

# LTE QoS DYNAMIC RESOURCE BLOCK ALLOCATION WITH POWER SOURCE LIMITATION AND QUEUE STABILITY CONSTRAINTS

Mohamed A. Abd El-Gawad<sup>1</sup>, Mohsen M. Tantawy<sup>2</sup> and Mohamed El-Mahallawy<sup>3</sup>

<sup>1,2</sup>Network Planning Department, National Telecommunication Institute, Cairo, Egypt  
m.algawad@gmail.com, ntimohsen@gmail.com

<sup>3</sup>Communication & Electronics Department, Arab Academy for Science, Technology &  
Maritime Transport, Cairo, Egypt  
mahallawy@ieee.org

## **ABSTRACT**

3GPP has defined the long term evolution (LTE) for 3G radio access in order to maintain the future competitiveness for 3G technology, the system provides the capability of supporting a mixture of services with different quality of service (QoS) requirements. This paper proposes a new cross-layer scheduling algorithm to satisfy better QoS parameters for real time applications. The proposed algorithm takes care of allocating resource blocks (RBs) with different modulation and coding schemes (MCS) according to target bit error rate (BER), user equipment supportable MCS, queue stability constraints and available transmit power constraints. The proposed algorithm has been valued, compared with an earlier allocation algorithm in terms of service rate and packet delay and showed better performance regards the real time applications.

## **KEYWORDS**

*Cross-layer Scheduling, LTE, OFDM, QoS, Packet Delay*

## **1. INTRODUCTION**

To face the ever growing demand for packet-based mobile broadband systems, the 3GPP has introduced the LTE (Long Term Evolution) specifications as the next step of the current 3G mobile networks. An enhanced access network (i.e., the E-UTRAN, Evolved-UMTS (Universal Mobile Telecommunications System) Terrestrial Radio Access Network) and an evolved core network have been defined [1-3].

The radio transmissions in LTE are based on the Orthogonal Frequency Division Multiplexing (OFDM) modulation scheme. In particular, the Single Carrier Frequency Division Multiple Access (SC-FDMA) and the OFDM Access (OFDMA) are used in uplink and downlink transmissions, respectively. Differently from basic OFDM, they allow multiple access by assigning sets of sub-carriers to each individual users. Moreover, OFDMA can exploit subsets of sub-carriers distributed inside the entire spectrum whereas SC-FDMA can use only adjacent sub-carriers. OFDMA is able to provide high scalability, simple equalization, and high robustness against the time-frequency selective fading of the radio channel. On the other hand, SC-FDMA is used in the LTE uplink to increase the power efficiency of user equipments (UEs) which are battery supplied [4].

LTE has a frame duration of  $T_f = 10$  ms and it is divided into equally size sub-frame, called Transmission Time Interval (TTI), lasting 1 ms. The whole bandwidth is divided into 180 kHz

physical RBs, each one lasting 0.5 ms and consisting of 6 or 7 symbols in the time domain (according to the OFDM prefix-code duration) and 12 consecutive sub-carriers in the frequency domain as shown in figure 1[5].

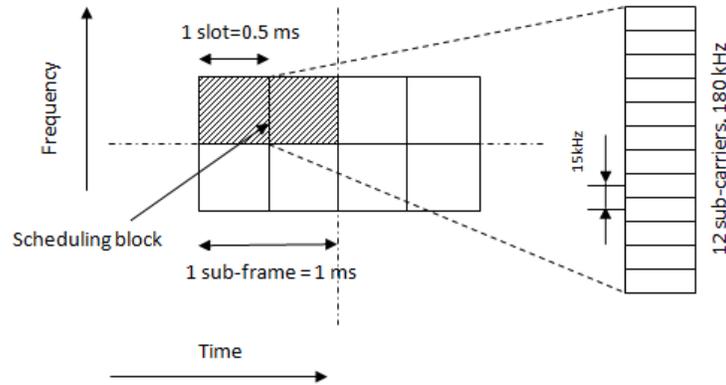


Figure 1. Illustration of a scheduling block in LTE downlink

The resources allocation is realized every TTI, that is exactly every two consecutive RBs; thus, resource allocation is done on a RB pair basis so during the remaining of this paper we use the term RB to denote two consecutive RBs in time domain that constitute one TTI.

Every TTI, the Channel Quality Indicator (CQI) is reported by the user measurement entity to the base station (BS) to provide time and frequency channel quality information for better spectral efficiency and resource allocation. For downlink RBs, users use the Physical Uplink Control Channel (PUCCH) to convey channel quality information to the BS. BS conveys downlink RBs allocations and MCS assignments to all users using the Physical Downlink Control Channel (PDCCH) [5]. For uplink RBs, the BS estimates the channel quality of the received uplink RBs and conveys uplink RB allocations to users using the PDCCH.

The information about the quality of the time and frequency variant channel is exploited by the link adaptation module to select, for each UE, the most suited MCS at the physical level, in order to maximize the spectral efficiency. This approach is known as Adaptive Modulation and Coding (AMC) and it has been adopted by several wireless technologies, such as EDGE [6] and WiMAX [7]. Considering that each modulation scheme (i.e., Quadrature Phase-Shift Keying (QPSK), 16-Quadrature Amplitude Modulation (QAM), and 64-QAM in LTE) corresponds to a fixed physical data rate, the link adaptation module can establish the maximum available physical data rate for each UE based on the received channel quality information, providing optimal resource allocation among all users.

In general, the most important objective of LTE scheduling is to satisfy the QoS requirements of all users by trying to reach, at the same time, an optimal tradeoff between utilization and fairness [8]. This goal is very challenging, especially in the presence of real time multimedia applications, which are characterized by strict constraints on packet delay and jitter.

Recent researches in the area of providing QoS aware resource allocation in multiuser OFDM systems can be subdivided in two broad categories: One, is to provide system-level QoS provisioning (i.e., to maximize the sum capacity of the system). Second, is to provide user-level QoS provisioning (i.e., to satisfy the QoS (data rate and BER) constraints of each user of the system). Although several works are done in the area of system-level QoS provisioning [9-11], less interest is shown in the area of user-level QoS provisioning.

User-level QoS allocation problem could be divided into allocating RBs and power to individual users under various constraints. Since, the joint optimization problem of allocating RBs and power optimally to satisfy user-level QoS under the total available power constraint, involves discrete assignment and is thus not convex in the unknowns of RBs allocation and the available power, making it an Non-deterministic Polynomial (NP)-hard problem [12].

As this problem is NP-hard and exhaustive search over all possible allocations is impractical for large scale systems, we are motivated to carry out RB allocation and power allocation algorithms separately and sequentially to provide service to heterogeneous users in LTE system.

In this paper, user-level QoS allocation algorithm has been proposed where the resources (RBs and MCS) are allocated separately and sequentially according to each user traffic profile which represents its channel condition, traffic load and QoS requirements so as to minimize the overall average packet transfer delay, while satisfying target BER, user equipment supportable MCS, queue stability constraints and available transmit power constraints.

The rest of this paper is organised as follow: Section 2 presents the system model. Section 3 describes the proposed QoS-based algorithm. Then, simulation results are given in section 4, and conclusion is made in section 5.

## 2. SYSTEM MODEL

Figure 2 shows the system model for LTE transmitter and receiver with cross-layer design to allocate RBs in a frame to different users. Frequency Division Duplex (FDD) is assumed to be the operation mode.

Each OFDM symbol contains a number of subcarriers in the set  $S_c = \{1, 2, \dots, C\}$  where  $C$  is the total number of subcarriers in an OFDM symbol. All RBs in the set  $S_r = \{1, 2, \dots, R\}$  are available for allocation to users where  $R$  is the total number of RBs in a frame.  $S_r^t \in S_r$  is the set representing all RBs in the  $t^{th}$  TTI whereas  $S_r^n \in S_r$  is the set representing all RBs in a frame allocated to a user  $n$ .

Therefore  $S_r^t \cap S_r^n$  is the set of RBs assigned to user  $n$  in the  $t^{th}$  TTI.  $S_t = \{1, 2, \dots, T\}$  is set of all TTIs where  $T$  is the total number of TTIs in a frame. Each RB is a grid with size equal to  $K < |S_c|$  consecutive subcarriers by  $F$  consecutive OFDM symbols.

Following the LTE standard constraints in [5], [13], all RBs have the same grid size, subcarriers in each RB have the same MCS, and all RBs allocated to a user in one TTI have also the same MCS. Each user  $n$  supports a maximum number  $M$ , of MCSs in the set  $S_\phi^n = \{\Phi_{n,1}, \Phi_{n,2}, \dots, \Phi_{n,M}\}$  where  $\Phi_{n,M}$  represents the number of data bits per subcarrier (i.e., 1, 2, 4, 6 for BPSK, QPSK, 16-QAM, and 64-QAM respectively).

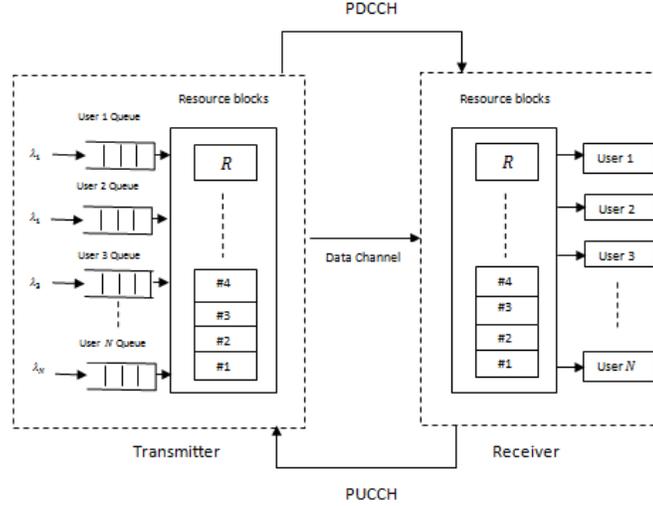


Figure 2. LTE model for cross-layer design and RB allocation

At the transmitter, there are  $N$  users each of them is represented as M/D/1 queue. There are  $Z$  traffic types and their priority set  $S_p = \{p_1, p_2, \dots, p_Z\}$ . Data packets arrive at queue  $n$  according to Poisson process with arrival rate  $\lambda_n$  bits per frame. Queue  $n$  packets are serviced with a service rate  $V_n$  bits per frame.

The average packet transfer delay of a user is given by [14]:

$$D_n = \frac{1}{V_n} \left[ 1 + \frac{\rho_n}{2(1-\rho_n)} \right], \forall n \in U, \quad (1)$$

Where  $\rho_n = \lambda_n / V_n < 1$ ,  $\lambda_n \neq 0$  and  $U = \{1, \dots, N\}$ .

An RB allocation matrix is denoted by  $a_{n,r}$  (i.e.  $a_{n,r}$  is a  $N$  rows by  $R$  columns matrix where  $a_{n,r} = 1$  if RB  $r$  is allocated to user  $n$  and  $a_{n,r} = 0$  if RB  $r$  is not allocated to user  $n$ ). The total service rate for user  $n$  can be calculated as:

$$V_n = K \cdot F \cdot \sum_{r=1}^R (a_{n,r} \cdot \Phi_{n,b_{n,r}}), \forall r \in S_r^n, \quad (2)$$

Where  $b_{n,r}$  is an MCS assignment matrix ( $N$  by  $R$ ) representing the MCS index assigned for each RB (i.e.,  $b_{n,r} = 2$  if RB  $r$  allocated to user  $n$  is assigned MCS  $\Phi_{n,2}$ ).

The transmit power of the  $r^{th}$  RB allocated to user  $n$  is denoted by  $P_{n,r}$ . That is,  $P_{n,r}$  represents the sum of transmit power allocated to all subcarriers of an OFDM symbol belonging to  $r^{th}$  RB allocated to user  $n$ .  $P_{n,r}$  can be expressed as [15]:

$$P_{n,r} = K \cdot \frac{\sigma^2 \Gamma}{h_{n,r}} (2^{\Phi_{n,b_{n,r}}} - 1), \quad (3)$$

Where  $\Gamma$  is a constant Signal-to-Noise Ratio (SNR) which is related with the target BER and  $\sigma^2$  denotes the variance of additive white Gaussian noise. The total transmit power of an OFDM symbol in the  $t^{th}$  TTI can be expressed as:

$$P_t = \sum_{r=1}^R \sum_{n=1}^N (a_{n,r} \cdot P_{n,r}), \forall r \in S_r^t, \quad (4)$$

All OFDM symbols in a TTI have the same transmit power due to the fact that each RB has the same MCS and thus the same number of allocated bits to each of its subcarrier. Each TTI can have different transmit power.

The objective of the allocation algorithm, which allocates RBs and assigns MCS to each RB, is to minimize the overall average packet delay for all users specially who has real time traffic while providing the BER requirement, queue stability constraint, power resource limitation, channel condition awareness, supportable MCS, and transport block constrains. This is an optimization problem that can be formulated as follow [15]:

$$(\mathbf{P}) \min_{\{V_n\}_{n=1}^N} \bar{D} = \frac{1}{\lambda_1 + \lambda_2 + \dots + \lambda_N} \sum_{n=1}^N \lambda_n \bar{D}_n$$

Subject to the following constrains:

$$\begin{aligned} C1: & \sum_{n=1}^N a_{n,r} \leq 1, \forall r \in S_r, \\ C2: & P_t \leq P_v, \forall t \in S_t, \\ C3: & \Phi_{n,b_{n,r}} \in S_{\Phi}^n, \forall n \in U, \forall r \in S_r^n, \\ C4: & \Phi_{n,b_{n,x}} = \Phi_{n,b_{n,y}} = \dots = \Phi_{n,b_{n,z}}, \forall n \in U, \forall x, y, z \in S_r^t \cap S_r^n, \forall t \in S_t, \\ C5: & V_n > \lambda_n, \forall n \in U \end{aligned} \quad (5)$$

where  $C1$  guarantees that each RB can only be used by at most one user;  $C2$  indicates that transmit power on all subcarriers of an OFDM symbol does not exceed the available transmit power  $P_v$ ;  $C3$  guarantees that an RB allocated to a user has its MCS supportable by that user;  $C4$  ensures that all RBs assigned to a user in a TTI (e.g. a transport block) have the same MCS; and finally  $C5$  guarantees that the service rate of a user is larger than its arrival rate. The problem  $\mathbf{P}$  is a complex combinatorial optimization problem that minimizes the weighted delay for all users. It is hard to find an optimal solution to this problem using known optimization techniques. For such a purpose, we propose a suboptimal solution to this problem as outlined in the next section.

### 3. THE PROPOSED QOS BASED ALGORITHM

The proposed algorithm has two phases, the first phase is the RB allocation phase while the second one is the MCS assignment phase.

In this algorithm the sub-carrier controller combines the CQI and the QoS information which is transferred from the traffic controller in the MAC layer to distribute the resources among the users with taking into account all constrains in (5).

The available resource blocks are allocated to users through an iterative process, where the total number of iterations is equal to the total number of RBs available at each frame. At each iteration only one RB is allocated to the user which maximizes the following proposed priority function.

$$PRF(n) = \frac{T_w(n) \cdot PLR(n) \cdot W_n}{T_w^{max} \cdot PLR_z^{thr} \cdot p_z}, \quad (6)$$

$$W_n = \frac{\lambda_n}{argmin_i \lambda_i} \forall n, \forall i \in U, \quad (7)$$

Step (1.b) sets each user  $PRF$ . The algorithm tries to spread RBs allocation to each user among TTIs as much as possible and at the same time select the best channel for a user. The reason for

the first criteria is that allocating different RBs in the same TTI for the same user does not utilize where  $T_w(n)$  is the packet waiting time in a queue  $n$  from the arriving instant to the scheduling instant,  $PLR(n)$  is the packet loss rate of user  $n$  which is the ratio between the transmitted and the dropped packets over the moving scheduling window ( $M_w$ ),  $W_n$  is the ratio of a user arrival rate to the minimum arrival rate,  $T_w^{max}_z$  is the maximum tolerated delay for the traffic type  $z$ ,  $PLR_z^{thr}$  is the maximum tolerated PLR for the traffic type  $z$  and  $p_z$  is the traffic priority.

The RB allocation phase starts by initializing the RBs allocation matrix,  $a_{n,r}$  the MCS assignment matrix,  $b_{n,r}$  the set of RBs allocated to user  $n$ ,  $S_r^n$ , a TTI counter,  $t$ , the set of users,  $U$ , and finally the priority function of each user,  $PRF$ .

the power resources efficiently since they must all have the same MCS (as required by C4). The second criteria stems from the fact that choosing an RB with the best channel condition is important for minimizing the transmit power.

In the second stage, the algorithm allocates one RB to each user and guarantees that each TTI has only one RB allocated to a single user. In the third stage, RB allocations are done in proportional to users'  $PRFs$ . First, a user is chosen according to its  $PRF$ .

Then, a TTI that has the minimum number of RBs allocated to that user is chosen (Step (3.b)). An RB from the selected TTI with the best channel condition to that user is allocated to it (Step (3.c)). The algorithm concludes the allocation phase by updating the allocation matrix  $a_{n,r}$  and the set of RBs allocated to each user  $S_r^n$ .

**Input:**  $\lambda_n, N, S_r, S_t, S_p, Z, h_{n,r}$

**Output:**  $a_{n,r}$

(1) Initialization

(a) Set  $a_{n,r} = 0, b_{n,r} = 0, S'_r = S_r, S_r^n = \phi, t = 1,$   
 $U = \{1, 2, \dots, N\};$

(b) Calculate  $PRF(n), \forall n, \forall i \in U;$

(2) For each  $n \in U,$

(a) Find  $r^* = \arg \max h_{n,r} \forall r \in S_r^t \cap S'_r;$

(b) Set  $a_{n,r^*} = 1, S'_r = S'_r - r^*, S_r^{n^*} = S_r^n + r^*;$

(c) add  $n$  to user subset  $U_i = \{U_1, U_2, \dots, U_Z\}$

(d) If  $t > |S_t|,$  Set  $t = 1,$  else Set  $t = t + 1;$

(3) While  $S'_r \neq \emptyset,$

(a) Find  $n^* = \arg \min_n \frac{|S_r^n|}{PRF(n)};$

(b) Find  $t^* = \arg \min_t |S_r^t \cap S_r^{n^*}|, \forall t \in S_t, S_r^t \neq \emptyset;$

(c) Find  $r^* = \arg \max_r h_{n^*,r}, \forall r \in S_r^{t^*} \cap S'_r;$

(d) Set  $a_{n^*,r^*} = 1, S'_r = S'_r - r^*, S_r^{n^*} = S_r^{n^*} + r^*$

Figure 3. Pseudo code for the RB allocation phase

The MCS assignment phase starts with initializing the TTI counter,  $t$ , as shown in the fourth stage. In the fifth stage, the algorithm gives priority to the users' subsets according to their traffic type so it begins with the subset which represents the most delay sensitive traffic and ends with the subset which represents the least delay sensitive traffic. Firstly the algorithm distributes  $V_n = \lambda_n + 1$  bits for each user among its allocated RBs in order to fulfill C5. For each user, the algorithm continuously selects the RB with the lowest MCS (Step (a.4.1)), finds the

corresponding TTI (Step (a.4.2)), and increments MCS for all RBs in this TTI belonging to the same user (Step (a.4.4)) until the user total service rate exceeds its arrival rate but in the same time checks the transmitted power level (Step (a.4.5), (a.4.6), (a.4.7)) in order to fulfill  $C2$  specially at the larger values of user arrival rates.

The six stage aims at assigning MCS for all RBs in a way to minimize the overall average packet delay. As in the fifth stage the algorithm gives priority to the users subsets according to their traffic type, this stage keep toggling between TTIs and in each TTI, it increases the MCS for those RBs in that TTI which yields the lowest average delay. In each TTI, our algorithm starts by calculating the total service rate for each user as currently assigned to its RBs (Step (a.1.1)) and the total service rate when the MCS for all its RBs in this TTI are incremented (Step (a.1.2) and (a.1.3)). Note that if an RB allocated to a user is already assigned the maximum supportable MCS, Step (a.1.2) has no effect and the total service rate as in Step (a.1.3) remains the same as that in Step (a.1.1).

As a result of incrementing the MCS, the algorithm calculates the corresponding reduction in weighted delay (Step (a.1.5)). If the transmit power, as a result of incrementing the MCS, exceeds the available transmit power or the RBs are already assigned the maximum supportable MCS, the user is excluded for any further processing during this TTI (Step (a.1.6)).

If all users in a given TTI are excluded, then all RBs in this TTI are considered to have reached the maximum possible MCS and the TTI is excluded from any further processing (Step (b)). Step (c) selects a user with the maximum weighted delay reduction,  $\Delta\bar{D}_{max}$  and its corresponding power increase  $\Delta P_{min}$  Step (e), a Cost function is calculated for all users as follow [15]:

$$C(\Delta\bar{D}_n, \Delta P_n) = \frac{\Delta P_n}{\Delta P_{min}} - \frac{\Delta D_n}{\Delta\bar{D}_{max}} \forall n \in U, \quad (8)$$

where  $\Delta\bar{D}_n$  is user  $n$  reduction in weighted delay and  $\Delta P_n$  is the corresponding increase in power. The Cost function  $C(\cdot)$  is a function that determines the overall cost between a reduction in weighted delay and the corresponding increase in power.

**Input:**  $\lambda_n, N, S_r, S_t, S_\Phi^n, U_i, h_{n,r}, \Gamma, \sigma^2, P_v, a_{n,r}$

**Output:**  $b_{n,r}$

- (4) Initialization: Set  $t = 1$ ;
- (5) For each  $U_i \forall i = \{1, 2, \dots, Z\}$ ,
  - (a) For each  $n \in U_i$ 
    - (a.1) Set  $b_{n,r} = 1, \forall r \in S_r^n$ ;
    - (a.2) Set  $V_n = \lambda_n + 1$ ;
    - (a.3) Set  $V_n = V_n - \sum_{r=1}^R (a_{n,r} \Phi_{n,b_{n,r}})$ ;
    - (a.4) while  $V_n > 0$ ,
      - (a.4.1) Find  $r^* = \arg \min_r b_{n,r} \forall r \in S_r^n$ ;
      - (a.4.2) Find  $t^* = \arg_t r^*, \forall t \in S_t$ ;
      - (a.4.3) Calculate  $V_n'$  and  $P_t'$  as in Eq. (2&4);
      - (a.4.4) Set  $b'_{n,r} = b_{n,r} + 1, \forall r \in S_r^t \cap S_r^n$ ;
      - (a.4.5) Calculate  $V_n''$  and  $P_t''$  as in Eq. (2&4);
      - (a.4.6) If  $(P_t'' > P_v)$  or  $(V_n'' = V_n)$ , Break;
      - (a.4.7) Set  $V_n = V_n - (V_n'' - V_n')$ ;
- (6) while  $S_t = \emptyset$ ,
  - (a) For each  $U_i \forall i = \{1, 2, \dots, Z\}$ ,
    - (a.1) For each  $n \in U_i$ 
      - (a.1.1) Calculate  $V_n'$  and  $P_t'$  as in Eq. (2&4);

- (a.1.2) Set  $b'_{n,r} = b_{n,r} + 1, \forall r \in S_r^t \cap S_r^n$ ;
- (a.1.3) Calculate  $V_n''$  and  $P_t''$  as in Eq. (2&4);
- (a.1.4) Calculate  $\Delta P_n = P_t'' - P_t'$ ;
- (a.1.5) Calculate  $\Delta D_n$  corresponding to  $V_n'$  and  $V_n''$ ;
- (a.1.6) If  $(P_t'' > P_v)$  or  $(V_n'' = V_n')$ ,  $U_i' = U_i - n$ ;
- (b) If  $U_i' = \emptyset$ , set  $S_t = S_t - t$  and GOTO Step(h);
- (c) Find  $n^+ = \arg \max_n \Delta \bar{D}_n, \forall n \in U_i'$ ;
- (d) Set  $\Delta \bar{D}_{\max} = \Delta \bar{D}_{n^+}$  and  $\Delta P_{\min} = \Delta P_{n^+}$ ;
- (e) For each  $n \in U_i'$ ;
- (e.1) calculate  $C(\Delta \bar{D}_n, \Delta P_n)$  as in Eq. (8);
- (f) Find  $n^* = \operatorname{argmin}_n C(\cdot), \forall n \in U_i'$ ;
- (g) Set  $b_{n^*,r} = b_{n^*,r} + 1, \forall r \in S_r^t \cap S_r^n$ ;
- (h) If  $t > |S_t|$ , Set  $t = 1$ , else Set  $t = t + 1$ ;

Figure 4. Pseudo Code for the MCS assignment Phase

In other words, it aims at finding a reasonable cost that balances between reduction in weighted delay and power increase. After calculating  $C(\cdot)$  for all users, the algorithm selects the user  $n^*$  with the minimum cost function and increments the MCS for its RBs in the given TTI as shown in Step (f) and (g) respectively. Finally, Step (h), increments the TTI counter to move to another TTI This phase concludes when all RBs in each TTI either reach the maximum supportable MCS or their total transmit power reach the available transmit power.

#### 4. SIMULATION RESULTS

To investigate the performance of the proposed algorithm simulation has been performed to an LTE system with the parameters listed in table 1.

We assume a variable channel gain that is uniformly distributed within [0.2 1]. We assume that each user supports MCSs in the set  $S_\phi^n = \{1,2,4,6\}$  corresponds to BPSK, QPSK, 16-QAM, and 64-QAM respectively. Arrival and service rates are in bits per frame and packet sizes are assumed to be fixed and of size 500 bits.

Also we assume that the simulation environment has 20 users, and are divided into two equal groups; group A (with index  $\{1, 2, \dots, 10\}$ ) and group B (with index  $\{11, 12, \dots, 20\}$ ) where the arrival rates,  $\lambda_A$  of users in group A are equal to each other and allowed to vary whereas the arrival rates,  $\lambda_B$  of users in group B are always fixed and set to 150 bits per frame.

Users with the even index  $\{2, 4, \dots, 20\}$  serve TCP-based traffic (e.g. Web), and users with the odd index  $\{1, 3, \dots, 19\}$  serve conversational voice traffic.

Table 1. Simulation Parameters

Parameter	Value
System Bandwidth	1.4 MHz
Number of RBs	60
Number of TTIs per frame	10
Number of RBs per TTI	6
RB bandwidth	180 KHz
Channel type, Channel variance	AWGN, 1mW
Signal to noise ratio	8dB
Available transmit power ( $P_p$ )	10W
Time slot (transmission time interval)	1 ms
Number of users	20
Applications	Web, Voice
Packet delay budget	Web=500 ms, Voice=200 ms
Maximum tolerated packet loss rate	Web=1e-6, Voice= 1e-2
Traffic priority	Web= 6, Voice=2
Cyclic prefix	Normal

The simulation results are demonstrating the performance of the proposed QoS algorithm and its comparison to the previous Resource Block Allocation (RBA) algorithm [15].

Figure 5 demonstrates the overall average packet delay versus the arrival rates  $\lambda_A$  for both the QoS proposed algorithm and the previous RBA algorithm; it's found that the average packet delay of the QoS proposed algorithm is slightly larger than the delay of the RBA algorithm specially when  $\lambda_A$  goes to large values.

The reason of this delay increment is as follow:

In RBA algorithm, the resources are distributed among users based only on their arrival rates and there is no differentiation among them based on their QoS requirements.

In QoS algorithm, the allocation is based on users' arrival rates and their QoS requirements (application type, delay budget, loss rate). In other words users serve Voice application with large arrival rate exhaust more resources to achieve their demands under the system constraints. In contrast users serve Web application with large arrival rate expect less resources (large delay) as shown in figure 6.

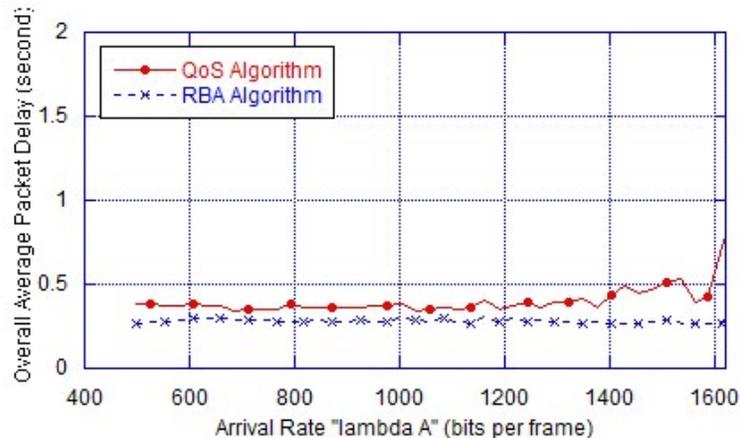


Figure 5. The average packet delay for the RBA and the proposed QoS Algo.

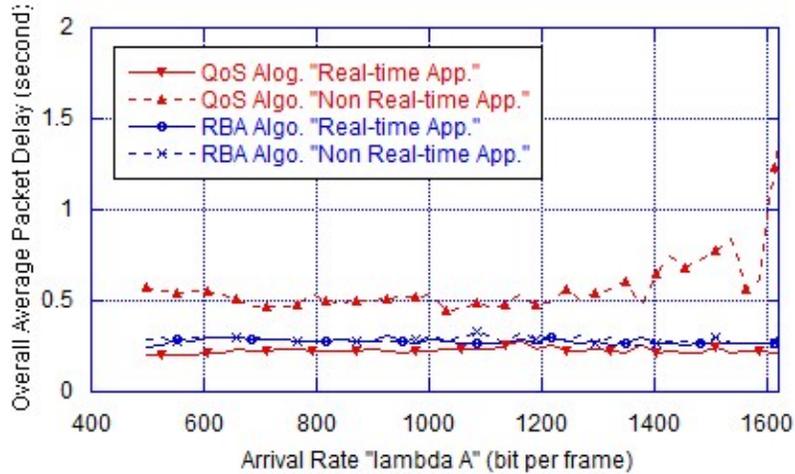


Figure 6. Differentiation between the real time application and the non real time application in term of the average packet delay.

In figure 7 throughput differentiation between the two applications is illustrated where in RBA algorithm there is no throughput differentiation. The throughput achieved by Voice users decreases while the arrival rate  $\lambda_A$  varies from  $\lambda_A=500$  to  $\lambda_A=1000$  and from  $\lambda_A=1000$  to  $\lambda_A=1600$  the throughput is approximately constant. In contrast throughput achieved by Web users increases while the arrival rate  $\lambda_A$  varies from  $\lambda_A=500$  to  $\lambda_A=1000$  and from  $\lambda_A=1000$  to  $\lambda_A=1600$  the throughput is approximately constant.

The behaviour illustrated in figure 7 is interpreted as follow: when  $\lambda_A$  goes to large values number of RBs which is assigned to Voice users has increased and due to our system limitation which discussed before in (5) (power limitation, all RBs in the same TTI must have the same MCS) the system is forced to downgrade the MCS of these RBs. On the opposite side, number of RBs which is assigned to Web users has decreased so the system tries to upgrade the MCS to achieve their demands.

In figure 8 when  $\lambda_A$  increases number of QPSK RBs is increased over the number of 16QAM RBs. Figure 9 shows that at  $\lambda_A = 1469$ , Voice users have large number of QPSK or 16QAM RBs but Web users have small number of 16QAM or 64QAM RBs.

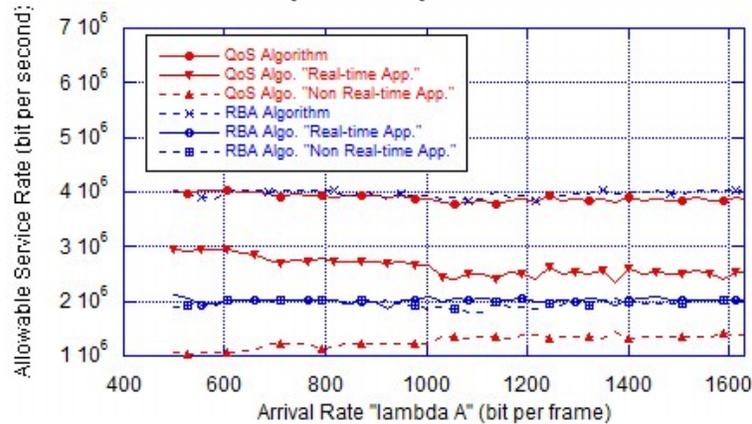


Figure 7. Service rate differentiation between traffic applications for the adaptive allocation strategies

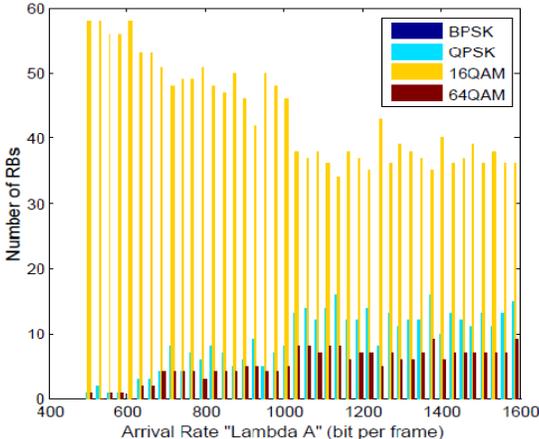


Figure 8. RBs distribution among the users variation with different arrival rates for the proposed QoS algorithm

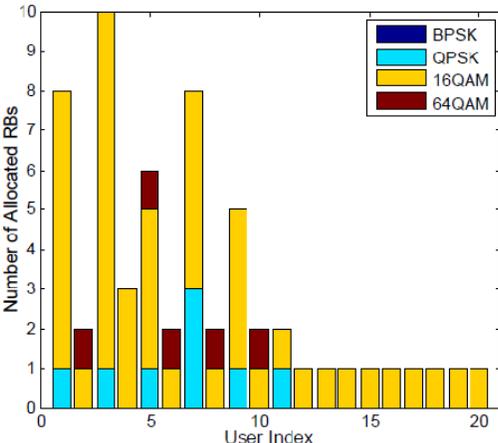


Figure 9. RBs distribution among the users for the proposed QoS Algorithm at “lambda A”=1469 (bit per frame)

Figures 10 ensure that the RBA algorithm doesn't differentiate between our applications and there is no biasing in RBs/MCS allocation toward any user or group of users.

Finally Fig. 11 shows that the maximum power value assigned in a frame doesn't exceed the threshold value which proves that  $C2$  is verified by our proposed algorithm.

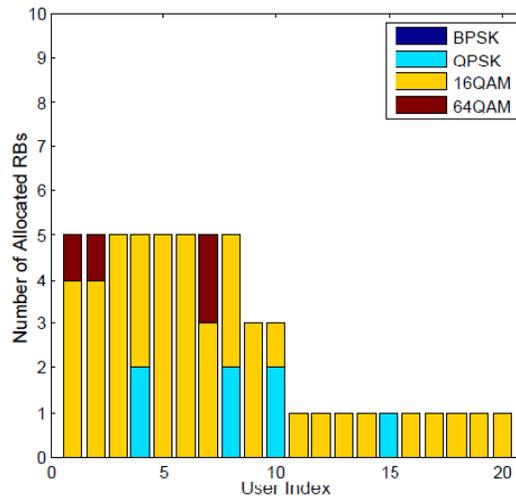


Figure 10. RBs distribution among the users for the legacy RBA Algorithm at “lambda A”=1469 (bit per frame)

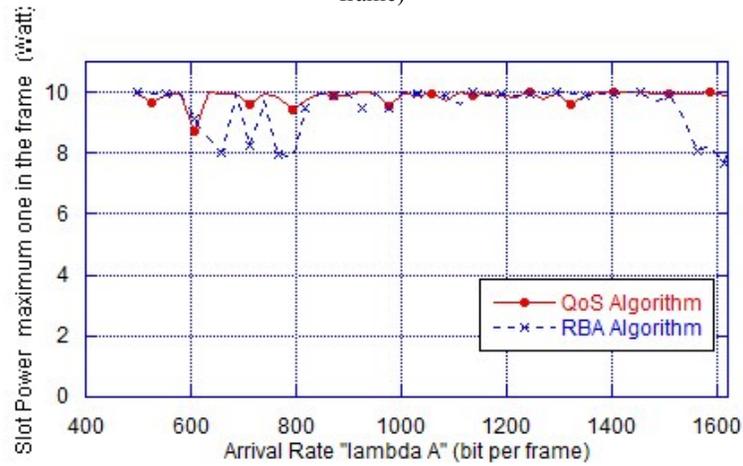


Figure 11. Slot power variation with different arrival rates for the different allocation strategies

## 5. CONCLUSION

In this paper, joint optimization problem of RBs and power resources allocation for an LTE system is considered. As such problems are complex in solving; most researches carry out subcarrier allocation and power allocation algorithms separately and sequentially so a suboptimal cross-layering algorithm for RB allocation and MCS assignment is proposed. Firstly the algorithm allocates RBs to different users in proportional to their CQIs and their traffic profiles, then, it performs MCS assignment for each RB in a way to minimize the overall average packet delay while taking into account queue dynamics, power limitation, channel condition and MCS capability of each user. Simulation results show that the proposed algorithm achieves more service rate and lower packet delay for the real time applications relative to a similar non-QoS previous algorithm and differentiates among users based on their QoS requirements. From the previous, our proposed algorithm shows that it has achieved user-level QoS provisioning.

## REFERENCES

- [1] 3GPP Technical Specification, (March 2008) "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description", Stage 2, TS 36.300, version 8.4.1.
- [2] 3GPP Technical Specification, (December 2006) "Service requirements for the Evolved Packet System (EPS)", TS 22.278, version 8.0.0.
- [3] 4gamericas, (2013) "Supporting Wireless Video Growth and Trends".
- [4] 3GPP Technical Specification, (June 2009) "Evolved universal terrestrial radio access (E-UTRA); physical layer procedures", TS 36.211, Version 8.7.0.
- [5] 3GPP Technical Specification, (March 2009) "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation", TS 36.211, V8.6.0.
- [6] T. Halonen, J. Romero, and J. Melero, (2003) GSM, GPRS and EDGE Performance: Evolution Towards 3G/UMTS, Wiley.
- [7] L. Nuaymi, (2008) WiMAX: Technology for Broadband Wireless Access, Wiley.
- [8] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, (2008) 3G Evolution HSPA and LTE for Mobile Broadband, Academic Press.
- [9] W. Rhee and J. M. Cioffi, (2000) "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation", IEEE VTC, Vol. 2, pp. 1085–1089.
- [10] H. Kim and Y. Han, (2005) "A proportional fair scheduling for multicarrier transmission systems", IEEE Commun. Letter, Vol. 9, No. 3, pp. 210–212.
- [11] M. Gawad, M. Tantawy and M. Mahallay, (2011) "New Cross-layer QoS-based Scheduling Algorithm in LTE System", In World Scientific and Engineering Academy and Society (WSEAS), pp. 190-194.
- [12] L.M.C. Hoo, J. Tellado, and J.M. Cioffi, (1998) "Dual QoS Loading Algorithm for DMT Systems Offering CBR and VBR Services", IEEE Global Telecommunications Conference, vol. 1, pp. 25-30.
- [13] 3GPP Technical Specification, (June 2011) "Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification", TS 36.321, Version 10.2.0.
- [14] L. Kleinrock, (1975) Queueing Systems, John Wiley & Sons, Inc.
- [15] H. Fattah, and H. Alnuweiri, (2009) "A Cross-Layer Design for Dynamic Resource Block Allocation in 3G Long Term Evolution System", In IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS), pp. 929-934.

## Authors

**Mohamed A. Abd El-Gawad** received the B.Sc. from faculty of engineering, Assiut University, in 2006. He joined the full time M.Sc program in the Faculty of Engineering, Arab Academy for Science & Technology and Maritime Transport (AAST). He is a Teaching Assistant in the National Telecommunication Institute, Cairo, Egypt. His current research interests are wireless communications and resource management



**Mohsen M. Tantawy** received the M.Sc. degree from Cairo University, Egypt in 1998 and the Ph.D. from Ain Shams University, Egypt in 2003. He is an associate professor in the Network Planning Department, National Telecommunication Institute, Cairo, Egypt. His current research interests include wireless communications and network security.



**Mohamed S. El-Mahallawy** received the B.Sc. and M.Sc. from Arab Academy for Science, Technology & Maritime Transport (AASTMT), in 1998 and 2002, respectively. He received the Ph.D. from Cairo University in 2008. He is an associate professor with the department of Electronics and Communications Engineering, AASTMT, Cairo, Egypt. His current research interests include signal processing, Artificial Intelligence, remote sensing.

