Performance Study of Multimedia Services Using Virtual Token Mechanism for Resource Allocation in LTE Networks

Mauricio Iturralde*, Tara Ali Yahiya†, Anne Wei‡ and André-Luc Beylot*

*Université de Toulouse, IRIT/ENSEEIHT
† Université Paris Sud 11, LRI
‡CNAM, Laboratoire Cédric
{mauricio.iturralde, andre-luc.beylot} @enseeiht.fr
tara.ali-yahiya@u-psud.fr, anne.wei@cnam.fr

***Abstract*—The LTE specification provides QoS of multimedia services with fast connectivity, high mobility and security. However, 3GPP specifications have not defined scheduling algorithms to support real time and non-real time application services. In this paper we propose a modified version of M-LWDF and EXP/PF scheduling algorithms based on token mechanism, which provide better performance to real time flows such as video and VoIP in downlink system. By taking the arrival rates of packets to queues into account, the proposed mechanism included in previous scheduling algorithms, can increase remarkably the bit-rate for multimedia services. Simulation results show that the proposed modified algorithms can achieve a throughput gain for real time services specially for video traffic. Performance evaluation is conducted in terms of system throughput and packet loss ratio (PLR).**

***Index Terms*—Wireless networks, quality of service, performance evaluation, long-term evolution, virtual token.**

I. Introduction

In conjunction with the useful growth of Internet, multimedia and real time services, Long Term Evolution (LTE) technology has been proposed to perform this ambitious task. LTE uses orthogonal frequency division multiple access (ODFMA) in the downlink. OFDMA divides the frequency band into a group of mutually orthogonal subcarriers, thereby improving like this system capabilities by providing high data rates, supporting multi-user diversity and creating resistance to frequency selective fading of radio channels.

The quality of service (QoS) of LTE must be satisfied by giving users the optimal balance of utilization and fairness. Real Time (RT) services need a high QoS level and non-Real Time (NRT) services must have a minimum bit-rate. To satisfy this demand, several packet scheduling algorithms have been proposed to support RT and NRT traffic for mobile and wireless networks such as M-LWDF, PF, EXP/PF [8][4][14]. In the above mentioned schedulers, a priority value is assigned to each connection regardless its type depending on certain criterion. The connection with the best criteria is scheduled first at the next transmission time interval (TTI). This approach has the advantage of low implementation complexity. Due to the different traffic characteristics and QoS requirements of RT and NRT flows, it is difficult to define a single priority criterion. To do this, it is necessary to use different algorithms for each type of service (RT and NRT).

The aim of this paper is to investigate the performance of the algorithms EXP/PF and M-LWDF modified respectively to use a virtual token mechanism. In previous publications, virtual token mechanisms have been proposed to guarantee a minimum throughput to non-real time services [12] and [13]. Given that multimedia services are becoming the most important applications in telecommunications technology (e.g video and VoIP), this work has adapted a scenario to use the token queues concept to improve the performance for multimedia services.

This paper is organized as follows. Section II describes the downlink system model in LTE, section III describes the aforementioned packet scheduling algorithms and their characteristics. There is also a discussion about the virtual token queues mechanism describing its adaptation to M-LWDF and EXP/PF. In section IV, a simulation environment scenario is presented, where the traffic model is described and a numerical results analysis is exposed. Section V concludes this paper.

II. Downlink System Model

The QoS aspects of the LTE downlink are influenced by a large number of factors such as: Channel conditions, resource allocation policies, available resources, delay sensitive/insensitive traffic. In LTE the resource that is allocated to a user in the downlink system, contains frequency and time domains, and is called resource block. The architecture of the 3GPP LTE system consists of several base stations called "eNodeB" where the packet scheduling is performed along with other radio resource management (RRM) mechanisms.

The entire bandwidth is divided into 180 kHz, physical resource blocks (RB's), each one lasting 0.5 ms and consisting of 6 or 7 symbols in the time domain, and 12 consecutive subcarriers in the frequency domain. The resource allocation is realized in every TTI, that is exactly every two consecutive resource blocks. In this way, resource allocation is made on a resource block pair basis.

Users report their instantaneous downlink channel conditions (e.g signal-to-noise-ratio, SNR) to the eNodeB at each

978-1-4244-8327-3/11/$26.00 ©2011 IEEE

TTI. At the eNodeB the packet scheduler performs a user selection priority, based on criteria such as channel conditions, HOL packet delays, buffers status, service types etc. The eNodeB has a complete information about the channel quality by the use of Channel State Information (CSI).

Packets arriving into the buffer at eNodeB, are time stamped and queued for transmissions based on a first-in first-out (FIFO) scheme. For each packet in the queue at the eNodeB buffer, the head of line (HOL) is computed, and a packet delay is computed as well. If the HOL packet delay exceeds a specified threshold, packets are discarded.

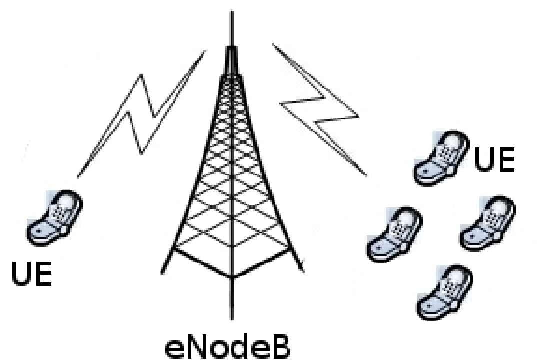

Figure 1. LTE downlink system

III. Scheduling Algorithms

Resource allocation among different flows in wireless networks is performed by several scheduling policies such as channel gain, average rate, and packets arrival delay. For elastic flows, PF is a good choice due to its capacity to allocate selecting users by their channel state. For real time services, algorithms based on packet delay such as M-LWDF have been proposed. For mixed flows, real time flows and non-real time flows, algorithms as EXP/PF or EXP rule have also been proposed. The scheduling algorithms under consideration in this work are specifically M-LWDF and EXP/PF. PF is also used as reference for this performance.

A. Proportional Fair (PF)

Proportional Fair algorithm [4], which is implemented in High Data Rate (HDR) networks was introduced to compromise between a fair data rate for each user and the total data rate. PF is a very suitable scheduling option for non-real time traffic. It assigns radio resources taking into account both the experienced channel quality and the past user throughput. The goal is to maximize the total network throughput and to guarantee fairness among flows.

$$j = \frac{\mu_i(t)}{\bar{\mu_i}} \tag{1}$$

where $\mu_i(t)$ denotes the data rate corresponding to the channel state of the user i at time slot t, $\bar{\mu_i}$ is the mean data rate supported by the channel.

B. Maximum-Largest Weighted Delay First (M-LWDF)

M-LWDF is an algorithm designed to support multiple real time data users in CDMA-HDR systems[8]. It supports multiple data users with different QoS requirements. This algorithm takes into account instantaneous channel variations and delays in the case of video service.

The M-LWDF scheduling rule tries to balance the weighted delays of packets and to utilize the knowledge about the channel state efficiently. At time slot t, it chooses user j for transmission as follows

$$j = \max_i a_i \frac{\mu_i(t)}{\bar{\mu_i}} W_i(t) \tag{2}$$

where $\mu_i(t)$ denotes the data rate corresponding to the channel state of the user i at time slot t, $\bar{\mu_i}$ is the mean data rate supported by the channel, $W_i(t)$ is the HOL packet delay and $a_i > 0, i = 1, ..., N$, are weights, which define the required level of QoS. According to [5], a rule for choosing a_i, which works in practice, is $a_i = -\log(\delta_i)T_i$. Here T_i is the largest delay that user i can tolerate and δ_i is the largest probability with which the delay requirement can be violated.

C. Exponential Proportional Fairness (EXP/PF)

Exponential Proportional fairness is an algorithm that was developed to support multimedia applications in an adaptive modulation and coding and time division multiplexing (ACM/TDM) system, this means that a user can belong to a real time service or non-real time service [14]. This algorithm has been designed to increase the priority of real time flows with respect to non-real time ones.

At time slot t, the EXP rule chooses user j for transmission as follows

$$j = \max_i a_i \frac{\mu_i(t)}{\bar{\mu_i}} \exp\left(\frac{a_i W_i(t) - \overline{aW}}{1 + \sqrt{\overline{aW}}}\right) \tag{3}$$

where all the corresponding parameters are the same as in the M-LWDF rule, except the term $\overline{aW}$ defined as

$$\overline{aW} = \frac{1}{N} \sum_i a_i W_i(t) \tag{4}$$

When the HOL packet delays for all the users do not differ a lot, the exponential term is close to 1 and the EXP rule performs as the Proportionally Fair rule. If for one of the users the HOL delay becomes very large, the exponential term overrides the channel state-related term, and the user gets a priority.

D. Token Queues Mechanism

It is essential that queue state information, such as queue length and packet delay, which is a reflection of traffic burstiness, to be utilized in scheduling packets. On the other hand, since the queue state information is tightly connected with QoS, wisely controlling queues is one of the most effective ways for QoS provisioning. As it is described earlier, M-LWDF and EXP/PF are making scheduling decisions based on the actual packet delays. However, we propose to modify these algorithms by combining them with virtual token in order not to take only the delay into consideration but also to provide a certain minimum throughput to flows. For doing so, a virtual token queue is associated to each flow, into which tokens arrive at constant rate r_i, the desired guaranteed minimum throughput of flow i. Let us define $V_i(t)$ to be the delay of the head of line token in the flow i token queue. Note that we do not need to maintain the token delays. As the arrival rates of tokens are constant,

$$V_i(t) = \frac{Q_i(t)}{r_i} \tag{5}$$

where $Q_i(t)$ is the token queue length (a counter value at time t). The value for r_i is 1 in our simulation scenario, like this, the same desired minimum throughput is given to all flows.

Then, we use the M-LWDF and EXP/PF rules with $W_i(t)$ being replaced by $V_i(t)$ respectively.

$$j = \max_i a_i \frac{\mu_i(t)}{\bar{\mu_i}} V_i(t) \tag{6}$$

$$j = \max_i a_i \frac{\mu_i(t)}{\bar{\mu_i}} \exp\left(\frac{a_i V_i(t) - \overline{aW}}{1 + \sqrt{\overline{aW}}}\right) \tag{7}$$

After the service of a real queue, the number of tokens in the correspondent token queue is reduced by the actual amount of data transmitted.

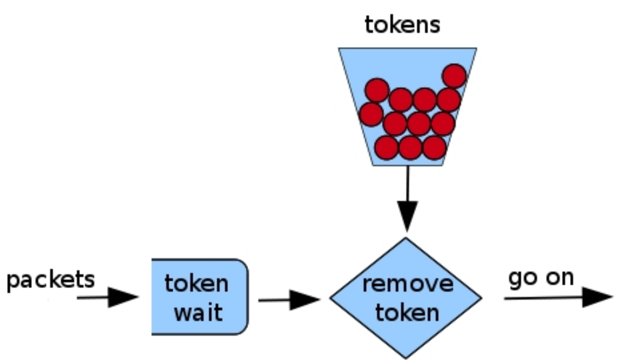

Figure 2. Token queues mechanism

IV. Simulation Environment

This paper investigates the performance of PF, M-LWDF, and EXP/PF in LTE. In this process a single cell with interference scenario is used. We have set up a scenario where there are 40% of users using video flows, 40% of users using VoIP flows and the remaining 20% are best effort and FTP flows. Users are constantly moving at speed of 3 kmph in random directions (random walk). LTE-Sim simulator is used to perform this process[11]. LTE-Sim provides a support for radio resource allocation in a time-frequency domain. According to [11], in the time domain, radio resources are distributed every Transmission Time Interval (TTI), each one lasting 1 ms. Furthermore each TTI is composed by two time slot of 0.5 ms, corresponding to 14 OFDM symbols in the default configuration with short cyclic prefix; 10 consecutive TTIs form the LTE Frame. Simulation parameters are shown in Table I.

Table I
LTE DOWNLINK SIMULATION PARAMETERS

Parameters	Values
Simulation duration	150 s
Flows duration	120 s
Frame structure	FDD
Radius	1 km
Bandwidth	10 MHz
Slot duration	0.5 ms
Scheduling time (TTI)	1 ms
Number of RBs	50
Max delay	0.1 s
Video bit-rate	242 $kbps$
VoIp bit-rate	8.4 $kbps$

A. Traffic Model

A video service with 242 kbps source video data rate is used in the simulation, this traffic is a trace based application that sends packets based on realistic video trace files which are available on [15]. For VoIP flows G.729 voice flow are generated by the VoIP application. In particular, the voice flow has been modeled with an ON/OFF Markov chain, where the ON period is exponentially distributed with mean value 3 s, and the OFF period has a truncated exponential probability density function with an upper limit of 6.9 s and an average value of 3 s [9]. During the ON period, the source sends 20 bytes sized packets every 20 ms (i.e., the source data rate is 8.4 kbps), while during the OFF period the rate is zero because the presence of a Voice Activity Detector is assumed. Best effort flows are created by an infinite buffer application which models an ideal greedy source that always has packets to send. The FTP application is modeled using a lognormal function with an upper limit of 5 Mbytes and a mean value of 2 Mbytes for the file size, reading time is modeled using an exponential distribution with a mean value of 180 s [3].

The LTE propagation loss model is composed by 4 different models (shadowing, multipath, penetration loss and path loss)[1]

- Pathloss: $PL = 128 : 1 + 37 : 6 \log(d)$ where d is the distance between the UE and the eNB in km.
- Multipath: Jakes model
- PenetrationLoss: 10 dB
- Shadowing: log-normal distribution (mean = 0dB, standard deviation = 8dB)

B. Numerical Results

In [13] [12] virtual tokens are used to guarantee a minimum throughput for non-real time users by modifying the EXP rule. In our case, the scenario aims to perform a virtual token method for real time flows. EXP/PF and M-LWDF schedulers have been chosen and modified to use a virtual token mechanism to improve their performance when using multimedia services such as video and VoIP. To compare results the following notation is used: "PF" represents the

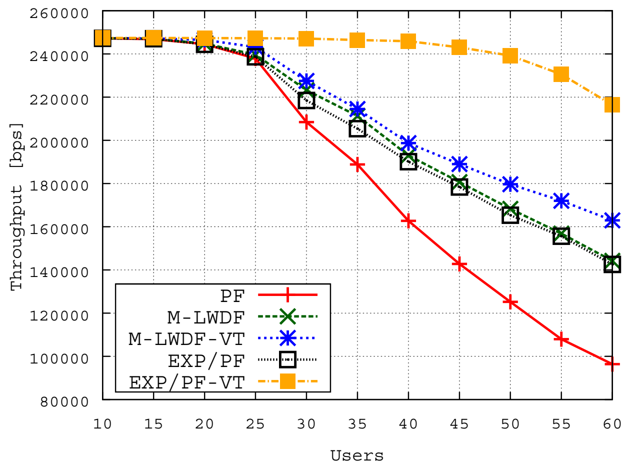

Figure 3. Average throughput per video flow

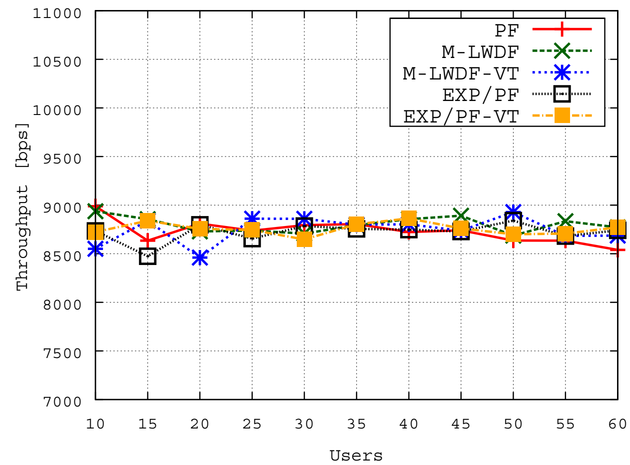

Figure 5. Average throughput per VoIP flow

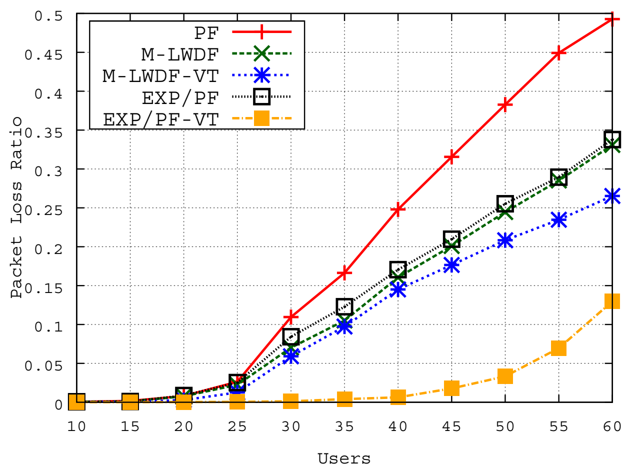

Figure 4. Packet loss ratio for video flows

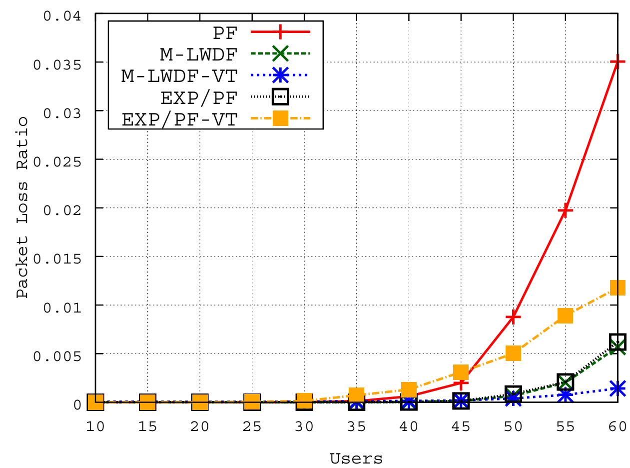

Figure 6. Packet loss ratio for VoIP flows

proportional fair algorithm (Equation 1), “M-LWDF” represents the classic M-LWDF algorithm (Equation 2), “EXP/PF” represents the classic EXP/PF (Equation 3), “M-LWDF-VT” represents the modified M-LWDF which use the virtual token mechanism (Equation 6), and "EXP/PF-VT" represents the modified EXP/PF which use the virtual token mechanism (Equation 7).

In this analysis, a percentage value is used to compare modified algorithms results with non-modified algorithm results.

For video flows, the aggregate throughput increases considerably by 35.1% for 60 users when EXP/PF-VT is used, compared to the non-modified EXP/PF (Figure 3). Packet loss ratio (Figure 4) decreases considerably by 61.6% when using EXP/PF-VT. Therefore, this modification shows a good performance for video flows. M-LWDF-VT shows a better throughput performance as well, increasing by 11%. M-LWDF-VT show a decrease of PLR by 20%.

For VoIP flows there is not a significant variation of throughput between modified algorithms (Figure 5). Packet loss ratio decreases when using M-LWDF-VT. It also presents a stable performance when the cell is charged by 60 users. Packet loss ratio increases when using EXP/PF-VT when the cell is charged by more than 56 users (Figure 6).

For non-real time flows, throughput decreases when the cell is charged by more than 30 users (Figure 7). EXP/PF-VT shows the highest and considerable difference from 30 to 60 users with a throughput decrease of 28.9%. M-LWDF-VT shows a throughput decreases up to 15.8% compared to the classic M-LWDF. Packet loss ratio increases for both algorithms (Figure 8), EXP/PF-VT being the one that presents the highest PLR increase, shows a rise up to 86.1% compared to EXP-PF when the cell is charged by more than 30 users.

It should be noted that there are no QoS guarantees for best effort flows. Therefore virtual token modification does not produce good results for best effort and FTP services because we set to 1 the value of r in Equation 5. Since we use an ON/OFF model, no packet arrives at the virtual queue when the state is set to OFF. This explains why there is almost no variation in throughput gain. Video flows will always have the largest virtual token queue, which justify the considerable throughput gain as shown in Figure 3. Helped by a virtual queue, the real queue is well served, which explains the low PLR obtained in simulations as shown in Figure 4.

All the efficiency obtained by video flows penalizes to non-

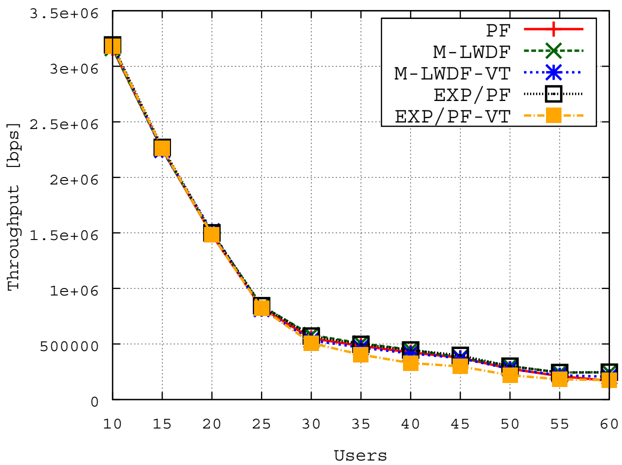

Figure 7. Average throughput per FTP and best effort flows

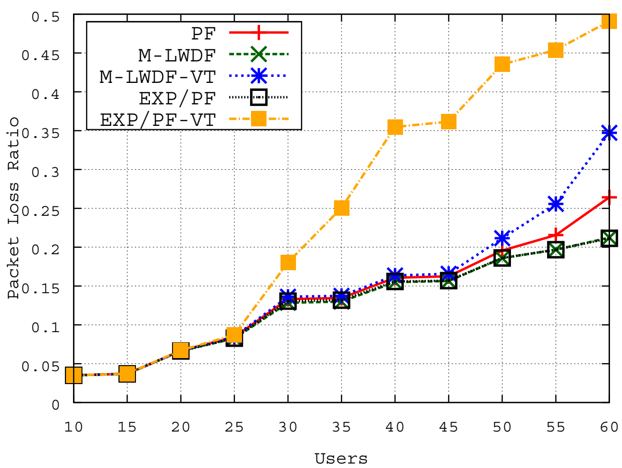

Figure 8. Packet loss ratio for FTP and best effort flows

real time flows. This happens due to the FTP model and the best effort flows (infinite buffer application) used. However, it should also be noted that LTE-Sim works under UDP traffic, so considering that FTP is normally implemented under TCP, the PLR rate could be lower than shown in Figure 8 due to the TCP retransmission control.

V. Conclusions

This paper aims to carry out the performance of resource allocation in downlink LTE system to enhance the QoS for real time services. In this paper a virtual token mechanism has been implemented along with some well known algorithms to improve the performance of multimedia services. We defined two performance metrics, namely, throughput and PLR. With respect to these measures, we found that EXP/PF-VT performs the best among the algorithms considered for video flows, nevertheless it penalizes non-real time flows. Moreover, M-LWDF-VT experiments a small performance increase for video flows, but the non-real time flows decrease, is not as high as when using EXP/PF-VT. However, this experimented inequity between RT and NRT flows, could be controlled by modifying the desired throughput value, in the virtual token mechanism. Our results show that token queue mechanisms used in conjunction with an intelligent scheduling algorithm can improve the efficiency of this algorithm for multimedia services. Future research could be focused on using virtual token mechanisms to balance the bit-rate distribution between RT and NRT flows.

References

[1] Tech. Specif. Group Radio Access Network 3GPP. "Physical layer aspect for evolved universal terrestrial radio access (utra) (release 7)". Technical report, 3GPP TS 25.814.

[2] Tech. Specif. Group Radio Access Network 3GPP. "Medium Access Control (MAC) protocol specification (release 9)". Technical report, 3GPP TS 36.321.

[3] Tech. Specif. Group Radio Access Network 3GPP. "Feasibility Study for Orthogonal Frequency Division Multiplexing (OFDM) for UTRAN enhancement (Release 6)". Technical report, 3GPP TS 25.892.

[4] J.-G. Choi and S. Bahk, "Cell-throughput analysis of the proportional fair scheduler in the single-cell environment". *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 766–778, Mar. 2007.

[5] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, and P. Whiting. "Providing quality of service over a shared wireless link". *IEEE Communications Mag.*, vol. 39, no. 2, pp. 150-154, Feb. 2001.

[6] H. Ekstrom. "QoS control in the 3GPP Envolved Packet System". *IEEE Communications Mag.*, vol. 47, no. 2, pp. 76-83, Feb. 2009.

[7] K. Pedersen, T. Kolding, F. Frederiksen, I. Kovacs, D. Laselva and P. Mogensen. "An overview of downlink radio resource management for UTRAN long-term evolution". *IEEE Communications Mag.*, vol. 47, no. 7, pp. 86-93, Jul. 2009.

[8] P. Ameigeiras, J. Wigard, and P. Mogensen. "Performance of the m-lwdf scheduling algorithm for streaming services in hsdpa". *IEEE Trans. Veh. Technol. Conf.*, vol. 2, pp. 999-1003, Sep. 2004. Los Angeles, USA.

[9] C. Chuah and R. H. Katz. "Characterizing packet audio streams from internet multimedia applications". *IEEE Int. Commun. Conf. (ICC)*, vol. 2, pp. 1199-1203, Apr. 2002. New York, USA.

[10] K. Kim, I. Koo, S. Sung, and K. Kim. "Multiple qos support using m-lwdf in ofdma adaptive resource allocation". *IEEE Loc. Metr. Net. Worksh.(LANMAN).*, pp. 217- 221, Oct. 2004. San Francisco, USA.

[11] G. Piro, L. Grieco, G. Boggia, F. Capozzi, and P. Camarda. "Simulating lte cellular systems: an open source framework". *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 498-513, Oct 2010.

[12] S. Shakkottai and A. Stolyar. "Scheduling Algorithms for a mixture of real time and non real time data in HDR". *Bells Laboratories*, 2004.

[13] K. Chang and Y. Han. "QoS-Based Adaptive Scheduling for a mixed service in HDR Systems". *IEEE Int. Symp. (PIMRC).*, vol. 4, pp. 1914-1918, September 2002. Lisboa, Portugal.

[14] R. Basukala, H. Mohd Ramli and K. Sandrasegaran. "Performance Analysis of EXP/PF and M-LWDF in Downlink 3GPP LTE System". *IEEE F. Asian Himalayas Conf.*, pp. 1-5, Nov. 2009. Kathmandu, Nepal.

[15] Video trace library. http://trace.eas.asu.edu/.

[16] B. Sadiq, R. Madan and A. Sampath. "Downlink Scheduling for Multiclass Traffic in LTE". *EURASIP Journal on Wirel. Comm. and Netw.*, vol. 2009 pp. 1-18, July. 2009.