A Novel Cross-Layer Scheduling Algorithm for Long Term-Evolution (LTE) Wireless System

Mohsen M. Tantawy, Adly S. Tag Eldien, and Rokaia Mounir Zaki

Abstract — The Third Generation Partnership Project (3GPP) is specifying the Long Term Evolution (LTE) of third-generation cellular systems to meet demands for higher user bit rates. In this paper, we propose a novel Quality of Service (QoS)-guaranteed cross-layer scheduling algorithm for LTE system that allocates resources as resource blocks and also provide the modulation and coding scheme, among users with different traffic loads. Numerical results demonstrate that the proposed algorithm can provide better behavior for QoS-based services than other previous resource allocation algorithms.

Key words — Cross-Layer scheduling algorithms, 3GPP LTE, OFDMA, QoS, GBR, Non-GBR.

I. INTRODUCTION

The 3GPP Long Term Evolution (LTE) represents a major advance in cellular technology. LTE is designed to meet carrier needs for high-speed data and media transport as well as high-capacity voice support well into the next decade. LTE is well positioned to meet the requirements of next-generation mobile networks. It will enable operators to offer high performance, mass-market mobile broadband services, through a combination of high bit-rates and system throughput – in both the uplink and downlink – with low latency [1].

Downlink and uplink transmission in LTE are based on the use of multiple access technologies: specifically, orthogonal frequency division multiple access (OFDMA) for the downlink, and single-carrier frequency division multiple access (SC-FDMA) for the uplink [2], [3]. OFDMA is a multiple access scheme based on orthogonal frequency division multiplexing (OFDM) digital modulation scheme where multiple user equipments (UEs) get assigned subcarriers or subsets of them in order to be served simultaneously. OFDMA is the most flexible scheme, which allows relatively easy assignment of radio resources of either portions of time or frequency to users [4], [5].

Adly S. Tag Eldien is with faculty of engineering, Benha University, Electrical Engineering Department, Cairo, Egypt (e-mail: adlymerg@yahoo.com).

Rokaia Mounir Zaki is with faculty of engineering, Benha University, Electrical Engineering Department, Cairo, Egypt (e-mail: rukaia.enam@feng.bu.edu.eg).

1Mohsen M. Tantawy is with the National Telecommunication Institute, Network planning Department, Cairo, Egypt (e-mail: mohsen@nti.sci.eg).

In LTE, the configured classes have been specified in two categories of bearers, Guaranteed Bit Rate (GBR) and Non-Guaranteed Bit-Rate (Non-GBR) bearers [6].

QoS Class Identifier (QCI) is an index that identifies a set of locally configured values for three QoS attributes: priority, delay and loss rate. QCI is signaled instead of the values of these parameters. The standard QCI classes are shown in Table I[7], [8]. When a connection (or bearer) is established between the UE and the LTE core network a QCI is specified. This defines whether the bearer is guaranteed bit-rate or not [5], [9].

| TABLE I STANDARDIZED QoS CLASS IDENTIFIERS (QCIs) FOR LTE [19]. |
|-----------------|-----------------|-----------------|-----------------|
| QCI             | Resource type   | Priority        | Delay budget    | Services        |
| 1               | GBR             | 2               | 100 ms          | VoIP            |
| 2               | GBR             | 4               | 150 ms          | Video call      |
| 3               | GBR             | 5               | 300 ms          | Streaming       |
| 4               | GBR             | 3               | 50 ms           | Real time gaming|
| 5               | Non-GBR         | 1               | 100 ms          | IMS signaling   |
| 6               | Non-GBR         | 7               | 100 ms          | Interactive gaming|
| 7               | Non-GBR         | 6               | 300 ms          | Applications with TCP: |
| 8               | Non-GBR         | 8               | 300 ms          | browsing, email, file |
| 9               | Non-GBR         | 9               | 300 ms          | download, etc.  |

Adaptive Modulation and Coding (AMC) is the mechanism used for link adaptation to improve data throughput in a fading channel. This technique varies the downlink modulation coding scheme based on the channel conditions of each user. Inside each subcarrier AMC is applied with three modulation schemes (QPSK, 16QAM and 64QAM) and variable code rates [10].

In order to exploit the advantages of OFDMA multiple access scheme and guarantee the QoS of different services with distinct traffic patterns and requirements, scheduling is importance in LTE. Scheduling algorithms are responsible for selecting which UEs will have access to the system resources and with which configuration [1], [5], [11]. Therefore, in this paper, we deal with downlink scheduling algorithms for capacity maximization in multiservice scenarios.

This paper is organized as follows: Section II presents an overview for the related works. Section III will establish the system model and analysis. In Section IV, the detailed cross-layer resource block allocation algorithm for LTE system is
presented. In Section V, we investigated the performance of the proposed scheme and compare it with the traditional fixed and fair schemes. Finally, we conclude the paper in Section VI.

II. RELATED WORK

In [12], the authors propose an adaptive proportional fair scheduling algorithm for LTE which adjusts the scheduling priority according to individual user’s channel condition. This method gives more scheduling probability to the users who are under poor channel condition for a long period of time, and avoids the users whose channel conditions are favorable occupying too much resource. It enhances the fairness with a limited degradation of whole system throughput.

As discussed in [13], a proposed cross-layer resource allocation algorithm which takes the channel quality variance, real-time services and non-real-time services and minimum transmission rate into account has high performance in terms of user fairness.

[14] has considered a flexible OFDMA wireless system, in which the fixed and fair allocation algorithms are explained. The results show the drawback of these algorithms for performance and fairness.

And in [15], the proposed cross-layer maximum weighted capacity (MWC) based resource allocation provides a much better QoS than maximum capacity (MC) and proportional fairness (PF) at a high total data arrival rate, while maintaining nearly the highest system capacity and costing a similar complexity.

III. SYSTEM MODEL

LTE system uses resource block (RB) which is the basic unit of exchanging information in both downlink and uplink [16]. The radio resource that is available for a user in the downlink 3GPP LTE system is defined in both frequency and time domains and is called a resource block (RB). In the frequency domain, the RB consists of 12 consecutive subcarriers (180 kHz total bandwidth) and in the time domain it is made up of one time slot of 0.5 ms duration as shown in Fig. 1.

Each 1ms Transmission Time Interval (TTI) consists of two slots (Tslotted). Each user is allocated a number of so-called resource blocks in the time–frequency grid. Slots consist of either 6 or 7 OFDM symbols, depending on whether the normal or extended cyclic prefix is employed. The more resource blocks a user gets, and the higher the modulation used in the resource elements, the higher the bit-rate [17], [18]. OFDM provides a physical basis for the multiplexed shared channels, where the total bandwidth B is divided into N subcarriers and each subcarrier have a bandwidth $B_s$ equal B/N. Let $m = \{1, 2, \ldots, N_{RB}\}$ denote the RB index set. We consider an OFDM system with K users, let $\Omega_k$ denote the index set of RBs allocated to user k and k = 1, ..., K. For simplicity, we assume that each RB is occupied by only one user and uniform power allocation across all subcarriers [13].

There are many ways to measure the capacity of wireless systems. A well-known definition of capacity is the one provided by Shannon which consists in the maximum achievable set of rates in multiple access channels with an arbitrarily small probability of error [5].

Fig. 1. The Downlink time–frequency resource grid [18].

One resource block contains $N_{RB}$ subcarriers; so the channel capacity of a RB in the OFDM multiplex can be expressed as :

$$C_m = N_{RB} log_2\left(1 + \frac{\bar{E}}{T_s N_0 B_n \alpha_m^2}\right)$$

where $N_0$ is the power spectral density of the additive white Gaussian noise (AWGN) channel (assuming perfect channel estimation), $\alpha_m$ is the multipath channel attenuation coefficient of the resource block m assumed here as Rayleigh distributed random variable, $T_s$ is the symbol time and $\bar{E}$ is the symbol energy. It has been considered that in an OFDM system the carrier spacing ($B_n$) is equal to the inverse of the symbol time ($T_s$).

$$C_m = N_{RB} log_2\left(1 + \frac{\bar{E}}{N_0 \alpha_m^2}\right)$$

The SNR ratio ($\text{SNR}_{m_k}$) is:

$$\text{SNR}_{m_k} = \left(\frac{\bar{E}}{N_0}\right) \alpha_{m_k}^2$$

effective signal-to-noise ratio for the $k$-th user at resource block m;

$$C_{m_k} = N_{RB} log_2\left(1 + \text{SNR}_{m_k}\right)$$

The capacity correction factor (F) is considered here,
where $T_{CP\ in\ TTI}$: the duration of cyclic prefix in one TTI and $T_{TTI}$: is the Transmission Time Interval.

So the capacity of subcarrier is updated to

$$C_{m,k} = F N_{RB}^{SC} B_n \log_2 (1 + \text{SNR}_{m,k})$$  \hspace{1cm} (5)

The maximum capacity for the $k$-th user is given by:

$$R_k = \sum_{m \in \Omega_k} C_{m,k}$$  \hspace{1cm} (6)

The resource allocation scheme is designed to maximize the system throughput ($J$):

$$J = \sum_{k=1}^{K} R_k$$  \hspace{1cm} (7)

IV. ALLOCATION ALGORITHM

The algorithms proposed in this paper are based on the estimation of the channel capacity belonging to the resource blocks. From the previous equations, the channel capacity that a user achieves by the assignment of a certain RB can be determined.

The first allocation algorithm is the fixed allocation algorithm that proposed in [14]. It approximates a uniform capacity distribution, where each user obtains the same number of RBs without any consideration to the channel parameters, so the users with the best channel conditions obtain the same number of resources with respect to the users with the worst channel conditions.

A. Fair allocation

To avoid the previous drawback another allocation algorithm is proposed in [14]. The algorithm assigns the RBs to each user depending on the best channel conditions and channel capacity.

Where:

- $S$: the set of free resource blocks;
- $R_k$: the capacity assigned to the $k$-th user.

1. Initialization:
   a) Set $R_k = 0$, $\Omega_k = \Phi$ \hspace{0.5cm} $k=1,...,K$;
   b) If $\text{SNR}_{n,k} \geq \text{SNR}_{m,k}$ (m,n $\in$ S) ; assign RB m to user k, i.e., add resource block m to $\Omega_k$. Remove RB m from S. Update $R_k$ according to (6).
2. Find the user $k$ so that $R_k \leq R_u$ for each user u; and repeat (1-b) for the corresponding user k.
3. Repeat step 2 until $S = \Phi$.

The described algorithm, after an initialization phase, assigns to each user a RB within which the user has the highest SNR. Subsequently all the remaining RBs are assigned through an iterative process: the user with the lowest amount of capacity is selected and a RB with the best SNR is allocated to him.

B. Proposed QoS-based allocation algorithm

The LTE QoS mechanisms follow a network initiated QoS control based on GBR and non-GBR bearers, which is a class-based packet forwarding treatment for delivering real-time and non-real-time traffic [6].

The previous algorithm may lead to the case where the users occupying air resources do not have a high demand for resources, while other users with urgent traffic demands are not allocated enough resources due to poor channel gains. The proposed algorithm considers the QoS information, e.g., the queuing delay.

So we generate a weight value for each user, which depends on the guaranteed QoS parameters as priority, delay and data rate. As in [15], the system throughput will be:

$$J = \sum_{k=1}^{K} W_k R_k$$  \hspace{1cm} (8)

where $W_k$ denotes the weight for user $k$ which indicates the QoS information for user $k$. Delay satisfaction indicator $(DS_k)$ indicates the connection delay for user $k$.

$$DS_k = L_\beta - S_k$$  \hspace{1cm} (9)

where $L_\beta$ is the delay bound for the traffic type, which is the class-$\beta$ QoS traffic, $S_k$ is the waiting time for the data of user $k$, which is the duration between the arriving time and the serving time for the data and neglecting the effect of guard slot.

Let $U_{k,m}$ denote the weight of the data corresponding to resource block $m$ of user $k$, and given by:

$$U_{k,m} = \begin{cases} \left[ \beta_m / DS_k \right] \log(D_{k,m} + 1) & DS_k > 0 \\ \beta_m \log(D_{k,m} + 1) & DS_k \leq 0 \end{cases}$$  \hspace{1cm} (10)

where $\beta_m$ is the class-$m$ QoS coefficient or priority (given in table 1) and $D_{k,m}$ is the amount of data arriving in RB $m$.

So the weight of user $k$ will be:

$$W_k = \sum_{m \in \Omega_k} U_{k,m}$$  \hspace{1cm} (11)

The proposed scheduling scheme, assigns a higher weight to the data packets with a less DS, i.e., the data with the least DS should be sent out first.

Letting $R_k/W_k$ denote the rate-to-weight ratio (RWR), we employ the following proposed RB allocation scheme, where the user with the lowest RWR is allowed to pick resource blocks in each iteration:

1. Initialization:
a) Set $R_k = 0$, $\Omega_k = \emptyset$ for $k=1, \ldots, K$; Sort $W_k$ in the descending order.

For $k=1$ to $K$:

b) If $SNR_{n,k} \geq SNR_{m,k}$ ($m,n \in S$); assign RB $m$ to user $k$, i.e., add RB $m$ to $\Omega_k$. Remove RB $m$ from $S$. Update $R_k$ according to (6).

2. Find the minimum $R_k/W_k$ ($k = 1, \ldots, K$), and repeat (1-b) for the corresponding user $k$.

3. Repeat step 2 until $S = \emptyset$.

The proposed algorithm, after an initialization phase, assigns to each user a resource block within which the user has the highest SNR. Subsequently all the remaining resource blocks are assigned through an iterative process: the user with the minimum RWR is selected and a resource block, which provides the best SNR to that user, is assigned to him.

V. NUMERICAL RESULTS

In this Section the numerical results obtained, the results show performance of the proposed scheduling scheme, in terms of the total system bandwidth efficiency as a function of SNR in LTE downlink direction.

Table II summarizes the system parameters. A simple channel estimation method is assumed and in which BS estimates instantaneous CQI with the previous CQI feedback.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>System bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Number of subcarriers</td>
<td>1024</td>
</tr>
<tr>
<td>Number of occupied subcarriers</td>
<td>600</td>
</tr>
<tr>
<td>Subcarriers per RB ($N_{RB}^{SP}$)</td>
<td>12</td>
</tr>
<tr>
<td>RB bandwidth</td>
<td>180 kHz</td>
</tr>
<tr>
<td>Number of RBS</td>
<td>50</td>
</tr>
<tr>
<td>Total transmit power</td>
<td>1 W</td>
</tr>
<tr>
<td>Power distribution</td>
<td>Uniform</td>
</tr>
<tr>
<td>Number of active users</td>
<td>9</td>
</tr>
<tr>
<td>Transmission Time Interval (TTI)</td>
<td>1 ms</td>
</tr>
<tr>
<td>Channel type</td>
<td>AWGN channel</td>
</tr>
<tr>
<td>Channel attenuation coefficient</td>
<td>Rayleigh distributed</td>
</tr>
<tr>
<td>Modulation/coding rate settings</td>
<td>QPSK: $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}$</td>
</tr>
<tr>
<td></td>
<td>16QAM: $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}$</td>
</tr>
<tr>
<td></td>
<td>64QAM: $\frac{2}{3}, \frac{3}{4}, \frac{4}{5}$</td>
</tr>
</tbody>
</table>

Fig. 2 shows the relationship between SNR and the total system capacity. It demonstrates performance of proposed algorithm, compared to fixed and fair allocation schemes. The maximum capacity achieved by the proposed algorithm is much higher than that for fixed scheme and approximately the same as the fair. It also demonstrates the performance of the proposed QoS-based allocation algorithm, compared to the fixed and fair allocation algorithms in bit per second per hertz. The fixed resource allocation scheme achieves low performance while the fair and the proposed scheduling algorithms achieve approximately the same overall capacity for all traffic types.

Fig. 3 shows the relationship between SNR and the total system capacity for Non-GBR and GBR traffic. In Fig. 3, the fixed and fair allocation algorithms do not provide any priority for GBR services over Non-GBR. On the other hand, the proposed QoS-based scheduling algorithm provides better capacity for GBR than that for Non-GBR services.
Fig. 4. Bit error rate versus SNR

Fig. 4, point out the better behavior of the proposed algorithm in comparison with the fixed one in terms of the BER parameter. It is also approximately the same as the fair scheme.

Through the second part of the simulation work, we fix the SNR at 20 dB and the number of active users varied from 2:74 user divided equally between the GBR and Non-GBR traffic types. The results in Fig. 5 show that, the total system capacity for the proposed and fair allocation algorithms is much better than that for the fixed allocation algorithm. As we predict; the capacity of the fixed scheme is constant with the increase of user numbers while the capacity of the proposed and the fair schemes are approximately the same and increase with increasing the number of active users.

As indicated from Fig. 6, the fixed and fair scheduling algorithms assign the same amount of capacity for both the GBR and Non-GBR users, so both of them do not provide any priority for GBR services. The proposed algorithm provides high performance for GBR services than Non-GBR services.

VI. CONCLUSION

In this paper, a Multi-User QoS Guaranteed Based on cross-layer resource allocation algorithm in downlink LTE system is proposed to ensure the QoS while satisfy the requirement of multi-service and maintain the throughput and fairness performances, based on QoS information from the data link layer. The results show that, the proposed algorithm has approximately the same performance compared to fair allocation algorithm and much better capacity than the fixed schemes. Moreover, it provides a better QoS and performance with guaranteed bit rate services in LTE systems than that of the fixed and fair algorithms.

REFERENCES


BIOGRAPHIES

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 currently an associate professor in network planning department in National Telecom. Institute (NTI), affiliate of the Ministry of Communication and Information Technology.

Adly S. Tag Eldien received the B.S. degree in Electronics and communication, Benha University in 1984 and the M.Sc. in computer based speed control of single phase induction motor using three level PWM with harmonic elimination, Benha University, in 1989. The Ph.D. in optimal robot path control, Benha University, in 1993. He is currently an Association prof. in shoubra faculty of engineering and Manager of Benha university network and information center. and his research interests include, Robotics, Networks, Communication.

Rokaia M. Zaki received the B.S. degree in Electrical Engineering, Communication from Zagazig University, in 2000 and the M.Sc. degree from Benha University, Egypt in 2007. Now she is a Student of Ph.D degree in Communication engineering in Benha University, Her research interests include, Mobile communication, wireless Networks.