Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2019) 27: 2021 – 2035

doi:10.3906/elk-1808-56

Mobility and traffic-aware resource scheduling for downlink transmissions in LTE-A systems

Önem YILDIZ*, **Radosveta İvanova SOKULLU**
Department of Electrical and Electronics Engineering, Faculty of Engineering, Ege University, İzmir, Turkey

Received: 06.08.2018 • **Accepted/Published Online:** 19.12.2019 • **Final Version:** 15.05.2019

Abstract: As new cellular networks support not only voice services but also many multimedia applications, the requirements for reliable data transmission at high speeds create heavy load on the system. Even though LTE/LTE-A technology takes action towards alleviating this load, it is still necessary to manage resources effectively because of the inadequacy of the available radio resources. Thus, the scheduler at the MAC layer of the base station plays a very important role in resource allocation to the user. In this study a novel algorithm for resource allocation in mobile environments is presented, with two variations addressing different input traffic. The idealized case (I-MAS algorithm) relates to the full-buffer model, while the realistic case (R-MAS algorithm) takes into consideration the specific characteristics of the incoming user traffic. The paper includes performance evaluation of the suggested algorithms in terms of mean and edge throughput, system fairness, and BLER and comparison with well-known algorithms like the round robin (RR) and best CQI (B-CQI) (full-buffer model) and their extensions for real-life traffic models, RR Traffic and B-CQI Traffic, respectively. When the simulation results are examined, it can be seen that the I-MAS and R-MAS algorithms maximize the throughput while at the same time distributing the resources fairly among the users. They also prove to be quite robust in mobile environments even at higher user speeds.

Key words: LTE-A, scheduling algorithms, throughput, fairness, mobility

1. Introduction

In recent years, both the capacity of wireless networks and the capabilities of mobile devices have increased in unprecedented proportions, transforming wireless and cellular communication from a luxury into a commodity. On the other hand, in line with user demand, the number and the scope of applications supported by mobile devices are growing incredibly fast. For example, online games can be played on smartphones, players can engage in video calls, or one user can browse a news website while another user carries out a video conference in the same physical environment. Enhanced voice services together with a greatly increased number of multimedia applications constitute an enormous portion of the traffic and create a major concern for cellular network operators. In order to cope with the growing traffic burden, service providers need to take into account the requirements of each traffic type (i.e. high bandwidth, low latency required for quality of service (QoS)) in relation to the specific application. Their efforts are supported by the LTE-A network standard developed by the 3rd Generation Partnership Project (3GPP), which allows to meet these requirements within the available frequency bands. A major approach in coping with the intensified mobile traffic using the limited available bandwidth is scheduling the resources. eNodeB is responsible for scheduling and allocating network resources in

*Correspondence: onem.yildiz@gmail.com

terms of bandwidth and time among the users. The operation is performed by the packet scheduler residing in the MAC layer of eNodeB. The objectives of the scheduling algorithms are efficient resource allocation, ensuring fairness among users, and achieving high system throughput as well as transmission within acceptable error and delay limits. The design of the scheduling algorithm is crucial for accurate and efficient resource planning and allocation. Furthermore, selection of the priority performance metrics is an important step that affects the choice of the specific scheduling algorithm.

Even though there are a number of such algorithms already suggested, the abundance of network scenarios and the plethora of constantly evolving user applications with a wide range of performance requirements makes the design of scheduling algorithms a hot research subject open to development.

In this study, we provide a detailed performance analysis of two algorithms frequently encountered in the literature, namely round robin (RR) and best CQI (B-CQI), which prioritize fairness and throughput performance metrics. Their advantages and disadvantages are laid out in a detailed and ratiocinated way. Based on that a novel scheduling algorithm is proposed, which achieves a good balance between user fairness and network throughput while at the same time ensuring successful data transmission with the least possible packet error in cases of user mobility. This new scheduling algorithm has been carefully scrutinized under various scenarios, involving real-time and nonreal-time traffic users as well as different cases of mobility and different network configurations. The effects of user speed and antenna configurations on the performance metrics have been compared and analyzed.

The remainder of the paper is organized as follows: Section 2 gives an overview of related work and states the contributions of this work; in Section 3 a relevant summary of the LTE-A standard and its adopted resource allocation procedures are presented. Section 4 defines the performance metrics while Section 5 describes the system model used. Details of the proposed algorithm are laid out in Section 6. Section 7 covers the simulation scenarios followed by results and discussion. The paper is concluded in Section 8.

2. Related works

For cellular networks the scheduling and resource allocation process is different for downlink and uplink multiple access. Since this study concentrates on downlink transmission, only the scheduling algorithms related to downlink are mentioned. These algorithms play an important role in resource allocation, and because different performance metrics can be optimized, a correct and efficient scheduling algorithm design is of utmost significance. Furthermore, defining the priority performance metric is an important step as it affects the choice of scheduling algorithm. Major goals are fair resource allocation among users, obtaining high system throughput, and achieving minimum error transmission.

Some of the classic scheduling algorithms available in the literature are resource fair (RF), proportional fair (PF), round robin (RR) and best CQI (B-CQI). The design of these algorithms is based on the assumption that users are "full-buffer users", i.e. their output buffers always have data waiting to be transferred in an unlimited amount so there is a continuous input data stream to the system. Since data transmission never ends, the number of users who want to benefit from the network always remains constant, as there is no such thing as a user leaving the system.

The RF scheduling algorithm [1] aims to distribute the available resources (in terms of resource blocks, RBs) equally among all users. In order to distribute N RBs between K users, $b_k = \frac{N}{K}$, the $\forall k$ parameter is calculated and resources are distributed to the users according to the obtained bk value. Here, if $\frac{N}{K}$ is a noninteger, some users will be allocated the number of RBs rounded to the highest integer less than $\frac{N}{K}$, which is expressed by $\lfloor \frac{N}{K} \rfloor$, while others will be allocated the number of RBs rounded to the lowest integer above the $\frac{N}{K}$ value, expressed by $\lceil \frac{N}{K} \rceil$. RF provides at least a minimum of service for all users by ensuring that resources are equally distributed, but it does not perform well in terms of throughput.

The PF scheduling algorithm [2] is aimed at balancing between fairness in resource allocation and high system throughput. For this purpose, a special parameter, M_n, is proposed, calculated as given in Eq. (1):

$$M_n = \frac{d_n}{r_n}. \tag{1}$$

In Eq. (1), n is the number of users, d_n is the instant data rate of the user, and r_n is the average throughput value of each user. Based on this equation, the PF algorithm allocates resources to the user with the highest M_n value that has utilized the least resources in the past. For this reason, even though the PF increases the system's fairness by guaranteeing service to users with minimum data rates, it reduces the cell throughput by allocating resources to users with the worst performance.

The RR scheduling algorithm [3, 4] is the one that achieves the highest degree of fairness in resource allocation among different users. However, this algorithm is not channel-aware, i.e. it assumes that the transmission environment does not change over the duration of the transmission and it does not take channel conditions into account when scheduling resources. RBs are assigned to the users in a sequential manner. In other words, distribution is made based on the assumption that users will benefit from the sources on an equal basis so the RR algorithm achieves a high degree of fairness among users. However, due to bad propagation channel conditions, some users might be experiencing poor transmission, so RR scheduling might result in low throughput performance of the system as a whole.

The B-CQI scheduling algorithm [5, 6] allocates RBs to the user with the best radio link. For the allocation process to occur, reference signals are initially broadcast by the eNodeB and are used by the UEs to measure their channel quality. The resultant CQI information is sent from the user to the base station, where high CQI values mean that the propagation channel conditions are good. The eNodeB allocates RBs to users starting with the one having the highest CQI, thus maximizing the aggregated system throughput. Users located away from the base station often report poor CQI values because of poor channel conditions. In principle, the B-CQI algorithm will not allocate resources for users on the cell edge when there are users located near the base station (with better CQI values). For this reason, the B-CQI scheduling algorithm significantly increases the throughput value for some users but deprives users close to the cell edge or users with low CQI from the resources. As a result, while the B-CQI algorithm maximizes cell throughput by utilizing the differences in users' channel conditions, the level of fairness is very low, especially because in many cases it does not provide any service to cell edge users.

A major deficiency of all the algorithms described above is that they have been designed and studied with the notion of full-buffer input packet queue, which is not necessarily true in real life. Furthermore, most studies do not exclusively consider the issues of user mobility and mobility-aware scenarios. The authors in [7] considered the basic traffic model for real-time services and a full-buffer model for nonreal-time services, but the study was limited to static users only. In [8], the VoIP traffic model was examined for a static user scenario. In [9], a new downlink scheduling algorithm considering real-time services was developed for nonmobile users. In the literature, there are a few studies including mobility-aware scenarios; however, they only consider a limited number of traffic models. The authors in [10] evaluated and compared various scheduling algorithms for mobile users using a full-buffer traffic model. In [11, 12], the authors evaluated the performance of the schedulers for nonreal-time flows and real-time flows such as VoIP and video traffic for different user speeds. In [13], the performance of several scheduling algorithms was considered including mobility but only for a single cell with interference scenario for different flows.

The main contributions of our work are as follows: a novel resource scheduling algorithm for the downlink transmission in LTE-A systems is proposed, which combines the advantages of RR (high degree of fairness) and B-CQI algorithms (channel awareness) while aiming at high system throughput and mobility awareness. Besides the "full-buffer" traffic model, the proposed algorithm is also examined in cases of more realistic input traffic models including delay-tolerant and delay-sensitive traffic flows. These improved models take into consideration the characteristics of these flows as well as the actual amount of information (bits) to be transmitted in the transmission buffer of each user in the network.

3. LTE resource scheduling and allocation procedures

The QoS offered to mobile users in wireless networks has become an important design issue because of the increasing density of mobile traffic and the difficulty of existing frequency bands to support this density. Efficient allocation of the available resources plays a key role. This network function resides with the scheduler at the MAC layer of the eNodeB and ensures optimal allocation of common radio resources among users for both downlink and uplink transmissions.

In LTE, resources are defined both in time and frequency domains. The resource unit in the time-frequency domain is the "Resource Block" (RB), expressed by 180 kHz bandwidth (12 subcarriers) and 0.5 ms duration (one slot). Resource allocation algorithms operate to distribute the available RBs to the users with the goal of meeting the specific QoS requirements under given system restrictions and optimization criteria.

LTE allows flexible user bandwidth allocation from a minimum of 1.4 MHz up to 20 MHz, where the number of RBs assigned to a user can vary accordingly. Resource scheduling is the process of distributing the available system resources in such a way as to provide the maximum required QoS among the users satisfying the optimization criteria of the network or the network provider. This process can be performed in two different ways: channel-aware and channel-unaware. Channel-unaware scheduling assumes that there are no errors in the transmission and that the channel characteristics do not change over time. Channel-aware scheduling, on the other hand, is done based on the assumption that the characteristics of the propagation channel between the eNodeB and the user vary continuously. The UE is responsible for sending information about its channel quality to the eNodeB on a regular basis.

4. Performance metrics

In order to evaluate the efficiency of resource distribution for a given scheduling algorithm, the following important performance metrics have been defined. Since there might be variations in the literature, explicit definitions are presented below.

The parameter "fairness" expresses how fairly the available resources are shared among the users in the network. According to Eq. (2), also known as Jain's fairness index [14], when resources are allocated to n users and the average throughput value of the ith user is X_i, the fairness parameter $f(X_i)$ is calculated as:

$$f(X_i) = \frac{(\sum_{i=1}^{n} X_i)^2}{n \times \sum_{i=1}^{n} X_i^2}. \tag{2}$$

The value of $f(X_i)$ must be between 0 and 1. For values of the fairness parameter close to 1 the resources are said to be distributed at a high degree of fairness among the users.

Throughput is another parameter that stands out in the evaluation of the system. Throughput in general describes the amount of data successfully delivered over the channel but different definitions are available in the literature. This study focuses on average user throughput parameters. Average user throughput ($AvgThr_{UE}$) is the number of bits successfully transmitted (N_{Bits}) from the eNodeB to the user during the simulation period and is calculated according to Eq. (3). Here, the term T_{Sim} refers to the entire simulation period:

$$AvgThr_{UE} = \frac{\sum N_{Bits}}{T_{Sim} \times 10^{-6}}. \tag{3}$$

The average aggregated user throughput in the system is obtained by averaging the throughput values for the users.

Another performance metric investigated in this study is cell edge throughput. In order to obtain the cell edge throughput, an empirical cumulative distribution function (ECDF) graph showing the UE throughput distribution was used. In this graph, the x-axis gives the user throughput values, while the y-axis gives the ECDF values. The first function value that is equal to or greater than 0.05 in the points where the throughput values in the x-axis correspond to the y-axis is accepted as the edge throughput value.

The last parameter investigated is block error rate (BLER), which is used to measure the amount of error in the received data. In this study, the user BLER ($BLER_{UE}$) is considered as defined in Eq. (4). A suitable numerical expression for this is derived from the ratio of the total number of received acknowledgment bits (R_{ACK}) to the total number of expected acknowledgment bits (E_{ACK}). When this ratio is subtracted from 1, the user BLER value is obtained. If the ratio is equal to 1, the BLER value will be 0 because the received and the expected ACK bits are equal:

$$BLER_{UE} = 1 - \frac{\sum R_{ACK}}{\sum E_{ACK}}. \tag{4}$$

5. System model

In this work we consider a general LTE topology consisting of 7 eNodeBs located to form a hexagonal structure at equal distances. Each eNodeB has 3 antennas, so a total of 21 cells have been created. Users are equally and randomly distributed in these cells and are mobile at speeds of 3 km/h, 30 km/h, and 120 km/h corresponding to walking, bicycle, and car speeds, respectively. For users with 3 km/h speed the PedA and for users with speeds 30 km/h and 120 km/h the VehA channel model was used. In addition, to make the mobility more meaningful for users, passing to neighboring cells is allowed based on handover.

From here on this section describes the main scenarios considered: the ideal case scenario (ICS) where traffic models of users are full-buffer and the real-life scenario (RLS) where the users have different traffic models and their different traffic characteristics are taken into account by the system, i.e. the resource allocation algorithm and procedures. The parameters for the ICS and the RLS and their related values are summarized in Table 1.

In the ICS, transmission mode is varied in addition to user speed. Transmission modes were defined as SISO and MIMO, with the number of receiving and transmitting antennas set to 4 in MIMO. In the RLS, in addition to SISO and MIMO, ML-MIMO as defined by LTE-A is used with 8×4 antenna configuration. Furthermore in RLS, all mobile users in the network are set to have the traffic models (ftp, http, VoIP, video streaming, and gaming). To ensure that all traffic types are taken into consideration the proportions of the mixed traffic are determined according to [15]. The specific parameters of each traffic model are also given in Table 2.

Table 1. Simulation parameters for scenarios.

Parameter	Value	
	Ideal case scenario	Real-life scenario
Simulation time	500 TTI	500 TTI
Number of trials	10	5
Number of eNodeBs	7	7
Distance between eNodeBs	500 m	500 m
Channel model	3GPP PedA, 3GPP VehA	3GPP PedA, 3GPP VehA
Tranmission mode	SISO, MIMO (4×4)	SISO, MIMO (4×4), ML-MIMO (8×4)
Number of UEs/sector	5	5
Frequency	2.6 GHz	2.6 GHz
Bandwidth	10 MHz	10 MHz
User speed (km/h)	3, 30, 120	3, 30, 120
Scheduling algorithms	Round robin, Best CQI, I-MAS	Round robin traffic, Best CQI traffic, R-MAS
Traffic model	Full-buffer	FTP, HTTP, VoIP Video streaming, Gaming

Table 2. Traffic model parameters.

Application	Percentage of users	Application characteristics
FTP	10%	Mean file size = 2 Mbytes Mean reading time = 180 s
HTTP	20%	Main object size: Mean = 10710 bytes Embedded object size: Mean = 7758 bytes Mean reading time = 30 s
Video streaming	20%	8 packets/frame Packet size: Mean = 10 bytes Packet sending interval = 6 ms Interarrival time between the beginning of each frame = 100 ms
VoIP	30%	AMR 12.2 codec Total voice payload = 40 Bytes Encoder frame length = 20 ms Voice activity factor (VAF) = 50% Mean talk spurt = 2 s
Gaming	20%	Largest extreme value distribution $f_x = \frac{1}{b} \times \exp(-(\frac{x-a}{b})) \times \exp(-\exp^{-}(\frac{x-a}{b}))$ x= $a - b \times \ln(-\ln(y))$, $y \in [0, 1]$ For packet arrival: a = 55 ms, b = 6 For packet size: a = 120 Bytes, b = 6 Delay constraint = 60 ms

6. Proposed algorithm

In this study, we propose a novel algorithm for downlink resource scheduling in LTE-A systems called MAS (mobility aware scheduling). It balances high fairness and maximizing system throughput taking into account user mobility and different traffic models. In this section, details of the algorithm and its two versions related to idealized input traffic and real-life input traffic (respectively I-MAS and R-MAS) are presented. For the I-MAS, a "full-buffer user traffic" model is use,d while for the R-MAS users' traffic can be real-time or nonreal-time. R-MAS prioritizes real-time users during the resource allocation process. Both versions of the proposed algorithm achieve a meaningful balance between efficiency and fairness and furthermore allow to maintain this balance even when users are mobile (from 3 km/h to 120 km/h). Thus, the main contribution of our work is a scheduling algorithm with a high degree of robustness, which ensures that data transmission is minimally impacted by different degrees of user mobility.

6.1. MAS algorithm for ideal case scenario (I-MAS)

The main concept of the I-MAS algorithm is a TTI-based resource allocation process. In the first TTI, the CQI feedback values are recorded as a matrix relating the number of UEs and the available RBs. Each CQI value saved in this matrix is transformed into a corresponding spectral efficiency value to form a new matrix reflecting both the channel conditions of the users and their throughput. Here, the spectral efficiency term defines the rate of transmitted information over the considered bandwidth. For each RB the spectral efficiency values are compared and the current RB is assigned to the user with the maximum efficiency value. If there is more than one user with the maximum value, one user is randomly selected. After that, the first TTI measures are taken to increase the fairness. At the beginning of each TTI, the algorithm marks the user that was allocated the max number of RBs in the previous TTI. If the same user has max spectrum efficiency in the current slot as well, an additional user-specific counter is activated. After half of the resource blocks are allocated, the algorithm checks whether all resources so far have been allocated to a single user. If that is the case, he is not allocated any more resources in the current TTI. The user having the second best channel conditions is defined and assigned the current RB. Control operations are performed continuously until all RBs are assigned for the current TTI. These operations provide a fair resource allocation among users for each TTI. Another control operation is performed at a different time span (total simulation duration) to ensure that allocations are made fairly on that scale as well. If 80% of the time has elapsed, the RBs are assigned to the next user in turn. In order to increase the fairness, in the beginning of the next TTI the first RB is assigned to the user following the last user served in the previous TTI. The 80% value is an empirical value, which allows to service users who have not benefited enough from the resources so far. During numerous trials it has been observed that this improves the throughput especially for users located at the cell edge and also increases the fairness in the cell.

6.2. Enhanced MAS algorithm for real-life scenario (R-MAS)

The R-MAS algorithm is an improved version of the I-MAS that takes into consideration the specifics of real-time and nonreal-time traffic users in the network. The traffic types are determined according to the details presented in Table 2.

The R-MAS algorithm also operates in a TTI-based manner. During the first step of the algorithm the number of real-time (VoIP, video, game) and nonreal-time (http, ftp) users is determined. If the users are only real-time (rUE) or nonreal-time traffic users (nrUE), the algorithm considers this "uniform traffic" and allocates resources by specifying the users with the best channel conditions for each RB during a TTI. If some of the users are real-time and some are nonreal-time traffic users, the algorithm detects "mixed traffic" and prioritizes real-time users. For this purpose the number of real-time users waiting for service in that TTI (N_{tbrUE}) is considered and the α parameter is determined according to Eq. (5):

$$\alpha = (N_{tbrUE} \times 10\%) \times N_{RB}. \tag{5}$$

The α parameter indicates how much resources should be allocated to real-time traffic users. Here, the 10% value is an empirical value. It allows real-time users an advantage of at least 10%. After α values of the RBs are allocated to real-time traffic users according to their spectral efficiency (tbrUE), the rest of the resources are allocated to the user with the best spectral efficiency regardless of its traffic model. The pseudocode of the algorithm for the first TTI is given in Figure 1.

```
while (current_TTI == 1) ∧ (RB_n ≤ N_RB) do
    calculate bits left;
    if sum(bits_left) > 0 then
        if (∀UE ∈ rUE) ∨ (∀UE ∈ nrUE) then
            RB_n is assigned to tbUE_max;
        else
            calculate N_tbrUE;
            if RB_n < α then
                RB_n is assigned to tbrUE_max;
            else
                RB_n is assigned to tbUE_max;
            bits_left_UE_n ← bits_left_UE_n − packetsize;
    else
        increase the count RB_n by 1;
```

Figure 1. Pseudocode for the first TTI.

For the TTIs after the first one, the information for the user assigned the largest number of resources in the previous TTI ($maxRB\_UE_{TTI-1}$) is also marked and checked during the following steps, similar to the I-MAS, to ensure sufficient fairness. If a single user has benefited from half of the available sources or the counter reaches $N_{RB} \times 50\%$, then the β parameter is calculated as given in Eq. (6):

$$\beta = \frac{(N_{RB} - counter)}{N_{RB}}. \tag{6}$$

The β parameter is recalculated for each RB for users whose counter is activated and multiplied by the so-called efficiency value (eff_tbrUE_{max}, for $RB_n < \alpha$; eff_tbUE_{max}, for $RB_n \geq \alpha$) of the user. If the throughput value of the user is still higher than the throughput value of the other users awaiting service, the resource allocation is made to this user again. However, if the new throughput value is lower than the other throughput values, this RB is now assigned to a new user. This allows other users to benefit from existing resources and the degree of fairness among TTI users is increased. The pseudocode of the algorithm can be seen in Figure 2. Once the simulation time reaches 80% of the total simulation time, the resources are allocated equally to all users awaiting service, which allows to increase fairness, especially for cell edge users.

```
while (current_TTI > 1) ∧ (current_TTI ≤ N_TTI × 80%) do
    counter ← bits_left ;
    while RB_n ≤ N_RB do
        calculate bits left ;
        if sum(bits_left) > 0 then
            if (∀UE ∈ rUE) ∨ (∀UE ∈ nrUE) then
                if maxRB_UE_{TTI-1} = tbUE_max then
                    increase the counter by 1 ;
                    if (RB_n > N_RB × 50%) ∧ (counter > N_RB × 50%) then
                        RB_n is assigned to second_tbUE_max ;
                else
                    RB_n is assigned to tbUE_max ;
            else
                calculate N_tbrUE ;
                if RB_n < α then
                    if maxRB_UE_{TTI-1} = tbrUE_max then
                        increase the counter by 1 ;
                        if (RB_n > N_RB × 50%) ∧ (counter > N_RB × 50%) then
                            if N_tbrUE > 1 then
                                eff_tbrUE_max ← eff_tbrUE_max × β ;
                    else
                        RB_n is assigned to tbrUE_max ;
                else if maxRB_UE_{TTI-1} = tbUE_max then
                    increase the counter by 1 ;
                    if (RB_n > N_RB × 50%) ∧ (counter > N_RB × 50%) then
                        eff_tbUE_max ← eff_tbUE_max × β ;
                    RB_n is assigned to tbUE_max ;
                else
                    bits_left_UE_n ← bits_left_UE_n − packetsize ;
        else
            increase the count RB_n by 1 ;
```

Figure 2. Pseudocode after the first TTI until 80% of the total simulation time.

7. Simulation results and discussion

This section presents the simulation results and analysis in two different groups related to the scenarios described in Section 5. Simulations were carried out using the object-oriented MATLAB-based LTE-A System Level Simulator [16]. Results for the performance metrics defined in Section 4 are presented for different user speeds and antenna configurations.

7.1. ICS results

Figure 3 shows the average user throughput values for users with different speeds for the SISO transmission mode. B-CQI provides the best performance in terms of throughput, closely followed by I-MAS. When the user speed increases, the average user throughput for all resource allocation algorithms decreases but the degradation of B-CQI is the sharpest.

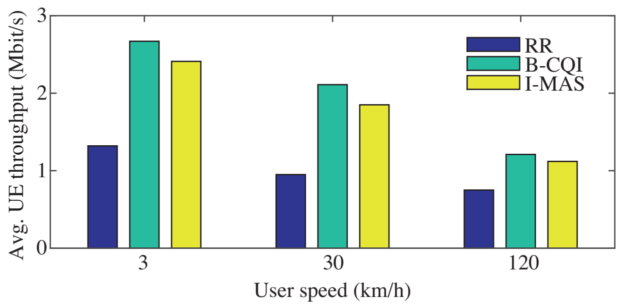

Figure 3. The relationship between average user throughput and user speed for SISO.

Figure 4 shows the average and edge user throughput values of users with different speeds for 4×4 MIMO transmission mode. The average throughput graph shows that even though B-CQI has the best performance, the performances of B-CQI and I-MAS are very close to each other. At high speeds, the average throughput performance of B-CQI sharply declines and the difference between the performances of B-CQI and I-MAS is reduced. RR, as expected, is the worst performing algorithm for MIMO systems. Furthermore, while the B-CQI algorithm does not provide any service to edge users at high speeds, the edge throughput graph shows that RR and I-MAS provide service to these users.

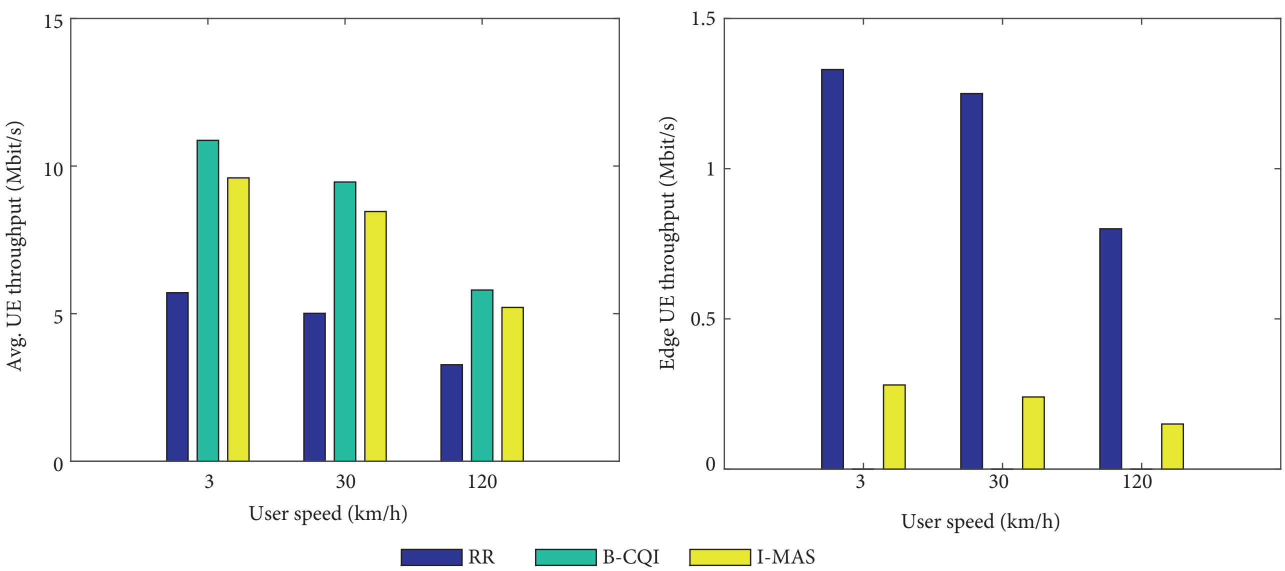

Figure 4. The relationship between throughput and speed for 4×4 MIMO.

Table 3 below summarizes the fairness values of resource allocation algorithms in SISO and MIMO transmission modes for 3 km/h and 120 km/h user speeds. As can be expected, RR allocation results in the highest fairness values for all speeds. I-MAS achieves a higher degree of fairness than that achieved by B-CQI. However, it can be observed that as the user speed increases, the fairness value of B-CQI decreases but the fairness values of RR and I-MAS increase.

Figure 5 shows the effect of user speed on user BLER values of RR, B-CQI, and I-MAS. The BLER values of the users are divided into 3 groups to allow better analysis. The 0–0.1 BLER range signifies the most successful transmission of the packets, while values of 0.4 and above indicate unacceptable error levels. As expected, with increased user speed, the percentage of users with high BLER increases. While RR performs worst at high speeds, I-MAS and B-CQI display quite similar results.

With the use of 4×4 MIMO (Figure 6) the percentage of users with acceptable BLER values at high speeds increases as compared to SISO. The highly computationally intensive B-CQI achieves the best performance for multiantenna techniques; however, the simple I-MAS has also proven to be mobility-resistant, following closely in the performance evaluation.

Table 3. Fairness values for different user speeds for SISO and MIMO in ICS.

Tranmission mode	Speed (km/h)	Resource allocation algorithms		
		RR	B-CQI	I-MAS
SISO	3	0.4487	0.2585	0.2962
	120	0.5199	0.2257	0.3246
MIMO	3	0.6164	0.2524	0.3413
	120	0.6711	0.2469	0.3698

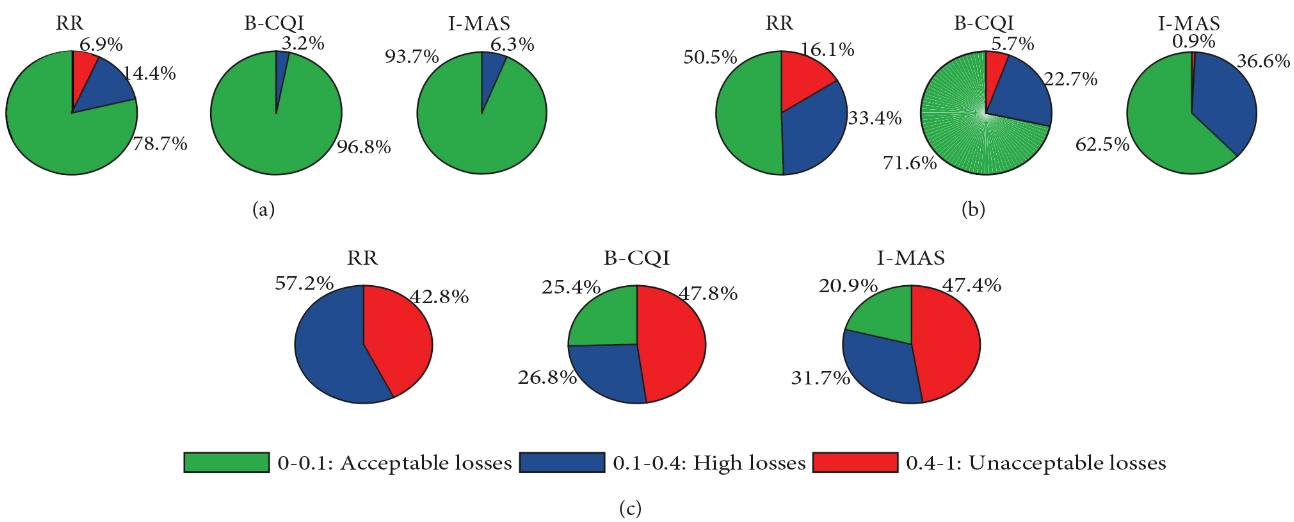

Figure 5. User BLER ratios at different user speeds for SISO in ICS: a) 3 km/h, b) 30 km/h, c) 120 km/h.

In conclusion, when all the BLER graphs for the ICS case are considered, it is obvious that with the introduction of multiple antenna techniques, a significant decrease in the percentage of users with unacceptable BLER values at all user speeds can be achieved.

7.2. RLS results

In this section, the performance of the above discussed algorithms is examined with realistic input traffic mixtures as given by 3GPP specifications [15]. In real-time scenarios, a user does not have bits to transmit all the time, so when he has exhausted the bits in his transmission buffer he leaves the system. Figure 7 gives the results for the average user throughput for different user speeds for SISO, MIMO, and ML-MIMO. While in the case of SISO there are no considerable differences, in the case of MIMO, the proposed R-MAS closely follows the best performing B-CQI.

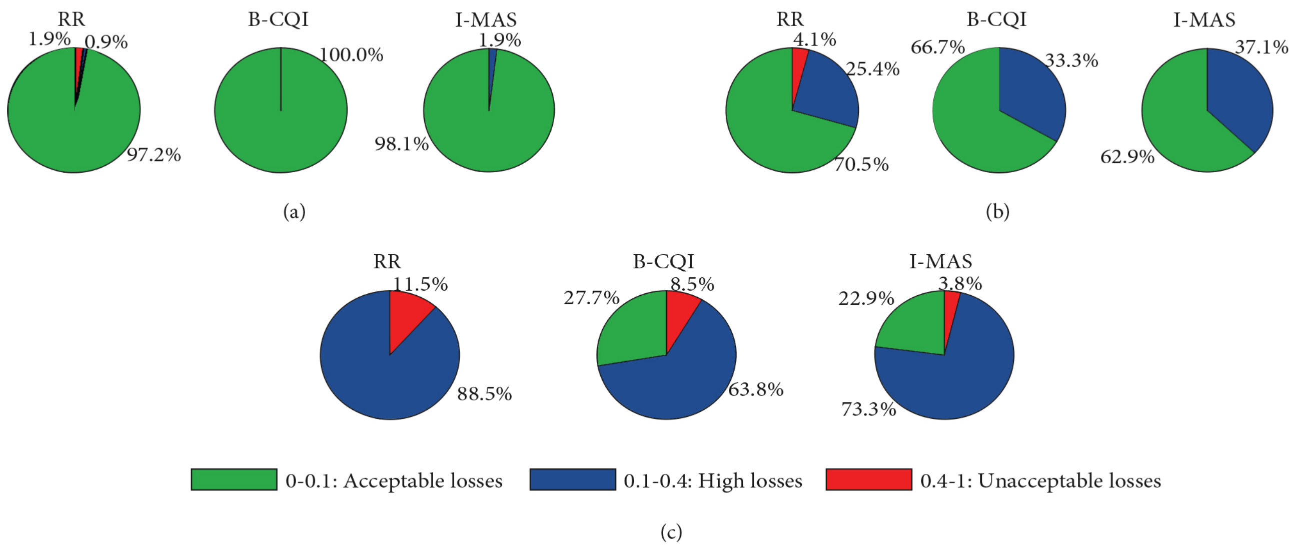

Figure 6. User BLER ratios at different user speeds for 4×4 MIMO in ICS: a) 3 km/h, b) 30 km/h, c) 120 km/h.

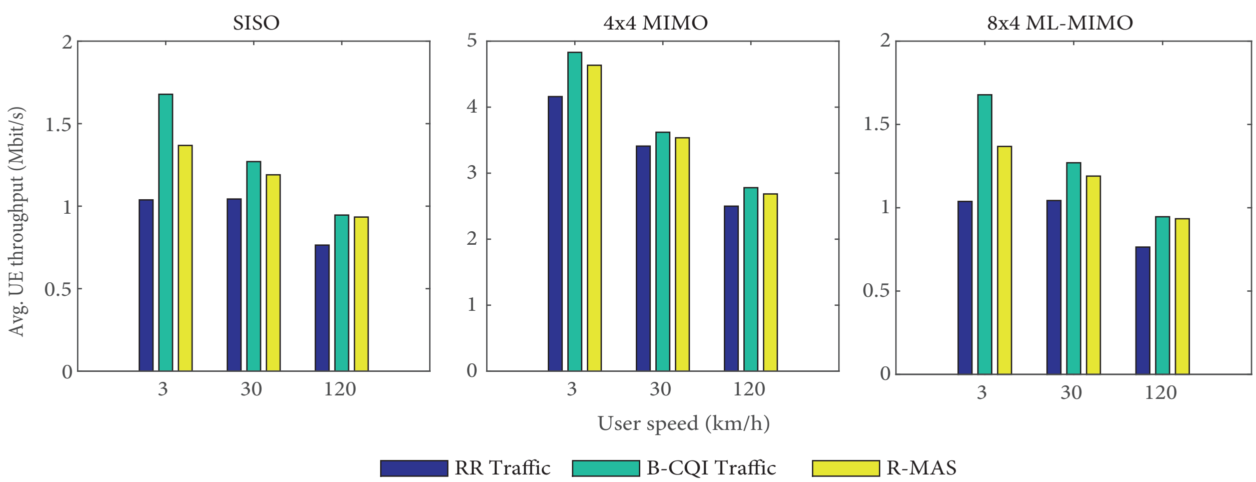

Figure 7. Average UE throughput versus speed for SISO, MIMO, and ML-MIMO.

In the case of MIMO, a significant increment in average throughput values is observed. The proposed R-MAS algorithm closely follows B-CQI Traffic, the best one in terms of throughput performance. Furthermore, with the introduction of an 8×4 multiantenna (ML-MIMO), R-MAS displays a very balanced performance against changing user speeds, and at high speeds it achieves almost the same performance as B-CQI Traffic.

Figure 8 shows edge throughput values for SISO, MIMO, and ML-MIMO. An interesting observation is that B-CQI Traffic allocates resources to edge users, which was not the case in ICS. This is due to the fact that users with the best channel conditions can exhaust their data and then users with worse channel conditions can be served. Another important point is the increase in edge user throughput values with the use of multiple antenna techniques. In particular, for the 4×4 and 8×4 multiple antenna configurations, R-MAS shows a considerable improved performance as compared to the SISO case. As a result of using multiple antennas, the system is much less affected when users move at high speeds. The proposed R-MAS algorithm is particularly robust against changing user speeds and performs well in ML-MIMO systems.

Table 4 summarizes the system fairness values of the algorithms for SISO, MIMO, and ML-MIMO, respectively. For all transmission modes, as the user speed increases, the fairness value increases, because the throughput values of the users are approaching each other. Increasing the number of antennas improves the user throughput values for the RR Traffic and R-MAS algorithms, so the system fairness value increases. However, in B-CQI Traffic, fairness drops because only the throughput values of the users with the best channel conditions increase and the other users still suffer from lack of resources. For this reason, as the difference in user throughput values gradually increases, the fairness value decreases.

Figure 9 shows that the BLER value changes for different user speeds of RR Traffic, B-CQI Traffic, and R-MAS algorithms for 10 MHz constant bandwidth. As can be seen from the graphs, at a low user speed, the large majority of users have acceptable (low) BLER values. However, as user speed increases, the percentage of users with high BLER values for all algorithms also increases. B-CQI Traffic exhibits the best performance while the RR Traffic algorithm exhibits the worst performance in terms of BLER for all speeds. The performance of R-MAS is average but quite robust, especially for higher mobility.

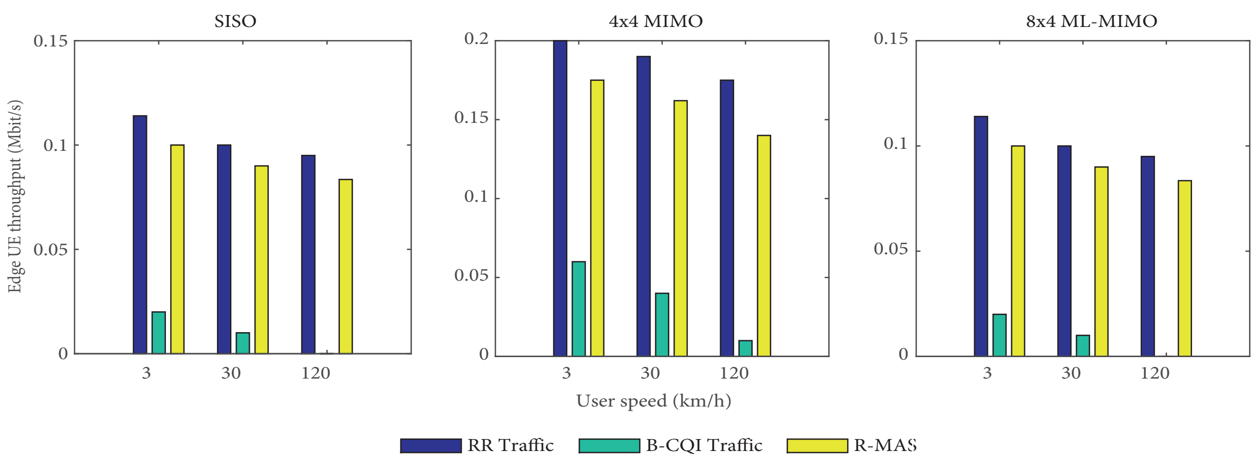

Figure 8. Edge UE throughput versus speed for SISO, MIMO, and ML-MIMO.

Table 4. Fairness values for different user speeds for SISO, MIMO, and ML-MIMO in RLS.

Tranmission scheme	Speed (km/h)	Resource allocation algorithms		
		RR Traffic	B-CQI Traffic	R-MAS
SISO	3	0.2764	0.2231	0.2445
	120	0.3771	0.2358	0.3134
MIMO	3	0.2882	0.2052	0.2691
	120	0.4272	0.2275	0.3552
ML-MIMO	3	0.3392	0.1826	0.2973
	120	0.4408	0.1954	0.3734

Figure 10 focuses on the effect of multiple antenna techniques on the BLER for high user mobility. For the RLS case with MIMO, all algorithms tend to reduce the percentage of users with unacceptably high BLER. The worst performing is RR Traffic, while R-MAS exhibits behavior very similar to that of B-CQI Traffic, which in general achieves the lowest BLER due to its channel-sensitive nature.

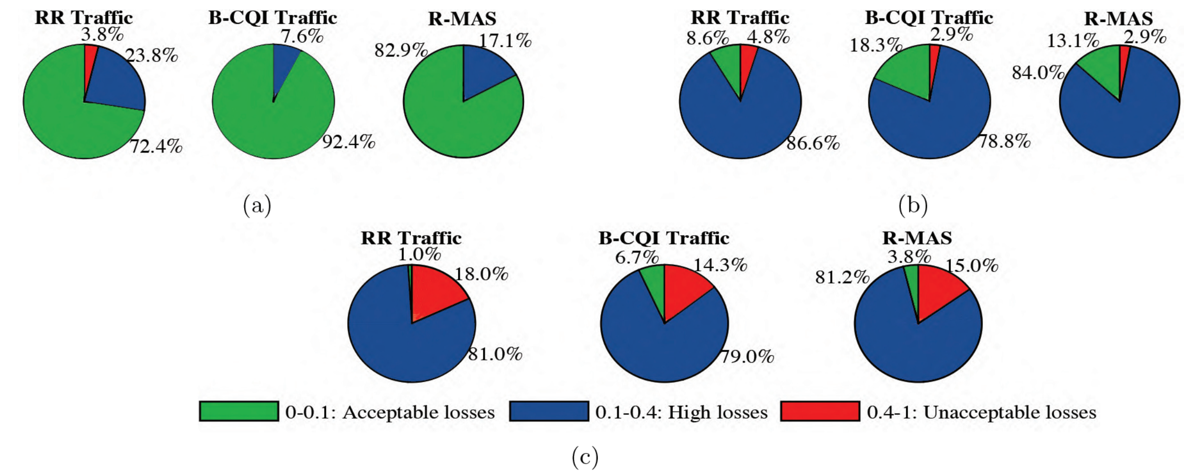

Figure 9. User BLER ratios at different user speeds for SISO in RLS: a) 3 km/h, b) 30 km/h, c) 120 km/h.

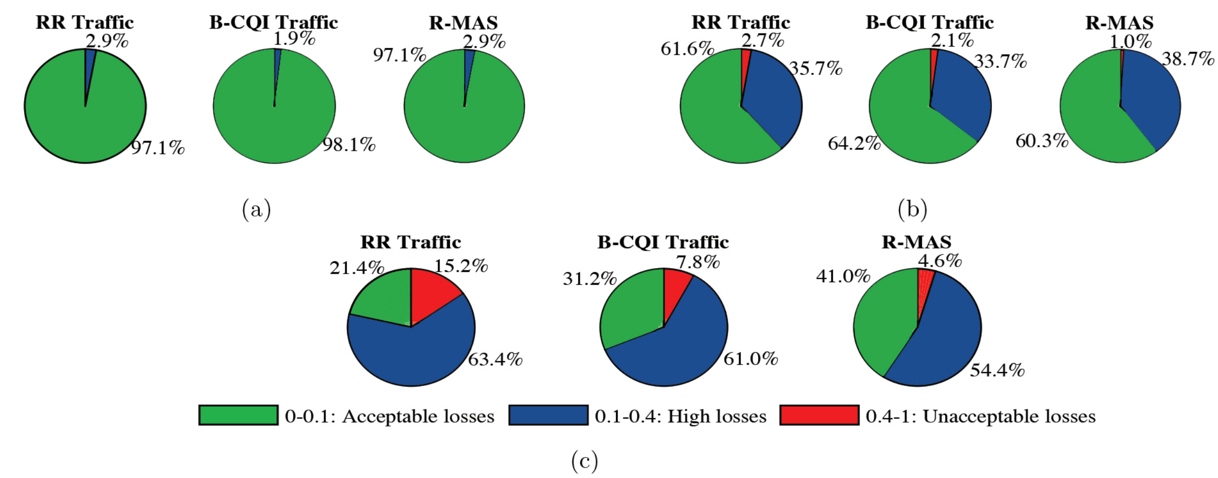

Figure 10. User BLER ratios at different user speeds for ML-MIMO in RLS: a) 3 km/h, b) 30 km/h, c) 120 km/h.

8. Conclusion

In this study, a novel mobility-aware resource allocation algorithm for LTE downlink has been developed with two variations for ideal environments and for real life, where the different characteristics of user traffic are taken into account. The algorithm aims to deliver the data at the highest possible rate and with the lowest number of errors and to provide a fair level of service to all users in the system. The proposed algorithm combines in a very balanced way the simplicity of the RR algorithm achieving high fairness and the high throughput performance of the computationally intensive channel-aware B-CQI algorithm. Two cases are considered: the ideal case scenario (ICS) and the real-life scenario (RLS). The simulation results have shown that for the ICS the proposed algorithm is fairly close to the B-CQI and with increased user speeds, especially at high speeds, the performance gap between the two is greatly reduced. For the RLS, user traffic is modeled as real-time and nonreal-time services. The proposed algorithm maximizes the average user throughput and performs very close to the B-CQI algorithm while also providing service to edge users similar to the RR algorithm. Furthermore, its performance is greatly increased by the introduction of MIMO and ML-MIMO. It is also quite robust against high user mobility and keeps the percentage of users with unacceptable (high) BLER values as low as possible.

References

[1] AlQahtani SA, AlHassany M. Performance modeling and evaluation of novel scheduling algorithm for LTE networks. In: 2013 IEEE 12th International Symposium on Network Computing and Applications; 22–24 August 2013; Cambridge, MA, USA. New York, NY, USA: IEEE. pp. 101-105.

[2] Kwan R, Leung C, Zhang J. Proportional fair multiuser scheduling in LTE. IEEE Signal Processing Letters 2009; 16: 461-464.

[3] Hussain S. Dynamic radio resource management in 3GPP LTE. MSc, Blekinge Institute of Technology. Karlskrona, Sweden, 2009.

[4] Capozzi F, Piro G, Grieco LA, Boggia G, Camarda P. Downlink packet scheduling in LTE cellular networks: key design issues and a survey. IEEE Communications Surveys & Tutorials 2013; 15: 678-700.

[5] Dikamba T. Downlink scheduling in 3GPP long term evolution (LTE). MSc, Delft University of Technology, Delft, the Netherlands, 2011.

[6] Trivedi RD, Patel MC. Comparison of different scheduling algorithm for LTE. International Journal of Emerging Technology and Advanced Engineering 2014; 4 (5): 334-339.

[7] Lima FR, Rodrigues EB, Maciel TF, Nordberg M. Resource Allocation for Improved User Satisfaction with Applications to LTE. Resource Allocation and MIMO for 4G and Beyond. New York, NY, USA: Springer, 2014.

[8] Zain ASM, Malek MFA, Elshaikh M, Omar N, Hussain AST. Performance analysis of scheduling policies for VoIP traffic in LTE-advanced network. In: 2015 International Conference on Computer, Communications, and Control Technology; 21–23 April 2015; Kuching, Malaysia. New York, NY, USA: IEEE. pp. 16-20.

[9] Samia D, Ridha B. A new scheduling algorithm for real-time communication in LTE networks. In: 2015 IEEE 29th International Conference on Advanced Information Networking and Applications Workshops; 24–27 March 2015; Gwangiu, South Korea. New York, NY, USA: IEEE. pp. 267-271.

[10] Shams AB, Abied SR, Hossain MF. Performance comparison of network layouts with mobile users under different resource scheduling techniques in downlink LTE. In: 2016 5th International Conference on Informatics, Electronics and Vision; 13–14 May 2016; Dhaka, Bangladesh. New York, NY, USA: IEEE. pp. 949-954.

[11] Dardouri S, Bouallegue R. Comparative study of downlink packet scheduling for LTE networks. Wireless Personal Communications 2015; 82: 1405-1418.

[12] Chayon HR, Dimyati KB, Ramiah H, Reza AW. Enhanced quality of service of cell-edge user by extending modified largest weighted delay first algorithm in LTE networks. Symmetry 2017; 9: 81. doi: 10.3390/sym9060081

[13] Dardouri S, Bouallegue R. Comparative study of scheduling algorithms for LTE networks. World Academy of Science, Engineering, and Technology International Journal of Computer, Information Science, and Engineering 2014; 8: 436-441.

[14] Jain R, Chiu DM, Hawe WR. A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer System. Hudson, MA, USA: Eastern Research Laboratory, Digital Equipment Corporation, 1984.

[15] T-Mobile. LTE Physical Layer Framework for Performance Verification. R1-070674, 3GPP TSG RAN, 48. St Louis, MO, USA: T-Mobile, 2007.

[16] Rupp M, Schwarz S, Taranetz M. The Vienna LTE-Advanced Simulators. Singapore: Springer, 2016.