divides resource types in the LTE network into two groups, those are Guaranteed Bit Rate (GBR) and Non-GBR. Table 1 represents the service classes in LTE network where voice service is a Guaranteed Bit Rate (GRB) service which has the second priority just after IP Multimedia Subsystem (IMS) signaling. However, in order to guarantee voice over LTE (VoLTE) quality is an extreme challenge. In order to meet different QoS requirements for these groups, several packet scheduling algorithms have been proposed. According to [4], the scheduling strategies for LTE downlink are divided into five groups including: (1) Channel-unaware strategies, (2) Channel-aware/QoS-unaware strategies, (3) Channel-aware/QoS-aware strategies,

Table 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)
	9		$10^{-6}$	TCP-based (e.g. www, e-mail, chat, FTP, P2P sharing, progressive video, etc.)

(4) Semi-persistent scheduling for VoIP support, and (5) Energy-aware strategies. For voice traffics which are very

sensitive to delay and PLR. So that, the Channel-aware/QoS-aware strategies are very essential for them. Several well-known scheduling algorithms for this group as FLS [5], M-LWDF [6], and EXP/PF [7]. In these schedulers, there is only the FLS which guarantees bounded delay for real-time flows, the remaining schedulers transmits users data in a given Transmission Time Interval (TTI) by assigning a computed priority metric that is specific to the scheduler functionality. However, due to the lack of delay budget and PLR attribute, they are not suitable for supporting simultaneously real-time and non real-time traffic [2]. In this paper, we propose a new downlink scheduling scheme for voice services in LTE network with the consideration of user perception. There are several papers which mentioned it. Specifically, authors in [8], [9] proposed a new Quality of Experience (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 users to access a cell. In [10], the authors a cross-layer design scheme that jointly optimizes three different layers of wireless protocol stack, namely Application, Medium Access Control (MAC) and Physical layers. The purpose of them is to maximize network resource utilization and user-perceived quality of service (also called Quality of Experience - QoE). These papers proposed a new LTE Downlink Scheduler but not based on the WB E-model. The paper which is closest to our paper described in [11]. 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 on the E-model. In addition, in the metric of the scheduler, there is no the presence of the MQS, and the authors evaluated only for VoIP traffic. In this paper, we extend the idea in [11] by using the WB E-model and propose to consider the MQS as an essential and effective factor for the metric. We used the WB E-model to predict the MOS score and use this score as a main factor in the metric. Besides, we see that, the MQS factor has significant effects on the system performance. In the LTE-Sim [12], this factor is fixed equal to 0. This means the MQS is infinite. So that, in the scheduling process, the MQS is not considered. 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 a essential factor in the metric of the scheduling algorithms.

The proposed scheduler selects UEs that based on their priorities which are computed according to the following factors: the maximum MOS, the minimum remaining queue size, the maximum delay, the channel condition. This means for the UE which has the higher MOS, the lower remaining queue size, the higher maximum delay, and the higher channel condition will have the higher priority. We assess the performance of the proposed scheduler with the FLS, M-LWDF and EXP/PF schedulers for a heterogeneous traffic including VoIP, Video, and non real-time service which is called INF-BUF user. The simulation results were implemented in the open source LTE system simulator called Long Term Evolution- Sim (LTE-Sim) [12] and are compared in terms of Delay, PLR, Cell throughput, and FI for the number of user from 10 to 50.

The rest of this paper is organized as follows: Overview of the system model is described in section 2. In section 3, we present the proposed scheduling scheme. The simulation results and performance evaluation of the proposed

scheduler are analysed in section 4. The conclusion and future work is represented in section 5.

## 2. THE SYSTEM MODEL

### 2.1 VoLTE traffic flow

#### 2.1.1 Protocol stack

In LTE network, the speech frame is packetized sequentially with network protocols, including Real-time Transport Protocol (RTP), User Datagram Protocol (UDP) and Internet Protocol (IP). And then, it will be encapsulated with other radio protocols such as Packet Data Convergence Protocol (PDCP), Radio Control Link (RLC) and Medium Access Control (MAC). All of these protocols will add their headers into the packetized speech packet.

In order to decrease data overhead of above protocols headers when the speech packet transmitted over a radio channel, Robust Header Compression (RoHC) is used. This will save bandwidth and enhance voice transmission over LTE network. RoHC compresses IP header from 40 bytes (with IPv4) and 60 bytes (with IPv6) down to 1 to 4 bytes [13], normally is 3 bytes. For voice traffic over LTE network, HARQ (Hybrid Automatic Repeat Request) technique is used at MAC layer for retransmission if FEC fails error correction. Each speech packet will be retransmitted at least from one to three times. The retransmission times depends on the error correction or the configured maximum times of retransmission.

The implemented model of VoLTE protocol stack used in LTE-Sim [12] is represented on Figure 1:

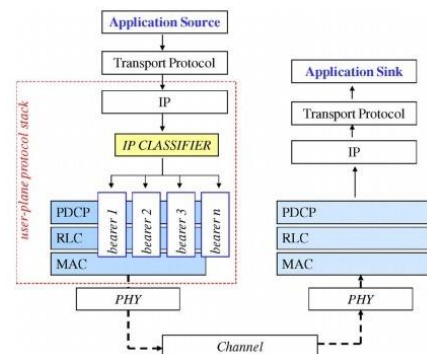


Fig. 1. Implemented model of VoLTE protocol stack used in LTE-Sim software

#### 2.1.2 Source codec

VoLTE uses AMR-WB as a vocoder. 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 detailed described in [14]. AMR-WB codec uses a sampling rate of 16 kHz, which covers 50-7000 Hz audio band- width. It has 9 different codec modes (from mode 0 to mode 8) corresponding to 9 source bit rates in range of 6.6-23.85 Kb/s. Each of them generates 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 bit rate (mode) such as described in Table 2.

**Table 2. Packet sizes of AMR-WB codec modes**

Parameter	AMR-WB bit rates (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
<b>Packet size (bits)</b>	<b>584</b>	<b>568</b>	<b>504</b>	<b>472</b>	<b>424</b>	<b>392</b>	<b>360</b>	<b>288</b>	<b>240</b>

In LTE network, AMR-WB codec is configured into 3 configurations [15] 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.

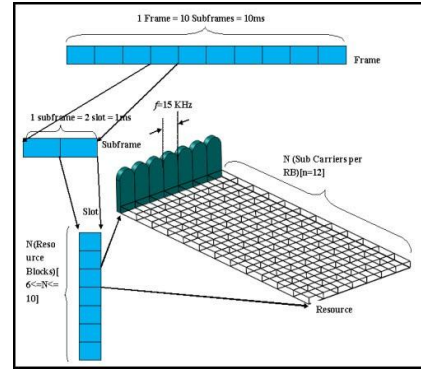
These configurations are used to simplify the negotiation of bit rate between the user equipment and the base station, thus will simplify the implementation and testing. The remaining bit rates can still be used for other purposes in mobile networks. In order to choose a bit rate, the receiver measures quality of radio channel. The channel quality indicator (CQI) is used for this purpose. It is defined as an equivalent carrier-to-interference (C/I) ratio. The C/I ratio 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.

### 2.1.3. LTE frame structure

Figure 2 represents the structure of LTE frame for the downlink air interface [16]. In LTE downlink, a frame is 10 ms in length and is divided into 10 sub-frames in time domain. So that, each sub-frame is 1 ms in length and is divided into 2 slots where each slot is 0.5 ms in length. In frequency domain, each slot is divided into a number of resource blocks (RBs) (from 6 to 10 RBs). Each slot contains 6 or 7 Orthogonal Frequency Division Multiplexing (OFDM) symbols in normal cyclic prefixes and extended cyclic prefixes, respectively. Each time slot in frequency domain is divided into bands of 180 kHz which contains 12 consecutive sub-carriers. Each RB is a basic unit of exchanging information in the downlink direction of LTE system. So that, a RB is the radio resource that is available for a user in LTE network and is defined by both frequency and time domains. The number of RBs in a slot depends on the bandwidth of the LTE system [17]. Each sub-frame represents a TTI which is the minimum transmission unit. Each TTI contains at least one transport block per UE. The RB size is the same one for all bandwidths [18].

## 2.1 WB E-model: Speech quality assessment for Wideband Audio

WB E-model is a computational model developed and standardized by ITU-T [19]. It is used to estimate the MOS for wideband audio quality. The output of the model is R-factor. The values of this R-factor in range of 0-129. And



**Fig. 2. Resource grid**

then, it is mapped to the MOS. The R-factor in WB E-model is defined as follows:

$$R_{wb} = R_{0,wb} - I_{s,wb} - I_{d,wb} - I_{e,eff,wb} + A \quad (1)$$

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  $R_{wb}$  factor is translated into the MOS as follows [19]: For  $R = R_{wb}/1.29$

$$MOS = \begin{cases} 1, & \text{if } R < 0 \\ 1 + 0.035 \times R + 7 \times 10^{-6} \times R \times (R - 60) \times (100 - R), & \text{if } 0 \leq R \leq 100 \\ 4.5, & \text{otherwise} \end{cases} \quad (2)$$

The relation between R-factor, user perception, and MOS is described in the Table 3.

**Table 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

$R_{wb}$  factor is then mapped to the MOS using Equation (2), and then, the MOS is mapped to the satisfaction level of the users. According to [19], for the wideband audio, the value of  $R_{0,wb}$  factor in equation (1) equals 129, thus, equation (1) can be rewritten as follows:

$$R_{wb} = 129 - I_{d,wb} - I_{e,eff,wb} \quad (3)$$

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 by the following equation [20]:

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

$$H(D_{e2e} - 177.3) \quad (4)$$

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

$$H(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{otherwise} \end{cases} \quad (5)$$

In equation (4),  $D_{e2e}$  represents the total end-to-end delay of speech packet. It is calculated via some functions in LTE-Sim software. The  $I_{e,eff,wb}$  is determined according to packet loss. According to [21],  $I_{e,eff,wb}$  is determined as follows:

$$I_{e,eff,wb} = I_{e,wb} + (129 - I_{e,wb}) \times \frac{P_{pl}}{P_{pl} + B_{pl}} \quad (6)$$

Where:  $I_{e,wb}$ : The respective impairment factor without any packet loss.  $P_{pl}$ : Packet loss rate. It is also calculated via some functions in LTE-Sim software.  $B_{pl}$ : A codec-specific factor which characterizes its robustness against packet loss. The values of  $I_{e,wb}$ ,  $B_{pl}$  and the relationship of them to the  $P_{pl}$  is represented on Figure 3 [21].

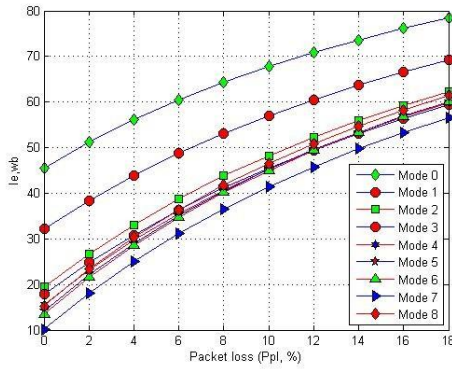


Fig. 3.  $I_{e,wb}$  vs.  $P_{pl}$  for nine modes of AMR-WB codec

The R-factor is then mapped to the MOS via equations (2). MOS is one of the important factors for the metric in the proposed scheduler.

## 2.2 The correlated Scheduling algorithms

The purpose of scheduling algorithm in LTE is to maximum system performance [22]. 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 [22]. The eNB offers a scheduling decision based on the CQI which is sent 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 obtain the metric, the scheduler usually need 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 [12]:

$$\bar{R}_i(k) = 0.8 \times \bar{R}_i(k-1) + 0.2 \times \bar{r}_i(k) \quad (7)$$

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 the proposed scheduling scheme including: FLS, M-LWDF, and EXP/PF schedulers. We select these scheduling

because they perform well and are suitable for voice services.

### 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 communicate 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:

$$V_i(k) = h_i(k) * q_i(k) \quad (8)$$

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  $h_i(k)$ .

### 2.2.2 The M-LWDF scheduler

M-LWDF scheduling algorithm is used to support multiple real-time services in CDMA-HDR systems [6]. For each real-time flow, by considering the maximum time  $\tau_i$ , the probability is defined as the maximum probability  $\delta_i$  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:

$$w_{i,j} = \alpha_i \times D_{HOL,i} \times \frac{r_{i,j}}{R_i} \quad (9)$$

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

### 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 [7]. 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:

$$w_{i,j} = \exp\left(\frac{\alpha_i \times D_{HOL,i} - X}{1 + \sqrt{X}}\right) \times \frac{r_{i,j}}{R_i} \quad (10)$$

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.

### 3. THE PROPOSED SCHEDULING SCHEME

In the proposed scheduling scheme, we consider the characteristics of VoIP service. This service is 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  $\alpha_i$ ,  $r_{i,j}$  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, the higher user satisfaction. MOS needs 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 [12], 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 of the scheduling algorithms. The lack of LTE-Sim software is that it supports only G.729 codec while VoLTE uses AMR-WB codec. G.729 has only a 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 AMRWB are changed according to channel condition (i.e. C/I ratio). So, the user perception is calculated at scheduled instant. In the proposed scheduling scheme, we proposed to complement AMR-WB codec into LTE-Sim by reconfiguring some parameters and modifying essential source codes. With the presence of AMR-WB, we can simulate VoLTE traffic more easily.

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 MOS and the lower ( $Q_{i,max} - Q_i$ ) values, the higher priority for the UE. The fixed maximum time  $D_{HOL,i}$  and the maximum probability  $\delta_i$  are included in the Equations (4) and (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:

$$w_{i,j} = \frac{MOS_i \times (Q_{i,max} - Q_i)}{\tau_i} \times \frac{r_{i,j}}{\bar{R}_i} \quad (11)$$

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 [12]. 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  $\tau_i$  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 the user perception. In fact, the AMR-WB mode is dynamically calculated and optimized at the AMR-WB encoder according to channel quality using rate adaptation control algorithm detailed in [14]. 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 proposed 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 chooses dynamically source mode according to channel quality. The threshold values of C/I ratio is chosen

according to [23] and [24]. The procedures for choosing AMR-WB mode and calculating the metric of the proposed scheduler are described in Algorithm 1 and 2 as follows:

---

**Algorithm 1** Change AMR-WB packet size according to the channel quality for the **Config-WB-Code 0**

---

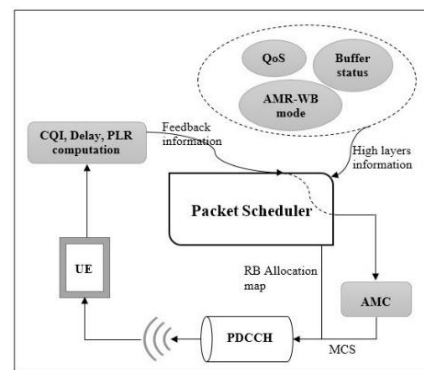
- 1: **procedure** UPDATE AMR-WB PACKET SIZE
  - 2: **Step 1:** Calculate C/I ratio: available in LTE-Sim
  - 3: **Step 2:** Update packet size of AMR-WB codec
  - 4:   **if** calculated C/I  $\leq$  6.5 **then**
  - 5:        $mode = 0$
  - 6:       Update Packet size = 30
  - 7:   **else if** calculated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

**Algorithm 2**  $W_{ij}$  calculation for voice user

---

- 1: **procedure**  $W_{ij}$  CALCULATION
  - 2: **Step 1:** Calculate  $Q_{i,max} - Q_i$ ,  $\tau_i$ ,  $r_{i,j}$ , and  $\bar{R}_i$
  - 3: **Step 2:** Calculate R-factor
  - 4:   Calculate  $I_{d,wb}$  based on Equation (4)
  - 5:   Calculate  $I_{e,eff,wb}$  based on the chosen  $mode$  using Formula (6).
  - 6:   Calculate R-factor based on Equation (3)
  - 7: **Step 3:** Calculate MOS score using Equation (2)
  - 8: **Step 4:** Calculate  $w_{i,j}$  based on Formula (11)
- 

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 model of the proposed scheduling scheme is represented such as on Figure 4.



**Fig. 4. Implemented model of the proposed scheduling scheme**

Figure 4 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 [4]. When 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 Head of Line packet delay (HOL) and the queue length (QL) are estimated. If the HOL packet delay or QL exceeds a specified thresholds

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 the scheduler making the decisions 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, specially user perception. The scheduling process in Figure 4 is described as followings:

- Each UE decodes the reference signals, calculates the Channel Quality Indicator (CQI), delay, PLR 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 Adaptive Modulation and Coding (AMC) module selects the best Modulation and Coding Scheme (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 Physical Downlink Control Channel (PDCCH).
- Each UE reads the PDCCH payload and in case it has been scheduled, accesses to the proper Physical Downlink Shared Channel (PDSCH) payload.

## 4. SIMULATION ENVIRONMENT AND PERFORMANCE EVALUATION

### 4.1 Simulation environment

#### 4.1.1 Traffic model

In our 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 and a Video flow at the same time. For the VoIP flow, a AMR-WB voice stream with bit-rates of 6.6, 8.85, and 12.65 kbps (Configuration A of AMR-WB codec which is mandatory and default for VoLTE service) were considered. The voice flow is a bursty application that is modelled with an ON/OFF Markov chain [25]. 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 [26] and it is also available in [12]. 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 [27].

- Path loss:  $PL = 128.1 + 37.6 \times \log(d)$ , 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.

#### 4.1.2 Simulation parameters

In this paper, we investigate the performance of the FLS, M-LWDF, EXP/PF and the proposed schedulers in LTE downlink direction. The simulation process is performed in a single cell with interference with the number of users in the interval [10, 50] which move randomly at a speed of 30 km/h.

In the simulation scenario, we consider each user using a VoIP, a Video, and a INF-BUF flow. In order to evaluate the system performance, we use LTE-Sim [12] for simulating. This is a open source framework for researchers and academic community. The basic parameters used in the simulation are represented in the Table 4.

Table 4. Simulation parameters

Simulation Parameters	Values
Simulation duration	100 s
Frame structure	FDD
Cell radius	1 km
Bandwidth	10 MHz
Video bit-rate	242 kbps
AMR-WB bit-rates	6.6, 8.85, 12.65 kbps
User speed	30 km/h
Number of users	10, 20, 30, 40, 50 UEs
Maximum delay	0.1 s
MQS	$10^5$ bytes
Traffic model	VoIP, Video, and INF-BUF
Packet Schedulers	FLS, M-LWDF, EXP/PF, and WE-MQS

### 4.2 Performance evaluation

In order to evaluate the performance of the proposed scheduling scheme, we use the LTE-Sim [12] 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.

#### 4.2.1 Delay

End-to-end delay (called one-way delay) is the time required for a packet to be transmitted from source to destination in the network. Figure 5 illustrates the delay of VoIP flow. It is clear that, the proposed 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. For this case, the FLS is quite special and unprecedented. 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.

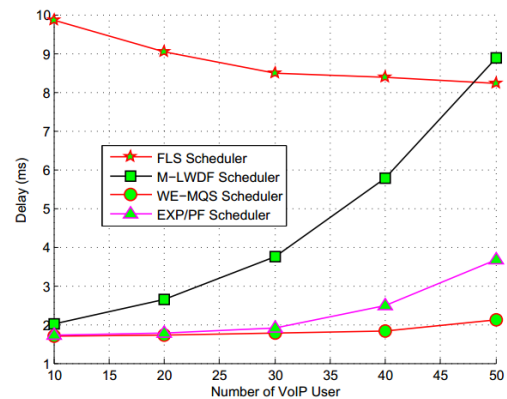


Fig. 5. Delay vs number of VoIP user

#### 4.2.2 Packet Loss Rate

PLR shows the failure of one or more transmitted packets to reach their destination across a network. Figure 6 represents the PLR of VoIP flow. When we set the MQS equal to  $10^5$  bytes (this factor is set equal to 0 as default in the LTESim), for the VoIP flow, all schedulers have the PLR decreasing then increasing when the number of UE increases. Normally, the PLR increases when the number of UE increases, thus, this case is unprecedented. However, it can be not stable in real LTE network because the LTE-Sim is quite similar to a real system. As shown in the Figure 6, all the schedulers have the PLR which are under 1%. This is very good for real-time services. 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 in range of 10..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.

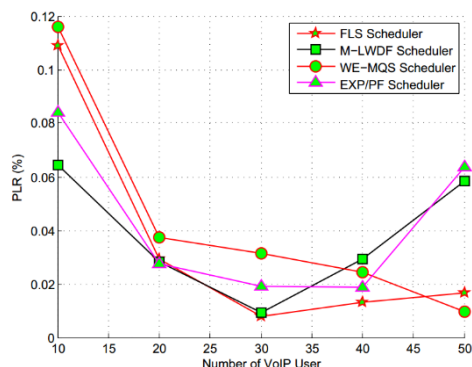


Fig. 6. Packet loss rate vs number of VoIP user

#### 4.2.3 Cell throughput

Throughput is a measurement of how many units of information a system can process in a given amount of time. As shown on Figure 7, for the VoIP flow, the cell throughput of all the schedulers increases when the NU increases. Although the proposed scheduler does not has the highest throughput, but it is always in top 2 or 3 of the highest throughput. There is not any scheduler which always has the highest throughput when the NU increases. The difference of throughput among the schedulers is not significant.

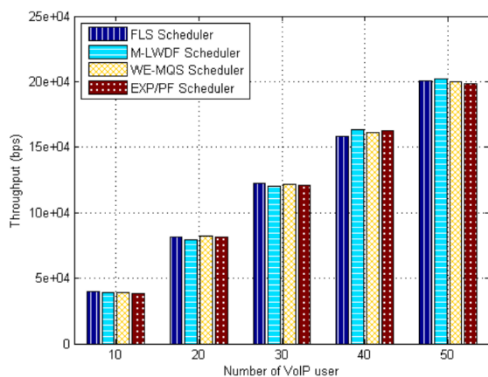


Fig. 7. Throughput vs number of VoIP user

#### 4.2.4 Fair index

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 on Figure 8, 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 throughput, for the FI, the difference among the schedulers is not also significant.

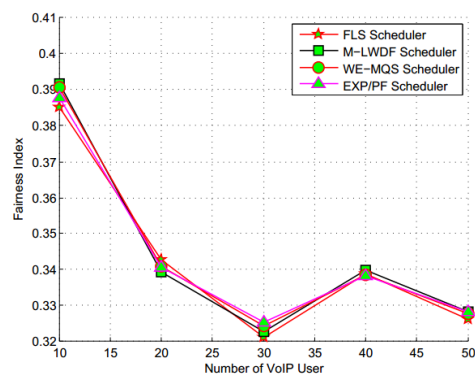


Fig. 8. Fairness Index vs number of VoIP user

#### 4.2.5 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 9, 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 too high. Therefore, all the schedulers are efficient, thus, they are very suitable for VoIP users.

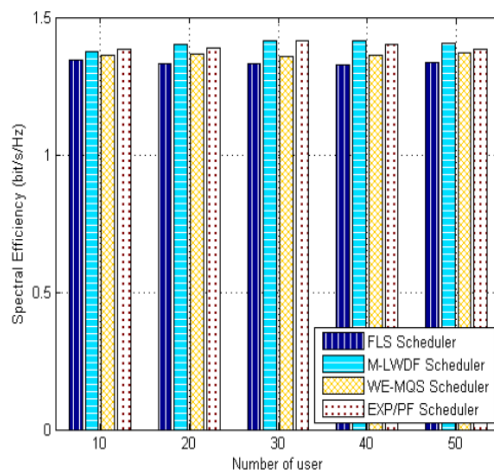


Fig. 9. Spectral efficiency vs number of user

## 5. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a new Channel-, QoS- and QoE-Aware scheduling scheme for voice users in the LTE downlink direction. The main idea in the proposed scheduler is the consideration of user perception (via the MOS score) and the MQS factors into the metric in our proposed

scheduler. The metric is based on the MOS score, the remaining queue size, the fixed maximum time, and the channel condition. 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. This overcomes the limitation of the LTESim that it supports only G.729 codec for VoIP application. It can be said 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 is difficult to say that which scheduler has the best performance for VoIP user. However, it is sure that, for delay, the proposed scheduler always keeps the lowest position for all the NU and this is similar to the case of PLR when the NU is more than 47. Therefore, it can be concluded that the proposed scheduler is also very suitable and efficient for voice services in the LTE downlink direction. For the future work, we will build a framework to measure voice quality and apply VoIP priority mode to the proposed scheduler.

## 6. REFERENCES

- [1] 3GPP. <http://www.3gpp.org>.
- [2] Salman Ali and Muhammad Zeeshan. A utility based resource allocation scheme with delay scheduler for lte service-class support. In *Wireless Communications and Networking Conference (WCNC)*, 2012 IEEE, pages 1450–1455. IEEE, 2012.
- [3] Mehdi Alasti, Behnam Neekzad, Jie Hui, and Rath Vannithamby. Quality of service in wimax and lte networks [topics in wireless communications]. *Communications Magazine*, IEEE, 48(5):104–111, 2010.
- [4] Francesco Capozzi, Giuseppe Piro, Luigi Alfredo Grieco, Gennaro Boggia, and Pietro Camarda. Downlink packet scheduling in lte cellular networks: Key design issues and a survey. *Communications Surveys & Tutorials*, IEEE, 15(2):678–700, 2013.
- [5] Giuseppe Piro, Luigi Alfredo Grieco, Gennaro Boggia, Rossella Fortuna, and Pietro Camarda. Two-level downlink scheduling for real-time multimedia services in lte networks. *Multimedia*, IEEE Transactions on, 13(5):1052–1065, 2011.
- [6] Pablo Ameigeiras, Jeroen Wigard, and Preben Mogensen. Performance of the m-lwdf scheduling algorithm for streaming services in hsdpa. In *Vehicular technology conference, 2004. VTC2004-Fall*. 2004 IEEE 60th, volume 2, pages 999–1003. IEEE, 2004.
- [7] Jong-Hun Rhee, Jack M Holtzman, and Dong-Ku Kim. Scheduling of real/non-real time services: adaptive exp/pf algorithm. In *Vehicular Technology Conference, 2003. VTC 2003-Spring*. The 57th IEEE Semiannual, volume 1, pages 462–466. IEEE, 2003.
- [8] Ali Alfayly, I Mkwawa, Lifeng Sun, and Emmanuel Ifeachor. Qoe-based performance evaluation of scheduling algorithms over lte. In *Globecom Workshops*, 2012 IEEE, pages 1362–1366. IEEE, 2012.
- [9] Ali Alfayly, Is-Haka Mkwawa, Lingfen Sun, and Emmanuel Ifeachor. Qoe-driven lte downlink scheduling for voip application. In *Consumer Communications and Networking Conference (CCNC)*, 2015 12th Annual IEEE, pages 603–604. IEEE, 2015.
- [10] Meysam Nasimi, Mohammad Kousha, and Fazirulhisyam Hashim. Qoe-oriented cross-layer downlink scheduling for heterogeneous traffics in lte networks. In *Communications (MICC)*, 2013 IEEE Malaysia International Conference on, pages 292–297. IEEE, 2013.
- [11] M Sajid Mushtaq, Brice Augustin, and Abdelhamid Mellouk. Qoe-based lte downlink scheduler for voip. In *Wireless Communications and Networking Conference (WCNC)*, 2014 IEEE, pages 2190–2195. IEEE, 2014.
- [12] Giuseppe Piro, Luigi Alfredo Grieco, Gennaro Boggia, Francesco Capozzi, and Pietro Camarda. Simulating lte cellular systems: an open-source framework. *Vehicular Technology*, IEEE Transactions on, 60(2):498–513, 2011.
- [13] Carsten Bormann, C Burmeister, M Degermark, et al. Robust header compression (rohc). Technical report, RFC 3095, June, 2001.
- [14] ITUT 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] 3GPP. Inband Tandem Free Operation (TFO) of speech codecs; Service description; Stage 3. TS 28.062, 3rd Generation Partnership Project (3GPP), 12 2009.
- [16] Richard Musabe and Hadi Larjani. Cross-layer scheduling and resource allocation for heterogeneous traffic in 3g lte. *Journal of Computer Networks and Communications*, 2014.
- [17] Freescale Semiconductor. Long term evolution protocol overview. White Paper, Document No. LTEPTCLOVWWP, Rev 0 Oct, 2008.
- [18] 3GPP. Evolved Universal Terrestrial Radio Access (EUTRA); Physical channels and modulation. TS 36.211, 3<sup>rd</sup> Generation Partnership Project (3GPP), 03 2010.
- [19] ITUT Rec. G. 107.1, wideband e-model. Int. Telecomm. Union, Geneva, 2011.
- [20] Cristian Olariu, Micheal O Foghlu, Philip Perry, and Liam Murphy. Voip quality monitoring in lte femtocells. In *Integrated Network Management (IM)*, 2011 IFIP/IEEE International Symposium on, pages 501–508. IEEE, 2011.
- [21] Sebastian Moller, Alexander Raake, Nobuhiko Kitawaki, Akira Takahashi, and Marcel Waltermann. Impairment factor framework for wide-band speech codecs. *Audio, Speech, and Language Processing*, IEEE Transactions on, 14(6):1969–1976, 2006.
- [22] Dakuri Chiranjeevi, V Mahender, B Kiran, and B Raja Kumar. Channel aware scheduling algorithms for 3gpp lte downlink.



- [23] Agilent Technologies. Available:[http://rfmw.em.keysight.com/rfcomms/refdocs/gsm/default.htm#gprsla\\_amr\\_bse\\_config.html#CBDDEBCD](http://rfmw.em.keysight.com/rfcomms/refdocs/gsm/default.htm#gprsla_amr_bse_config.html#CBDDEBCD), May 2010.
- [24] Agilent Technologies. Available:[http://rfmw.em.keysight.com/rfcomms/refdocs/gsm/gprsla\\_wb\\_amr\\_bse\\_config.html](http://rfmw.em.keysight.com/rfcomms/refdocs/gsm/gprsla_wb_amr_bse_config.html), May 2010.
- [25] Chen-Nee Chuah and Randy H Katz. Characterizing packet audio streams from internet multimedia applications. In *Communications, 2002. ICC 2002. IEEE International Conference on*, volume 2, pages 1199–1203. IEEE, 2002.
- [26] 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.
- [27] 3GPP. Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN). TR 25.913, 3rd Generation Partnership Project (3GPP), 12 2009.