# A ML based downlink scheduling for 5G Networks

Fang-Yie Leu[1*], Kun-Lin Tsai[2], Heru Susanto[3,4], Jung-Chun Liu[1]

[1] Department of Computer Science, Thunghai University, Taichung, Taiwan
leufy@thu.edu.tw
[2] Department of Electrical Engineering, Thunghai University, Taichung, Taiwan
kltsai@thu.edu.tw
[3] School of Business, University Technology of Brunei, Bandar Seri Begawan BE, Brunei Darussalam
uus.khusni@lipi.go.id
[4] Research Center for Informatics, The Indonesia Institute of Sciences, Cibinong, Indonesia
jcliu@thu.edu.tw

**Abstract**

Presently, owing to the deployment of 5G networks, people wish that high quality streaming can be available in the near future. However, data transmission between BS and UE is one of the biggest challenges for high-quality streaming. The key reason is that a BS is often given limited bandwidth. To mitigate the problem, in this study, we propose a downlink scheduling mechanism, named the Machine-learning based Scheduling Scheme (QSRAS) which dynamically adjusts QoS parameters for a base station. Our simulation demonstrates that the QSRAS outperforms those of state-of-the-art systems on throughputs and delays.

**Keywords:** ML, Downlink scheduling, 5G networks, QoS parameters

# 1 Introduction

Currently, online streams and multimedia are popularly used everywhere in the world. Also, a large number of IoT systems, such as smart cities, industry automation and intelligent vehicles, will be established and connected to 5G networks. In order to offer users with greater bandwidth and tolerable delays, LTE employs Scheduling and Resource Allocation (SRA) to manage wireless resources for UEs so as to make the limited radio resources of a BS be effectively utilized.

The 7th International Conference on Mobile Internet Security (MobiSec'23), Dec. 19-21, 2023, Okinawa, Japan, Article No. 53

*Corresponding author: Department of Computer Science, Thunghai University, Taichung, 407224, Taiwan, Tel: +886-4-2359-0121

In this study, w we propose a downlink scheduling mechanism, named the Q-learning based Scheduling Resource Allocation System (QSRAS) which dynamically adjusts QoS parameters for a base station. Our simulation demonstrates that the QSRAS outperforms those of state-of-the-art systems on throughputs and delays.

This paper is organized as follows. Section 2 briefly describes related studies of this paper. Section 3 presents the architecture of the QSRAS. Our simulation and the results are stated and discussed in Section 4, respectively. Section 5 concludes this paper and addresses out future studies.

# 2 Relate Work

In the following, we will briefly describe the related studies of this paper.

## 2.1 Scheduling and Resource Allocation (SRA)

Basically, when UE is connected to a BS, Channel State Information (CSI) will be exchanged between these two network components every Transmission Time Interval (TTI). As receiving CQI from UE, BS chooses an appropriate modulation and coding scheme (MCS) for this UE following Adaptive Modulation and Coding (AMC) defined in [1]. After that, the SRA allocates suitable resources of BS to serve the UE, aiming to achieve better system performance.

## 2.2 Packet Scheduling Algorithms

According to [2], radio resource management and scheduling solutions can be divided into two types: QoS unaware and QoS aware.

### 2.2.1 QoS Unaware

QoS unaware deals with parameters related to system fairness, including CSI, average transmission rate, etc., aiming in effectively allocating radio resources and scheduling UE transmission.

Max CQI [3] allocates radio resources to UEs according to the conditions of the channel allocated to UEs at current TTI. In fact, Max CQI can effectively maximize the overall throughputs of data transmission without guaranteeing resource allocation fairness among UEs. With RR, the circulated schedule offers the best user fairness. But its throughputs are not maximum due to lacking channel quality. PF maximizes throughputs and achieves fairness by estimating current channel quality and transmission rates before it assigns RBs to UEs.

### 2.2.2 QoS-Aware

This scheduling category takes different parameters into account for scheduling decisions. The following will be their explanation.

#### (1) Delay Aware

Delay aware scheme is more suitable for real-time transmission (RT) flows, such as, self-driving car and remote surgery. QoS class identifier (QCI) [4] can be applied to different types of RT flows, including delivery delay, packet loss, flow priority, etc. Typical delay aware approaches, like Modified earliest deadline first (M-EDF) [5], Modified Largest Weighted Delay First (M-LWDF) [6], Exponential Proportional Fair (EXP-PF) [6] and Exponential Rule (EXP-Rule) [7], can be found in literature. However, these approaches do not balance QoS services when flows are hybrid, i.e., mixing RT and NRT (non-real-time transmission). Generally, RT flows which are given higher priority should be delivered first.

**A. EXP-PF**

This scheduler calculates weight $\mathrm{W}_{i,j}(t)$ for flows, aiming to enhance the priorities of RT flows by adopting exponential function, $D_{HoL}$ and PF rule and $\mathrm{h}(t)$, defined in [2], which is the delay of RT flows, as well as $\mathrm{N}_i$ which is the number of RT flows currently in the system. The EXP-PF truly lowers delay time for RT flows.

**B. M-LWDF**

This scheme deals with PF and Head of Line (HoL) Delay when delivering RT flows, while utilizing PF to process NRT flows, meaning that M-LWDF takes care of RT and NRT flows for scheduling at the same time. Its weight parameter $\mathrm{W}_{i,j}(t)$ can be found in [2], where $D_{HoL,i}(t)$ is HoL delay, and $\alpha_i$ decides the degree of delay. $\delta_i$ is is the maximum probability that $D_{HoL,i}(t)$ may exceed the delay threshold $\tau_i$ defined beforehand. If a RT-flow packet's delay time in the MAC queue exceeds $\tau_i$, the M-LWDF drops the packet to avoid prolonging the flow's delays.

### (2) Queue Aware

Generally, Queue Aware schemes emphasizes fairness for overall UEs presently in a system, particularly focusing on the length of a MAC queue, attempting to highlight throughputs of RT flows, while guaranteeing the lowest transmission rate for NRT flows. Virtual Token Modified Largest Weighted Delay First (VT-M-LWDF) is a typical Queue aware scheme by improving M-LWDF. Its $\mathrm{W}_{i,j}(t)$ (Nasralla, 2020) is calculated by substituting the delay of RT Flows in M-LWDF by queue length. This parameter shows the number of packets right now in a queue. Actually, the VT-M-LWDF is more suitable for multimedia services, like VoIP and jitter sensitive applications.

### A. Target bitrate

The Target bitrate is developed for maximizing throughputs of all UEs in a system, like those in [8][9]. It minimizes target transmission rates for RT flows, while maximizing throughputs for NRT flows. However, when serving RT and NRT flows at the same time, their rules are hard to guarantee the QoS requested by users.

### B. Hybrid and others

Many hybrid algorithms have been relaesed. Nasvalla [2] presented a hybrid scheduling approach to improve M-LWDF by applying SRA algorithm to RT and NRT flows simultaneously for lowering NRT delays. It adopts a cross-layer high-efficient algorithm introduced in [10] which achieves the best performance of the SRA with a dynamic programming method. Actually, it is an enhanced version of a greedy algorithm described in [11]. As we know, a greedy method may be stuck on a local minimum, rather than a global one.

Feki and Zarai [12] presented a Q-learning SRA for determining which scheduling algorithm is suitable for current TTI. Based on system fairness indexes, throughputs, and other thresholds, PF is chosen when the underlying system is now balanced between throughputs and fairness. Max CQI will be used when system throughputs have not met UE's QoS requirements. When fairness does not achieve UE's requests, RR will be chosen.

## 2.3 The Q-Learning

Reinforcement learning, as a machine learning technique, focuses on the interaction between itself and current working environment. The purpose is maximizing the rewards for the following processes. The key concept of Q-learning is to enhance good behaviors, and weaken poor ones. The $Q$ function defined in Eqs. (1) and (2) [12] for a given $Q(\text{state}, \text{action})$ is shown below.

$$Q(s_t, a_t) \longleftarrow Q(s_t, a_t) + \alpha \times \delta_t \quad (1)$$

$$\delta_t = r_{t+1} + \gamma \max_a Q(s_{t+1}, a_t) - Q(s_t, a_t) \quad (2)$$

in which $\delta_t$ is a temporal error; $Q(s_t, a_t)$ as a function evaluates current actions; $\alpha$ is the learning rate, $0 \leq \alpha \leq 1$, for ranking the certitude of values estimated previously; $r_{t+1}$ is an reward received from the environment and $\gamma$represents the weight that influences future rewards by immediate rewards, $0 \leq \gamma \leq 1$. It is a discount factor showing the importance of future rewards. If $\gamma = 0$, this learning approach only accepts immediate reward.

# 3 The QSRAS Architecture

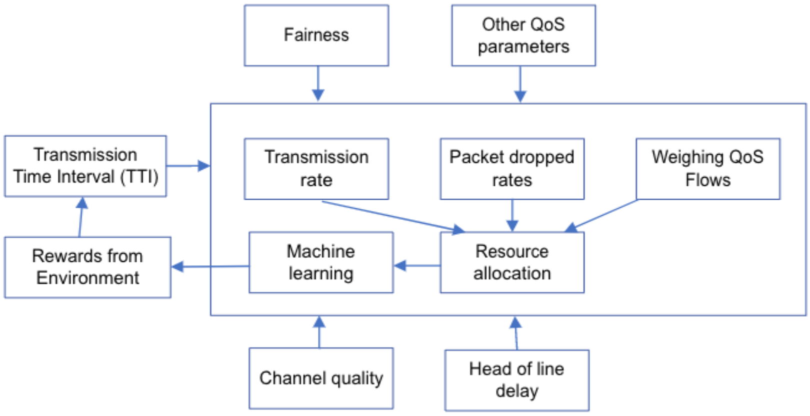


**Fig. 1.** The architecture of the QSRAS.

The architecture of the QSRAS is shown in Fig. 1. It balances system throughputs and delays for UEs. The approach used is updating Q-Table with Q-learning technique to guarantee UEs' QoS requirements, while continuing serving NRT applications. The QSRAS adopts channel quality and HoL delays as its key QoS parameters. In the QSRAS, UE's weight of $\text{flow}_i$ at $\text{subcarrier}_j$ at time $t$, denoted by $\mathrm{W}_{i,j}(t)$, is calculated by using Eq. (3). In each TTI, the higher the weight, the higher the UE' s priority.

$$\mathrm{W}_{i,j}(t) = \mathrm{W}_{i,j}(t-1) + \alpha\left(\mathrm{R}_{i,j}(t) - \mathrm{W}_{i,j}(t-1)\right) \tag{3}$$

where $\mathrm{W}_{i,j}(t-1)$ represents the weight of $\text{flow}_i$ at $\text{subcarrier}_j$ at time $t-1$; $\mathrm{R}_{i,j}(t)$ is the immediate reward responded from the underlying environment for $\text{flow}_i$ at $\text{subcarrier}_j$ at time $t$; as described previously, $\alpha$ is the learning rate, $0 \le \alpha \le 1$. When the value of $\alpha$ is larger, the impact on $Q$ by reward $r$ is also higher. To balance QoS required by RT flows, $\alpha > 0.5$ to improve the impact on the next TTI by current reward, consequently producing greater impacts on the weight of RT flows at time *t*.

$\gamma$ in Eq. (2) indicates the weight influencing future rewards by current immediate rewards. But, we only consider immediate reward. In other words, this reward is determined only by current environment, i.e., $\gamma = 0$. The reward $\mathrm{R}_{i,j}(t)$ received from current environment is shown in Eq. (4).

$$\mathrm{R}_{i,j}(t) = \frac{r_{i,j}(t)}{\overline{R}_i(t)} \times \mathrm{D}_{\mathrm{HoL},i}(t) \tag{4}$$

in which $r_{i,j}(t)$ defined above is the instantaneous transmission rate of $\text{flow}_i$ at $\text{subcarrier}_j$ at time *t*, $\overline{R}_i(t)$ is defined in [2], $\mathrm{D}_{\mathrm{HoL}_i}(t)$ is $\text{flow}_i$'s HoL delay at time *t*. When allocating an RB to UE, e.g., $\text{UE}_l$, SRA retrieves the maximum weight, e.g., $\mathrm{W}_{l,k}(t)$, $1 \le k \le m$, in Matrix table $(n \times m)$ and then allocates $\text{RB}_k$ to $\text{UE}_l$. In other words, $r_{i,j}(t)$ is utilized to enhance system throughputs, and $\mathrm{D}_{\mathrm{HoL},i}(t)$ improves the weight for delay in Matrix table.

Since this learning technique takes previous weight as a part of its parameters, its diversification of its weights is smoother than those of the other algorithms.

# 4 Experiment and Analysis

In the following, 5G-air-simulator, an open-source simulator [13], will be used to perform our experiments. The metrics employed includes packet loss rate (PLR), delays, throughputs and spectrum efficiency. Assuming that the number of packets sent by sender (received by the receiver) is K (Q). PLR is defined as (K-Q)/K. Throughput is defined as the number of bits the receiver receives per second. Delay is the time with which a packet travels from the sender to the receiver. Spectrum efficiency is defined as the number of bits that a spectral frequency in Hz can carry. We will also evaluate the performance of the QSRAS and some state-of-the-art SRA algorithms in 5G.

## 4.1 Simulation

The parameter settings of the 5G-air-simulator are listed in Table 1. In a cellular system, seven cells are joined as a hexagon, and a number of UEs are connected to a BS. UEs move toward the directions randomly selected with the moving velocity of 30 km per hour. An UE sends video packets toward their destinations simultaneously. The simulation last 30 s, and a flow is transmitted 25 s.

**Table 1.** Simulation Parameter settings

| Parameters | Values |
|---|---|
| Base station Bandwidth | 20 MHz |
| Number of RBs | 100 RBs |
| Number of UEs | 10~60 |
| Max Delay | 100 ms |
| Frame Structure | FDD |
| Cell radius | 1000 m |
| Video bitrate | 440 kbps |
| UE’s moving speed | 3, 30, 75, 120, 240 and 360 km/h |

## 4.2 Evaluation

In the following experiment, we evaluate performance of some state-of-the-art SRA algorithms, including PF, M-LWDF, EXP-PF, RR and QSRAS, given different moving speeds of UE. The evaluation metrics as mentioned above include PLRs, delays, system throughputs and spectral efficiencies. Each experiment is performed 100 times.

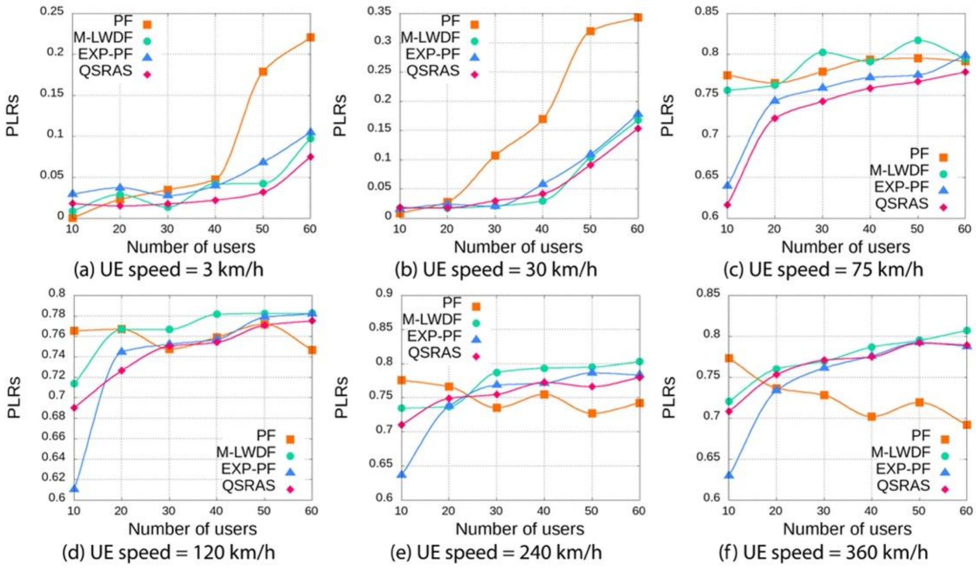


**Fig. 2.** Packet loss rates of the 4 test schemes.

After testing, the PLRs of the 4 schemes are shown in Fig. 2. When the amount of users is high, the video quality decreases since PLRs are relatively high. Since they compete the RBs of the base station they are connected to. With lower moving speeds, the QSRAS performs better, even not completely outperforming other evaluated systems. PF's PLRs are high when speeds are low. But on 360km/h, the values are lower than those of other schemes since it does not deal with QoS parameters, thus able to transmit data directly. Other three schemes need to check QoS parameters before allocating RBs to UEs and choose corresponding signal modulation approaches, e.g., 256 QAM or QPSK. Further, when UE's moving speeds are higher than 75 km/h, PLRs are heavy.

Delays of the 4 test schemes are illustrated in Fig. 3, in which the schemes without QoS receive longer delays, meaning that if SRA algorithms would like to reduce packet transmission delays and enhance user experience, the parameter of delays need to be one of the QoS parameters. The PF's delays are long since no QoS parameters are involved.

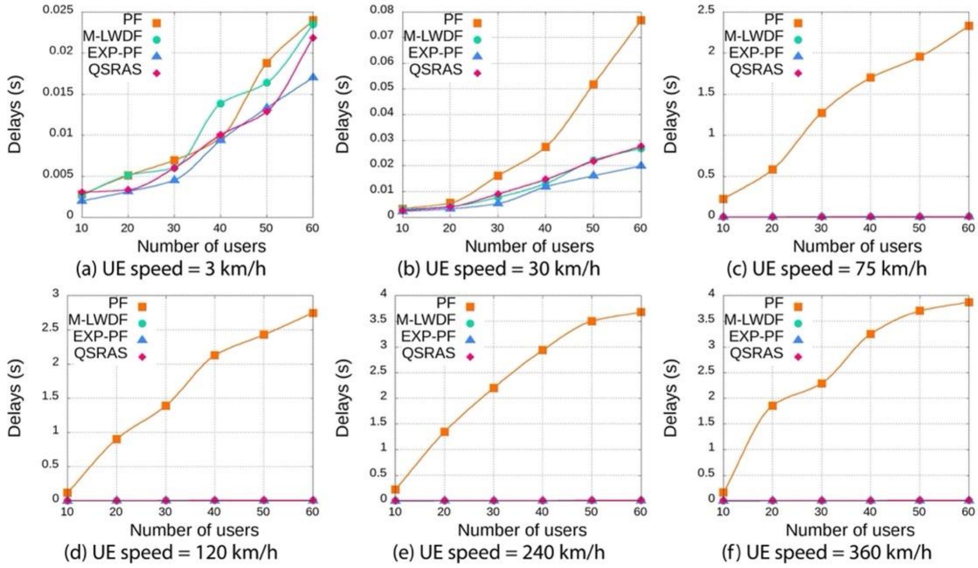


**Fig. 3.** Delays of the 4 test schemes.

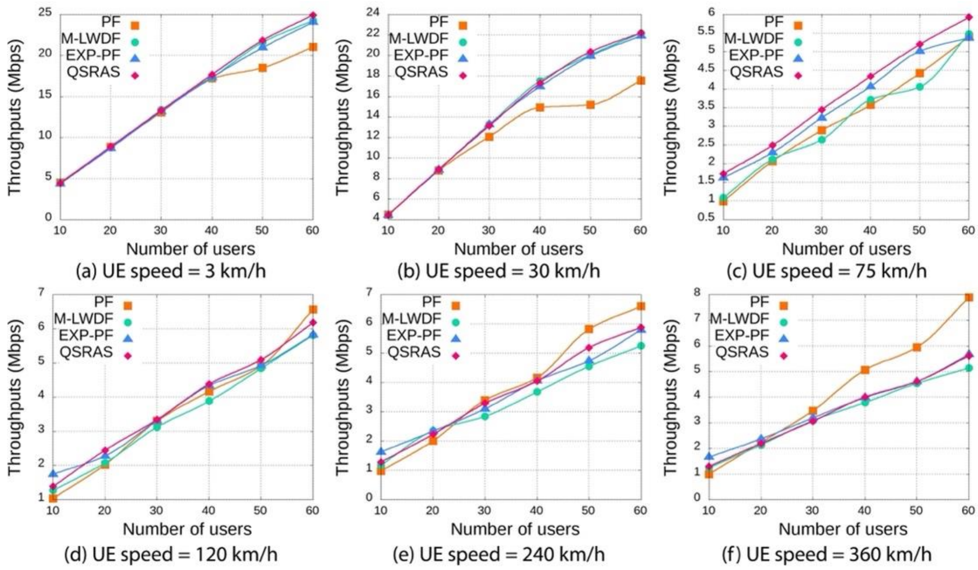


**Fig. 4.** Throughputs of the 4 test schemes.

Throughputs are shown in Fig. 4. As number of users rises, throughputs follow for all test schemes. When UEs' moving speeds are low and traffic is heavy, i.e., involving many more users, the PF's throughputs are lowered. As moving speeds are up to 360 km/h, the PF outperforms the other three due to its low PLRs (see Fig. 2).

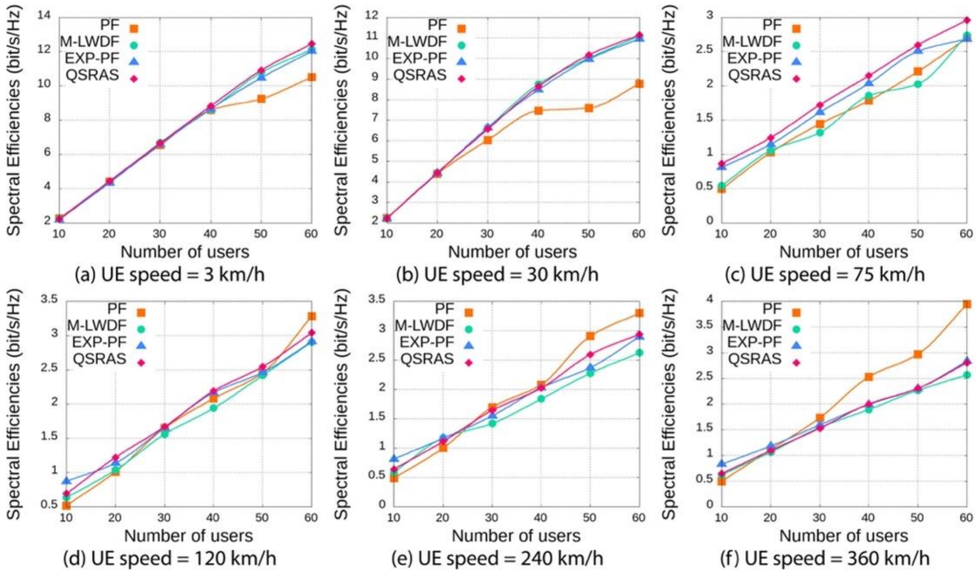


**Fig. 5.** System spectral efficiencies.

The system spectral efficiencies of the 4 test schemes for a network are depiced in Fig. 5. Spectral efficiency actually is one of the performance indicators for base station resource utilization. As number of users is higher, spectral efficiencies increase for all test schemes. When the number of UEs is 60 and UEs' moving speeds are low, the spectral efficiencies can achieve 11-12 bit/Hz. However, at high speeds, the values are reduced to 2-4 bits/Hz., meaning UE's moving speeds are an important factor for spectral efficiencies.

# 5 Conclusions and Future Studies

This study presents a Q-learning_based downlink SRA system, attempting to improve the performance for traffic of different categories and enhance its system capacities. In [10], the scheduling/resource allocation approaches with various QoS parameters are discussed, including delay aware, queue aware, target bit-rate aware, Hybrid aware and others. The QSRAS deals with channel quality, previous delays and throughputs and computes the weight with Q-learning approach for RB scheduling and allocation. Our simulation evaluates different downlink scheduling/resource allocation schemes. The evaluation metrics include PLRs, transmission delays, system throughputs and spectral efficiencies.

Our simulation shows that the schemes of delay awareness demonstrate well. Also, the QSRAS balances QoS and traffic of different categories. Now, we dare to conclude that QoS balanced scheduling/resource allocation schemes are better in delivering RT flows. In the future, we would like to optimize and extend this study and derive the behaviors and reliability models of the QSRAS so that users can comprehend the behaviors and reliability before using it. These constitute our future studies.

**Acknowledgments.** This study is financial support in part by Ministry if Technology and Science/ National Science and Technology Council, Taiwan under the grants NSTC 112-2221-E-029-007 and MOST 109-2221-E-029-017-MY2.

# References

[1] 3GPP 38 Series TS 138 214V16.5.0 5G; NR; Physical layer procedures for data.
[2] Nasralla, M.M. (2020). A Hybrid Downlink Scheduling Approach for Multi-Traffic Classes in LTE Wireless Systems. Retrieved from IEEE Access, vol. 8 (pp. 82173-82186). IEEE.
[3] Gueguen, C. and Baey, S. (2009). A fair MaxSNR scheduling scheme for multiuser OFDM wireless systems. In proceedings of the IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, Tokyo, Japan (pp. 2935-2939). IEEE.
[4] 3GPP 23 Series TS 123 501V16.8.0 5G; System architecture for the 5G System.
[5] Hamed, M.M., Shukry, S., El-Mahallawy, M.S. and El-Ramly, S.H. (2014). Modified earliest deadline first scheduling with channel quality indicator for downlink real-time traffic in LTE networks. In proceedings of the third International Conference on e-Technologies and Networks for Development (ICeND2014), Beirut, Lebanon (pp. 8-12).
[6] Basukala, R., Ramli, H.A.M. and Sandrasegaran, K. (2009). Performance analysis of EXP/PF and M-LWDF in downlink 3GPP LTE system. The First Asian Himalayas International Conference on Internet, Kathmundu, Nepal(pp. 1-5).
[7] Ang, E., Wee, K., Pang, Y. et al. (2015). A performance analysis on packet scheduling schemes based on an exponential rule for real-time traffic in LTE. Journal of Wireless Communication Network, 201.
[8] Monghal, G., Pedersen, K.I., Kovacs, I.Z. and Mogensen, P.E. (2008, May). QoS oriented time and frequency domain packet schedulers for the UTRAN long term evolution. In Proceedings of VTC Spring-IEEE Vehicle Technology Conference, Singapore (pp. 2532–2536). IEEE.
[9] Skoutas, D.N. and Rouskas, A.N. (2010, April). Scheduling with QoS provisioning in mobile broadband wireless systems. In Proceedings of European Wireless Conference (EW), Lucca, Italy, (pp. 422–428).
[10] Vora, A. and Kang, K.D. (2018). Effective 5G Wireless Downlink Scheduling and Resource Allocation in Cyber-Physical Systems. Technologies 6, no. 4: 105.
[11] Femenias, G., Riera-Palou, F., Mestre, X. and Olmos, J.J. (2017). Downlink Scheduling and Resource Allocation for 5G MIMO-Multicarrier: OFDM vs FBMC/OQAM. IEEE Access, vol. 5, (pp. 13770-13786). IEEE.
[12] Feki, S. and Zarai, F. (2017). Cell Performance-Optimization Scheduling Algorithm Using Reinforcement Learning for LTE-Advanced Network. In proceedings of the IEEE/ACS 14th International Conference on Computer Systems and Applications (pp. 1075-1081). IEEE.
[13] Martiradonna, S., Grassi, A., Piro, G. and Boggia, G. (2020). 5G-air-simulator: an open-source tool modeling the 5G air interface. Computer Networks (Elsevier), vol. 173 (107151).