2022 International Conference on Computer and Drone Applications (IConDA) | 978-1-6654-9235-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICONDA56696.2022.10000358

Dynamic Resource Block Allocation Techniques for Simultaneous EMBB and URLLC Traffic

Thaahirah Shireen Mohamed Rasied
Department of Electrical and Computer Engineering
International Islamic University Malaysia
Kuala Lumpur, Malaysia
hadibahmr@iium.edu.my

Huda Adibah Mohd Ramli
Department of Electrical and Computer Engineering
International Islamic University Malaysia
Kuala Lumpur, Malaysia

Farah Diyana Abdul Rahman
Department of Electrical and Computer Engineering
International Islamic University Malaysia
Kuala Lumpur, Malaysia

***Abstract*—Efficient utilization of Resource Blocks (RBs) whilst simultaneously meeting the desired Quality of Service (QoS) of Ultra-Reliable Low Latency Communication (URLLC) and Enhanced Mobile Broadband (EMBB) applications in the downlink Fifth Generation (5G) is an open research problem. Therefore, this paper proposes two dynamic Resource Block (RB) allocation techniques known as Channel-aware Dynamic RB Allocation (CDRA) and Urgency-aware Dynamic RB Allocation (UDRA) in our attempts to address the problem. These techniques dynamically allocate the number of RBs for packet transmission according to a user channel quality (CDRA) as well as packets urgency (UDRA) to achieve the 1 ms latency with 10^{-5} Packet Loss Ratio (PLR). An extensive computer simulation was conducted based on realistic radio propagation and interference affected in a cell using the new frame structure, mini-slot 0.143 ms Transmission Time Interval (TTI) of 15 kHz Subcarrier Spacing (SCS), dynamic Control Channel (CCH) and short Hybrid Automatic Repeat Request (HARQ) features. The simulation results showed the efficacy of UDRA over CDRA where UDRA has explicitly improved the multimedia QoS namely PLR, latency, and throughput in the downlink 5G network.**

***Keywords*—latency, 5G, resource allocation, URLLC, EMBB**

I. INTRODUCTION

Research towards achieving the goals of Fifth Generation (5G) and beyond to support diverse multimedia use cases especially Enhanced Mobile Broadband (EMBB) and Ultra Reliable Low Latency Communication (URLLC) applications, has gained much attention in the literature [1]. EMBB is the expansion of current Fourth Generation (4G) broadband promising a higher transmission rate and supporting more multimedia applications. The use case is expected to support exceptionally high transmission rates (i.e., 20 Gbps for data-hungry multimedia applications such as 3D videos, Ultra-High Definitions, U-HD, screens, Augmented Reality, AR). URLLC is a high prospectus use case that requires support for unprecedented latency and reliability [2]–[4]. It focuses on time-sensitive applications which include mission-critical applications, self-driving cars, industrial automation, and control equipment or devices for smart factories.

The Key Performance Indicator (KPI) for URLLC where URLLC multimedia packets must be transferred and correctly arrived at the user end within 1 ms latency and 10^{-5} Packet Loss Rate (PLR) was defined in Release 15. Concurrent support of both EMBB and URLLC in one shared network will be a challenging issue to be addressed, as two conflicting factors, namely, latency and reliability must be satisfied, and additionally, throughput must be maximized.

Packet scheduling is of the most vital 5G Radio Resource Management (RRM) functions. It is accountable for intelligently selecting users for transmissions of their multimedia packets in each Transmission Time Interval (TTI) [5] such that scarce 5G radio resources are effectively used whilst the desired Quality of Services (QoS) of every user is satisfied. Numerous packet scheduling algorithms have been developed for the effective transmission of multimedia traffic in the legacy mobile wireless network. These include well-known algorithms namely Maximum Rate (Max-Rate) [6], Round Robin (RR) [7], Maximum-Largest Weighted Delay First (M-LWDF) [8], Proportional Fair [9], Exponential Proportional Fairness (EXP/PF) [10] and other QoS aware and delay based PS algorithms [11], [12]. These well-known algorithms presented excellent performances in managing resources under the legacy mobile cellular network requirements and were later adapted to simultaneously meet the desired multimedia QoS in the 4G mobile cellular networks.

Contrary to the downlink legacy 4G that operates packet scheduling in every 1 ms TTI, packet scheduling is performed in a variety of TTI in the downlink 5G. Downlink 5G was chosen as it remains the dominant demand for data transmission for various multimedia applications. Moreover, conflicting to the 4G that executes packet scheduling on each Resource Block (RB), the packets of a chosen 5G user will be transmitted on a group of RBs which is fixed depending on the chosen TTI such that the group of RBs use will be narrower or wider to meet the stringent demand of the diverse 5G multimedia use cases. These changes were made to allow efficient usage of the scarce 5G bandwidth. Given the changes made, this paper seeks to explore dynamic techniques of RBs allocation, rather than fixed RB allocation employed in the current technique, such that the expensive bandwidth can be further utilized. The fixed RB allocation allocates a fixed number of RBs for packet transmission and the number of RBs to be allocated is dependent upon the chosen TTI.

The remaining sections of the paper are organized as follows: Section II elaborates on the downlink 5G network while section III details the proposed RB allocation techniques. Section IV elucidates the assumptions made for the simulation. Section V discusses the results of the proposed RB allocation techniques followed by a conclusion that summarizes the paper in Section VI.

II. DOWNLINK 5G NETWORK MODEL

As satisfying URLLC requirements shall be more critical, achieving low latency comes with a trade-off of throughput degradation. To reduce latency, the 3GPP has proposed a

978-1-6654-9235-5/22/$31.00 ©2022 IEEE

flexible numerology frame structure which is contrary to the legacy 4G fixed frame structure [13]. Flexible numerology is discussed in sub-section II-B. Like the legacy 4G, the 5G utilizes the Orthogonal Frequency Division Multiple Access (OFDMA) as its multiple access scheme to allow users to dynamically multiplex in the time-frequency domains of the frame structure. The OFDMA divides the bandwidth of each radio spectrum into sub-carriers of one specific frequency spacing. In the time domain, Cyclic Prefix (CP) that functions as a guard interval is inserted after each OFDMA symbol. This allows the OFDMA signals to be resistant to Inter Symbol Interference (ISI) due to multi-path propagation that degrades signal quality at the receiver [14]. The smallest radio resource unit that is used for data/control/signaling transmission in the downlink 5G is called RB.

A. Legacy 4G Frame Structure

An RB in the legacy 4G has a fixed 1 ms duration [5] and consists of 14 OFDM symbols for normal CP in the time domain. The frequency domain of an RB contains 12 sub-carriers with 15 kHz SubCarrier Spacing (SCS), each adding up to a 180 kHz width. A total of 168 Resource Elements (REs) (14*12=168 REs) are available in each RB when using normal CP as shown in Fig. 1. Most of the REs are assigned to carry downlink data whereas the rest known as Control Channel Overhead (CCH) are used for control and signaling functions.

B. 5G Frame Structure

In contrast to the legacy 4G, the 5G supports scalable numerology which consists of flexible SCS, as shown in Table I, as well as a flexible frame structure. The flexible frame structure allows time-frequency multiplexing of multimedia users while allowing them to adaptively transmit according to their channel conditions [15]. The scalable frame structure introduced in the downlink 5G is by the factor: 15 x 2n kHz, where n is an integer $(0\leq n \leq 2)$. The 15 kHz width is the SCS used in legacy 4G with a TTI of 1 ms (see Table I). The numerology accommodates 14 OFDM symbols per TTI slot. The maximum TTI size for scheduling for a specific SCS equals the time duration for 14 OFDM symbols while the minimum TTI size can be reduced according to certain specifications and details can be found in [13]. This numerology can be reduced to mini-slots TTI of 0.5 ms 0.286 ms and 0.143 where 7,4 and 2 OFDM symbols can be transmitted per TTI respectively, in one RB. Short TTI is significantly effective with the penalty of higher channel control overhead (CCH). Higher numerologies could be adapted to reduce the CCH, yet, this can only be adapted according to the deployment scenario [13].

TABLE I. LIST OF SCS WITH ITS TTI AND MAXIMUM BANDWIDTH FOR 5G RADIO SPECTRUM [16]

Frequency	***SCS***	***Slot durations (TTI)***	***Max. bandwidth***
0.45 GHz-6 GHz	15 kHz	1 ms	50 MHz
	30 kHz	0.5 ms	100 MHz
	60 kHz	0.25 ms	200 MHz
24 GHz-52.6 GHz	60 kHz	0.25 ms	200 MHz
	120 kHz	0.125 ms	400 MHz

The numerology of 15 kHz SCS with 0.143 ms TTI which is similar to the work in [17], [18] was chosen in this paper to better support delay-sensitive URLLC use case. This is because the use case demands ultra-low latency such as 1 ms and the 15 kHz SCS with 0.143 ms numerology seems promising to achieve this strict uRLCC latency. This frame structure accommodates 2 OFDM symbols which is equivalent to 0.143 ms in a subframe which is smaller than 14 OFDM symbols which is equivalent to 1 ms per subframe in legacy 4G. Since this numerology contains 2 OFDM symbols, each time slot is made of 1 ms/ 14 *2=0.143 ms. It is the time ratio of legacy 4G which contains14 OFDM symbols in 1 ms time frame. Therefore, the number of REs in one RB for 2 OFDM symbols of 15 kHz SCS is 2 x 12=24 REs. An illustration of a number of REs for one RB, in legacy 4G and the chosen 5G numerology 15 kHz, 0.143 ms mini-slot TTI were given in Fig. 1.

C. Short Round Trip Time (RTT) of Hybrid Automatic Repeat Request (HARQ)

The HARQ method requires users to send Acknowledgement / Negative Acknowledgement (ACK/NACK) feedback corresponding to the received packets to the base station. The 5G base station is known as gNB. ACK indicates the transmitted packets are received correctly. In case of decoding failure of the received packets, NACK is sent to the gNB to allow retransmission of the packets. It is assumed in this work that the downlink 5G employs Incremental Redundancy of HARQ (HARQ-IR). This HARQ-IR retransmits packets using a different set of coding bits based on recomputed Signal-to-Interference-Noise-Ratio (SINR) and the mapped Channel Quality Information (CQI) levels of currently available RBs. It is expected that HARQ-IR to improve system throughput in a limited feedback system, especially systems that require low

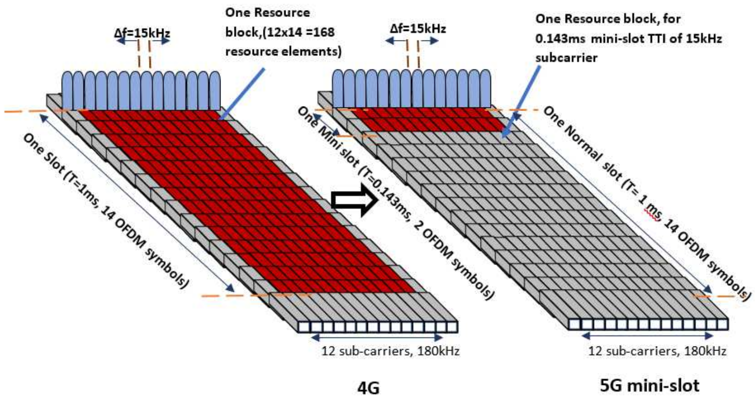

Fig. 1. (left) Frame structure fixed 4G LTE-A per RB (1ms TTI) and (right) frame structure of new 5G mini-slots (0.143 ms TTI) per RB. Both numerology using 15 kHz SCS

latency. It is to be noted, the TTI length for HARQ RTT specified in legacy 4G was 8 TTI, meaning that even using a 0.143 ms TTI, will not be possible to achieve the required latency for URLLC packets which is less than 1 ms. Thus, in line with [19], the RTT is reduced from 8 TTI to 4 TTI per downlink transmission.

III. DYNAMIC RB ALLOCATIONS

The legacy 4G allocates one RB which contains a fixed number of REs (168 REs) per RB. Out of these 168 REs, 14% or the equivalent of 24 REs are reserved for CCH [17]. The CCH contains information related to link adaption parameters needed for decoding [18]. This is not an issue in the legacy 4G as packet scheduling is performed on each RB of 168 REs per TTI. However, in our chosen 5G numerology, a flexible frame structure is utilized and the number of REs per RBs allocated for a user in a TTI depends on the chosen TTI. For example, if the numerology of 15 kHz SCS with 0.143 ms TTI is used, then only 24 REs per RB will be available. This will not be sufficient for the transmission of CCH. Therefore, in line with [20], two dynamic RB allocation techniques known as Channel-aware and Urgency-aware Dynamic RB allocations, are proposed in this paper (see Table II). These dynamic RB allocations are performed after user selection has taken place where they combine a certain number of RBs for the transmission of user packets. The group of RBs shall be allocated according to the UEs channel quality for sufficiency to accommodate a reasonable number of data packets.

A. *Channel-aware Dynamic RB Allocation (CDRA)*

The number of RBs to be allocated for transmission of packets is based on the estimated channel quality of a chosen user (either URLLC or EMBB) at each TTI. For a higher CQI (good channel quality), the number of RBs needed is less in comparison to the user located at the cell edge (lower CQI or low channel quality) where more RBs are needed.

B. *Urgency-aware Dynamic RB Allocation (UDRA)*

The UDRA technique is performed after the CDRA technique has taken place. This UDRA technique will only be performed if there URLLC users are having low channel quality and have waited in the buffer for more than 2TTIs. In this technique, the number of allocated RBs will be increased slightly to cater to the urgency of URLLC packets. UDRA aims to improve data transmission to meet the QoS of URLLC traffic as well as avoid unnecessary RB allocation ensuring optimization in spectrum efficiency in downlink 5G network. As an example, a chosen user was estimated with CQI 7. Therefore, based on Table II, 7 RBs will be used to transfer its data bits per Transport Block (TB). The data bits that could be accommodated in that TB is per chosen MCS based on its mapped CQI and number of REs for data payloads will be transmitted as per the equation below:

$$DB_{i,k}(t) = Eff_{i,k}(t) * \ RED_k\,(t) \tag{1}$$

where $DB_{i,k}(t)$ is the number of bits of *ith* user for *kth* data packet at *tth* TTI, $Eff_{i,k}(t)$ is the efficiency (in bits/RE) of *ith* user for *kth* data packet at *tth* TTI, and RED_k is the total number of REs specified for *kth* data packet.

Thus, 7 RBs will be allocated (7 x 24REs = 168 REs) for CQI 7 resulting in 96 REs for data payloads. The efficiency for the estimated CQI is 1.4766, leading to 141 bits of data payload. Note that, higher CQI users (due to stronger signals) require fewer RBs because the probability of successful data transmission is higher when using higher modulation schemes such as 16 Quadrature Amplitude Modulation (QAM) or 64 QAM) during transmission. This simulation assumed a small payload data transmission ensuring satisfactory QoS for both URLLC and EMBB traffic in a limited bandwidth network. If most of the users are located at cell boundaries (low CQI), the allocated bandwidth needs to be expanded to ensure successful data transmission (the lower modulation technique used guarantees a reduction in packet loss). Otherwise, URLLC QoS cannot be satisfied.

TABLE II. DYNAMIC RB ALLOCATION AND CCH FOR 15 KHZ SCS WITH 2 OFDM SYMBOLS (0.143 MS TTI)

CQI	***SINR (dB)***	***Control Channel overhead (CCH)***	***Resource Blocks Allocated (REs)***	
			CDRA	***UDRA***
1-3	(-∞, -6.936)	8x36=288 REs	14 RB (336)	20 RB (480)
4,5	(-6.936, -3.18)	4x36=144 REs	10 RB (240)	14 RB (336)
6,7	(-3.18, 4,694)	2x36=72 REs	7 RB (168)	7 RB (168)
8-10	(4,694, 10.366)	1x36=36 REs	4 RB (96)	4 RB (96)
11-14	(10.366, 19.829)	1x36=36 REs	3 RB (72)	3 RB (72)

IV. ENVIRONMENT OF SIMULATION

The impacts of CDRA and UDRA techniques were evaluated based on the PLR, throughput, and latency metrics in the modeled downlink 5G network for simultaneous support of the URLLC and EMBB use cases. These metrics are considered common when evaluating mobile wireless performance and are mathematically expressed as follows.

$$PLR = \frac{\sum_{i=1}^{i=K}\sum_{t=1}^{t=T} PD_i(t)}{\sum_{i=1}^{i=K}\sum_{t=1}^{t=T} PS_i(t)} \tag{2}$$

$$Throughput = \frac{1}{T}\sum_{i=1}^{i=K}\sum_{t=1}^{t=T} \mathrm{PR}_i(t) \tag{3}$$

$$L_{i,m} = \frac{\sum_{m=1}^{m=M} LXQ_{i,m} + PDel_{i,m}}{M} \tag{4}$$

$$LXQ_{i,m} = \begin{cases} 4 * TTI, without\ retransmission \\ 6 * TTI, with\_one\_retransmission \end{cases} \tag{5}$$

where $PD_i(t)$ is the size of discarded packets of *ith* user at *tth* time, $PS_i(t)$ is the size of all arriving packets at the gNB buffer of *ith* user at *tth* time, $PR_i(t)$ is the size of all correctly received packets of *ith* user at *tth* time, K is the total number of users, T is the total simulation time, $L_{i,m}$ is the latency of *ith* user of *mth* packet, $LXQ_{i,m}$ is the time duration for processing without packet queuing or delay of *ith* user of *mth* packet, $PDel_{i,m}$ is the queuing delay in gNB waiting for transmission of *ith* user of *mth* packet, M is the total number of packets and TTI is 0.143 ms. It is assumed that URLLC packets can be successfully received to a user either by first or one retransmission.

The simulation was conducted in a low load scenario with 1.26 Mbps with a fixed number of users (20 URLLC users:20 EMBB users) to accomplish the strict latency of 1 ms and PLR of 10^{-5} as per QoS requirements for URLLC. The impact of other user distributions as well as increasing the number of URLLC and EMBB users will be investigated in future works. The FTP model 3 and VoIP were used to represent the

URLLC and EMBB applications, respectively. This delay critical URLLC application requires URLLC packets to arrive at the user end within 10 ms duration (packet delay threshold) while the packet delay threshold of EMBB is more relaxed and hence was set to 100 ms. The assumptions and parameters used in this performance evaluation are briefed in Table III.

TABLE III. PARAMETERS OF DOWNLINK 5G SIMULATION

Simulation Parameters	**Values**
Topology and cell radius	1 Hexagonal Cell, 250m
Bandwidth and Number of CCs	10 MHz (each CC 5 MHz) and 2 CCs
Carrier Frequencies	2, 2.6 GHz,
RBs	50 RBs (12 sub-carriers per RB (180 kHz))
PHY Numerology	15 kHz sub-carrier spacing configurations
TTI size	0.143ms (2 OFDM symbols mini-slot)
eNB Transmit Power	43.01dBm
CQI (Channel Quality Reporting)	Periodic CQI every 5 ms, with 2 ms latency
HARQ	TYPE II HARQ with Incremental Redundancy (HARQ-IR)
BLER Target	10%
Traffic Model (packet arrivals)	URLLC: FTP3 type (every 16 ms) EMBB: Vo1P (every 20ms)
QoS (delay threshold & Max PLR)	URLCC: 10 ms & 10^{-5} EMBB: 100 ms & 10^{-2}
Link-to-system mapping	Based on effective exponential SINR mapping (EESM)
URLLC user and EMBB user proportion	40 users (20 URLLC: 20 EMBB)
UE Distribution	Uniformly distributed with 3km/hr mobility in Outdoor

For more detailed and comprehensive performance evaluations, this paper selected four well-known packet scheduling algorithms and implemented the dynamic RB allocation techniques for each. Given that this paper aims to simultaneously support the delay-sensitive URLLC and EMBB multimedia traffic, a weight criterion (*w*) was added to the priority calculation to allow these packet scheduling algorithms to prioritize URLLC users over EMBB users. The mathematical expressions of these selected algorithms that include weight insertion are given in Eq. 6 – Eq. 9 and these algorithms are represented as M-Max Rate, M-PF, M-MLWDF, and MEXP-M_LWDF hereafter.

$$\vartheta_{i,k}(t) = w_i * \ r\_avg_{i,k}(t) \tag{6}$$

$$\vartheta_{i,k}(t) = w_i * \frac{r\_avg_{i,k}(t)}{R_i(t)} \tag{7}$$

$$\vartheta_{i,k}(t) = w_i * \alpha_i * \ W_i(t) * \frac{r\_avg_{i,k}(t)}{R_i(t)} \tag{8}$$

$$\vartheta_{i,k}(t) = w_i * \alpha_i * \frac{r_{avg_{i,k}}(t)}{R_i(t)} * \left(\frac{T_i}{T_i - W_i(t)}\right) \tag{9}$$

where $\vartheta_{i,k}(t)$ is the significance metric of *ith* user on *kth* radio spectrum at *tth* TTI, w_i is the weight for *ith* user, $r\_avg_{i,k}(t)$ is the average transmission rate of *ith* user over all RBs on *kth* radio spectrum at *tth* TTI, $R_i(t)$ is the average rate of *ith* user at *tth* TTI (see [5]), $W_i(t)$ is the delay of the Head of Line (HOL) packet of *ith* user at *tth* TTI, and T_i is the application dependent buffer delay threshold of ith user i.

It is assumed that the weight criterion (w)for URLLC users is 5 and for EMBB users are 1. These values were chosen to ensure that in all situations especially during EMBB users having good channel quality while URLLC users are situated at the cell edge (poor reception), the considered packet scheduling algorithms will prioritize URLLC as compared to EMBB users. These weight values can be varied accordingly (i.e., if more than one type of URLLC application is evaluated).

V. RESULTS AND DISCUSSION

The results in the impacts of our proposed RB allocation techniques on PLR, latency, and throughput over well-known packet scheduling algorithms when simultaneously supporting URLLC and EMBB traffic in the downlink 5G are demonstrated in this section. Note that the discussed results are mainly on URLLC users due to the challenging QoS requirements in comparison to EMBB QoS, which are more relaxed. The EMBB performance was not shown in these results given that the desired QoS of all EMBB users was achieved. Fig. 2 illustrates the PLR performance of URLLC users, and it can be seen in the figure that the desired URLLC PLR threshold of 10^{-5} is only possible to be reached when the system implements the UDRA technique. Introducing weight criterion in packet scheduling algorithms allows the downlink system to prioritize URLLC traffics compared to EMBB. However, the occurrence of packet loss among users located at the cell edge is still high in CDRA due to insufficient of RBs for data transmission using a lower modulation scheme. Thus, UDRA which identifies URLLC users that need more RBs ensures that 10^{-5} PLR threshold could be met by all packet scheduling algorithms. Meanwhile, M-MaxRate demonstrates the worst PLR in both techniques due to the characteristic of the algorithm that considers channel quality only in priority selections in comparison to other packet scheduling algorithms that also account for packet delay and throughput.

The URLLC latency achieved by all evaluated packet scheduling algorithms is shown in Fig. 3. It can be noticed both techniques are capable to maintain URLLC latency below 1 ms. However, UDRA can significantly reduce the waiting time by 0.17 ms as compared to the CDRA. This achievement is reflected in all assessed packet scheduling algorithms. Consequently, it can be established that UDRA allows the packet scheduling algorithms to significantly improve the network performance whilst ensuring the URLLC QoS is maintained.

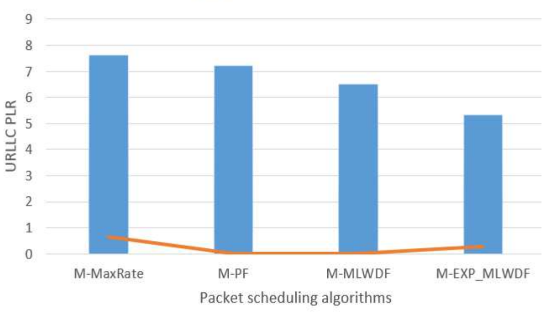

Fig. 2. PLR of URLLC user packets vs. packet scheduling algorithms

The throughput performance of every evaluated packet scheduling algorithm when implementing UDRA is presented in Fig. 4. The maximum throughput of 940.875 kbs was obtained in M-PF and M-MLWDF which maps to 0.07% improvements over CDRA. The impact of employing weight criterion allows the packet scheduling algorithms to prioritize URLLC users while short HARQ and dynamic RB allocation allow these evaluated packet scheduling algorithms to significantly improve the URLLC performance and satisfy the QoS requirements of all users simultaneously. Dynamic RB allocation with mini-slot comes with a tradeoff of an increase in signaling overheads during transmissions that leads to throughput degradation in the 5G as compared to the legacy 4G.

VI. CONCLUSION

Motivated by the challenge of meeting the rigid desired URLLC QoS of 1 ms latency with 10^{-5} PLR, this paper proposed two RB allocation techniques that can dynamically allocate several RBs to a chosen user according to the user channel quality and packet urgency. Simulation results demonstrated that the proposed UDRA technique is more effective than CDRA in achieving the desired URLLC QoS over all modified channel-aware packet scheduling algorithms, namely M-PF, M-MaxRate, M-MLWDF, and M-EXP-MLWDF. Future works include further investigations of these techniques when the number of users has varied as well as its performance comparison with the fixed RB allocation technique when simultaneously supporting the EMBB and URLLC multimedia traffic.

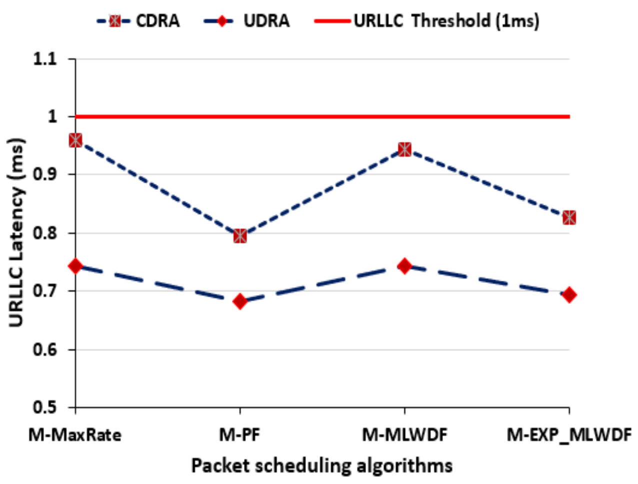

Fig. 3. Latency of URLLC user packets vs packet scheduling algorithms

ACKNOWLEDGMENT

This research was fully funded by the Kementerian Pendidikan Malaysia under Grant FRGS/1/2019/TK04/UIAM/02/2.

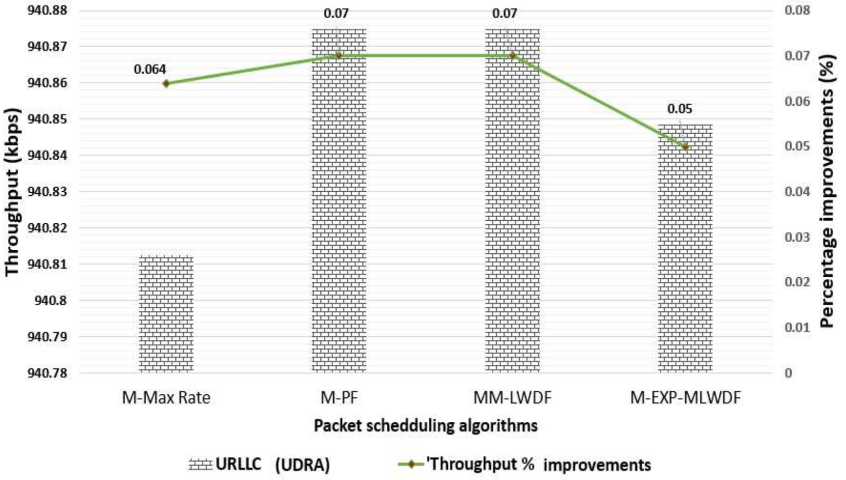

Fig. 4. Throughput of URLLC user packets and percentage of throughput improvements vs packet scheduling algorithms

REFERENCES

[1] X. Zhang, J. Wang, and H. V Poor, "Statistical delay and error-rate bounded QoS control for muRLLC over 6G multimedia mobile wireless networks in the non-asymptotic regime," in *56th Annual Conference on Information Sciences and Systems,* (CISS), pp. 206–211, 2022.

[2] A. A. Esswie and K. I. Pedersen, "Analysis of outage latency and throughput performance in industrial factory 5G TDD deployments," in *IEEE Vehicular Technology Conference*, (IEEE VTC), pp. 1–6, Dec. 2021.

[3] R. B. Abreu, G. Pocovi, T. H. Jacobsen, M. Centenaro, K. I. Pedersen, and T. E. Kolding, "Scheduling enhancements and performance evaluation of downlink 5G time-sensitive communications," *IEEE Access*, vol. 8, pp. 128106–128115, 2020.

[4] D. Panno and S. Riolo, "An enhanced joint scheduling scheme for GBR and non-GBR services in 5G RAN," *Wireless Networks*, vol. 26, no. 4, pp. 3033–3052, Jan. 2020.

[5] H. A. M. Ramli, "A study on packet scheduling algorithms for healthcare contents over fifth generation (5G) mobile cellular network," *International Journal of Electronics and Telecommunications*, (IJET) vol. 66, no. 4, pp. 729–735, 2020.

[6] H. A. M. Ramli, K. Sandrasegaran, R. Basukala, R. Patachaianand, and T. Sohana Afrin, "Video streaming performance under well-known packet scheduling algorithms," *International Journal of Wireless and Mobile Networks*, (IJWMN) vol. 3, no. 1, pp. 25–38, 2011.

[7] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey," *IEEE Communications Survey Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.

[8] E. M. A. Esheikh, M. H. Habaebi, H. A. M. Ramli, M. H. A. Arafa, and K. I. Ahmed, "Maximum-largest weighted delay first algorithm for heterogeneous traffic in 4G networks," *Indones. J. Electr. Eng. Informatics*, vol. 6, no. 2, pp. 190–199, Jun. 2018.

[9] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *IEEE Vehicular Technology Conference*, (IEEE VTC), vol. 3, pp. 1854–1858, 2000.

[10] J. H. Rhee, J. M. Holtzman, and D. K. Kim, "Scheduling of real/non-real time services: Adaptive EXP/PF algorithm," *IEEE Veh. Technol. Conf.*, vol. 57, no. 1, pp. 462–466, 2003.

[11] N. D. Adesh and A. Renuka, "Adaptive downlink packet scheduling in LTE networks based on queue monitoring," *Wireless Networks*, vol. 25, no. 6, pp. 3149–3166, 2019.

[12] I. Angri, M. Mahfoudi, A. Najid, M. El Bekkali, A. Naji, and M. El Bekkali, "Exponential MLWDF (EXP-MLWDF) downlink scheduling algorithm in LTE for high mobility and dense area scenario," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 3, pp. 1618–1627, 2018.

[13] J. Sachs, G. Wikstrom, T. Dudda, R. Baldemair, and K. Kittichokechai, "5G radio network design for ultra-reliable low-latency communication," *IEEE Netw.*, vol. 32, no. 2, pp. 24–31, Mar. 2018.

[14] S. Ahmadi, "Downlink physical layer functions," *LTE-Advanced*, pp. 399–720, 2014.

[15] I. Klaus, K. Pedersen, G. Pocovi, J. Steiner, and A. Maeder, "Agile 5G scheduler for improved E2E performance and flexibility for different network implementations," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 210–217, 2018.

[16] J. F. De Valgas, J. F. Monserrat, and H. Arslan, "Flexible numerology in 5G NR: Interference quantification and proper selection depending on the scenario," *Mob. Inf. Syst.*, vol. 2021, pp. 1–9, 2021.

[17] G. Pocovi, K. I. Pedersen, and P. Mogensen, "Joint link adaptation and scheduling for 5G ultra-reliable low-latency communications," *IEEE Access*, vol. 6, pp. 28912–28922, 2018.

[18] G. Pocovi, B. Soret, K. I. Pedersen, and P. Mogensen, "MAC layer enhancements for ultra-reliable low-latency communications in cellular networks," in *IEEE International Conference on Communications Workshops*, pp. 1005–1010, 2017.

[19] A. Karimi, K. I. Pedersen, N. H. Mahmood, G. Pocovi, and P. Mogensen, "Efficient low complexity packet scheduling algorithm for mixed URLLC and eMBB traffic in 5G," in *IEEE Vehicular Technology Conference*, (IEEE VTC), pp. 1–6, 2019.

[20] K. I. Pedersen, M. Niparko, J. Steiner, J. Oszmianski, L. Mudolo, and S. R. Khosravirad, "System level analysis of dynamic user-centric scheduling for a flexible 5G design," in *IEEE Global Communications Conference*, (IEEE GLOBECOM), pp. 1–6, 2016.