International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

An Overview of MAC Scheduling Techniques in 5G-NR

Dr. Ashok Kumar A R[1], Minaz I[2]

[1]Associate Professor, Dept of CSE, RVCE, Bangalore
[2]MTech, CNE, RVCE, Bangalore

Abstract

The advent of 5G technology has brought about a paradigm shift in wireless communication, enabling unprecedented data rates, ultra-low latency, and massive connectivity. At the heart of this technological revolution lies the 5G New Radio (5G-NR) standard, a key component that facilitates the seamless interaction between devices and networks. Within the 5G-NR architecture, the Medium Access Control (MAC) layer plays a pivotal role in optimizing the utilization of radio resources and ensuring efficient data transmission. This comprehensive overview delves into the intricate realm of MAC scheduling techniques within the context of 5G-NR. Scheduling, a fundamental aspect of wireless communication, is the process through which resources are allocated to different user devices to enable simultaneous and efficient data transmission. NR is essentially a scheduled system where the scheduler in the gNB controls downlink and uplink transmissions. This paper describes the details around dynamic scheduling, including associated functionality such as buffer-status report and power-headroom reports.

Keywords: Scheduling, dynamic scheduling, preemption, buffer-status report, power-headroom report, scheduling request, DRX

Introduction

NR functions as a meticulously organized system, operating on a predetermined schedule that orchestrates the allocation of temporal, frequency, and spatial resources, as well as the specification of transmission parameters like data rates, to designated devices. The process of scheduling within NR can manifest as either dynamic or semi static in nature. Within the dynamic mode, the scheduler makes per-interval decisions, such as within a slot of time, determining the transmission and reception participants. This approach enables swift adjustments to swiftly shifting traffic demands and the quality of radio channels, effectively harnessing the available resources. In contrast, the semi static variant of scheduling involves furnishing transmission parameters to devices beforehand, eschewing the dynamic real-time allocation approach.

The ensuing discussion will delve into the realms of dynamic downlink and uplink scheduling within NR, encompassing the nuanced concept of bandwidth adaptation.

1 Dynamic Downlink Scheduling

In the context of NR (New Radio), the downlink scheduler plays a crucial role in dynamically controlling the transmission of devices. Each scheduled device is provided with a scheduling assignment that includes essential information as discussed in [1][9]. Usually, the scheduling assignment is transmitted just before the data on the PDSCH, but it can also be scheduled in later OFDM symbols within the same slot or in later slots, allowing for bandwidth adaptation.

To obtain channel condition information, the gNB can utilize various methods, but typically it relies on CSI reports from the device. Assessing the correlation between spatial channels of different devices is also important for multiuser MIMO and spatial isolation estimation [6]. Uplink sounding with SRS transmission can be used along with channel reciprocity assumptions to assess downlink channel quality. Other quantities like signal-strength measurements for different beam candidates can also be used.

For downlink, buffer status and traffic priorities are easily available as the scheduler and transmission buffers are in the same node. Retransmissions are usually prioritized over new data transmission, especially for data flows of the same priority. The scheduler also has the flexibility to select the transmission duration.

Different downlink schedulers may coordinate their decisions to enhance overall performance. This coordination involves avoiding transmission on certain frequency ranges in one cell to reduce interference towards another cell. In dynamic TDD, cells can also coordinate the transmission direction (uplink or downlink) to mitigate interference situations explained in [2]. In the case of carrier aggregation, scheduling decisions are made per carrier, and the scheduling assignments are transmitted separately for each carrier. A device scheduled to receive data from multiple carriers simultaneously receives multiple PDCCHs [3]. The PDCCH received can either point to the same carrier (self-scheduling) or to another carrier with a different numerology (cross-carrier scheduling) referred in Fig 1.1 and discussed detailly in [4][5]. Timing offsets in the scheduling assignment are considered for cross-carrier scheduling of carriers with different numerologies.

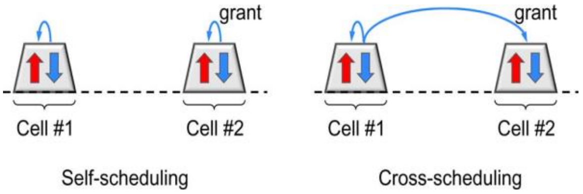

FIGURE 1.1 Self-scheduling and cross-carrier scheduling.

1.1 Bandwidth Adaptation

NR (New Radio) supports an extensive transmission bandwidth, spanning up to several hundred megahertz on a single carrier. This capability proves highly beneficial for quickly delivering substantial data payloads. However, for smaller payload sizes or instances when monitoring the downlink control channels is not scheduled, such a broad bandwidth becomes unnecessary. To address this, NR implements receiver-bandwidth adaptation as discussed in [7], allowing devices to utilize a narrow bandwidth solely for monitoring control channels and opening up the full bandwidth only when significant data transmission is scheduled.

The activation of the wideband receiver is achieved through the use of the "bandwidth part indicator" field in the DCI (Downlink Control Information). When the bandwidth part indicator points to a different segment than the one currently active, the active bandwidth part is changed accordingly (refer to Fig. 1.2). The time required to effect this change depends on various factors, such as whether the center frequency needs retuning or not, but typically it takes around one slot. Once the new, wider bandwidth part is activated, the device utilizes it for its operation

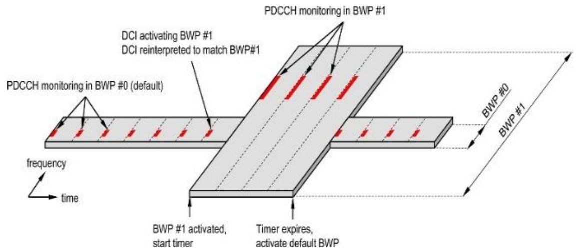

FIGURE 1.2 Illustration of bandwidth adaptation principle.

After completing the data transfer that requires a wider bandwidth, the same mechanism can be employed to revert back to the original bandwidth part. Another option is to utilize a timer for handling the bandwidth-part switching, eliminating the need for explicit signaling. In this case, one of the bandwidth parts is designated as the default. If no default bandwidth part is explicitly set, the initial bandwidth part obtained from the random-access procedure serves as the default. Upon receiving a DCI indicating a bandwidth part other than the default, the timer is activated. Once the timer expires, the device automatically switches back to the default bandwidth part. Typically, the default bandwidth part is narrower, which contributes to reduced device power consumption.

1.2 Downlink Preemption Handling

Dynamically scheduling involves making scheduling decisions within time slots, influenced by subcarrier spacing. NR (New Radio), however, achieves low latency more efficiently by enabling transmission over partial slots starting from any OFDM symbol. This maintains low latency without sacrificing resistance to time dispersion. [1][8]. Fig. 1.3 illustrates an example of this approach. Device A is initially scheduled for a transmission. While this is happening, urgent data for device B arrives, so a transmission for device B is scheduled. Ideally, device B's transmission uses separate resources, but in busy times, it might use some of device A's resources. This is called "preemption" where device B's urgent data takes priority, potentially affecting device A's data reception

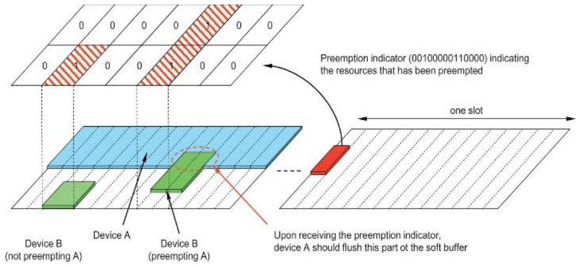

FIGURE 1.3 Downlink preemption indication.

In NR, there are several ways to address this issue. One approach involves using hybrid-ARQ retransmissions. if Device A can't decode data due to limited resources, it signals the base station (gNB) with a "no-go." The gNB responds by resending the data later, either as a full block or just the affected parts, using a technique called hybrid-ARQ retransmissions. This helps ensure successful data delivery.

Another option is to Notify Device A that some of its resources are being used for other purposes by sending a preemption indicator in a slot following the data transmission. The preemption indicator, conveyed through DCI format 2-1, includes a 14-bit bitmap [8]. The bitmap's interpretation is adaptable:

each bit signifies either one OFDM symbol with the entire bandwidth part in the time domain or two OFDM symbols with half the bandwidth part.

2. Dynamic Uplink Scheduling

In dynamic scheduling, the uplink scheduler performs a role similar to the downlink scheduler. It manages which devices can transmit, how they use uplink resources, and sets transmission parameters. Unlike downlink, where power is centralized at the base station, uplink distributes power among devices. Devices have lower max uplink power than the base station, affecting scheduling. Thus, even with much data, there might not be enough power for all transmissions. Scheduled devices get grants specifying time, frequency, and spatial resources for uplink shared channel (UL-SCH) and transport format [9]. Uplink data transmission occurs only with a valid grant otherwise, it's not allowed.

The uplink scheduler has full authority over determining the transport format to be used by a device, and the device must adhere to the scheduling grant. Uplink scheduling primarily revolves around individual devices rather than radio bearers. This arrangement is depicted in the right part of Fig. 1.4, where the scheduler controls the transport format, and the device manages logical channel multiplexing discussed in [9][10].

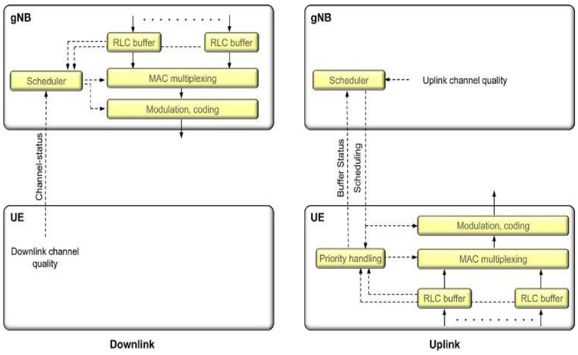

FIGURE 1.4 Downlink and uplink scheduling in NR.

The time allotted for uplink transmission is indicated in the DCI for the device [12]. For half-duplex devices, they must switch their transmission direction before sending data in the uplink. Depending on the uplink-downlink allocation, multiple uplink slots may require scheduling using multiple grants transmitted during the same downlink occasion. Therefore, the timing field in the uplink grant plays a crucial role.

The device needs a specific amount of time for preparation before transmission, as depicted in Fig. 14.5. Shorter preparation time is preferred for better performance, but it can't be too short due to device complexity limits. LTE used to offer over 3 ms for preparation [13], but NR (New Radio) has reduced this time significantly, aiming for lower latency. Fig. 1.1 shows the delay from grant reception to uplink data transmission, linked to subcarrier spacing, though not always directly proportional. Two device capabilities are defined, all must meet baseline requirements, while some can opt for a more aggressive timeline, valuable for low-latency applications[11]

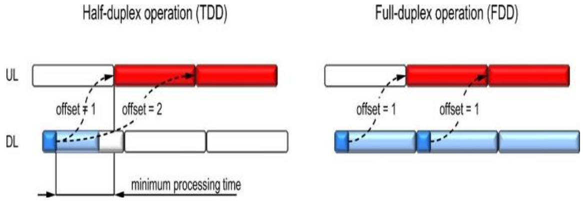

FIGURE 1.5 Example of uplink scheduling into future slots.

2.1 Uplink Priority Handling

The MAC multiplexing feature combines different priority logical channels into a single data block for transmission. When not all data can fit the uplink grant, prioritization is needed. Uplink uses predefined rules with network-set parameters [14][15]. This ensures the grant is for an uplink carrier, not a specific channel within it.

Upon receiving an uplink grant, two steps are taken. Firstly, the device identifies the eligible logical channels that can be multiplexed using the grant. This defines an implicitly derived profile, where only the logical channels meeting the grant's restrictions are allowed to be transmitted at that moment. Additionally, transmission without a dynamic grant may impose restrictions on logical channel multiplexing. Linking the multiplexing rule to the PUSCH (Physical Uplink Shared Channel) duration in 3GPP is motivated by the ability to control whether latency-critical data can utilize a grant intended for less time-critical data. An improved approach involves separately requesting a transmission during a short PUSCH duration for the latency-critical data, achievable by appropriately configuring the maximum PUSCH duration [17]. With the higher-priority configuration of the logical channel carrying latency-critical traffic, the non-latency-critical service won't hinder the transmission of critical data during the short PUSCH duration.

Restricting uplink carriers for certain logical channels is motivated by potential differences in propagation conditions and dual connectivity. After determining the set of logical channels eligible for transmission based on the current grant and mapping-related parameters, the next step involves distributing resources among these logical channels with data to transmit. This distribution relies on priority-related parameters for each logical channel, including priority, prioritized bit rate (PBR), and bucket size duration (BSD) [16]. The product of PBR and BSD represents a minimum number of bits that should be transmitted for a given logical channel within a certain time period. During each transmission instant, the logical channels are served in decreasing priority order, while attempting to meet the minimum bit transmission requirement. If there is excess capacity after serving all logical channels up to their respective bucket sizes, it is distributed in strict priority order

2.2 Scheduling Request

The uplink scheduler requires knowledge of devices that have data to transmit and need to be scheduled. To determine whether a device has data to transmit and should be granted resources, the scheduler relies on what is known as a scheduling request. Scheduling requests are utilized for devices that do not already possess a valid scheduling grant. Devices with a valid grant provide more detailed scheduling information to the gNB.

A scheduling request serves as a flag raised by a device to request uplink resources from the uplink scheduler. Since the device making the request does not have any PUSCH (Physical Uplink Shared Channel) resource by definition, the scheduling request is transmitted on the PUCCH (Physical Uplink Control Channel) using preconfigured and periodically recurring PUCCH resources dedicated to that particular device[18]. By employing this dedicated scheduling-request mechanism, when a device receives data with higher priority than the data already existing in its transmit buffers and it lacks a grant to transmit this new data, it transmits a scheduling request at the next available opportunity. Upon receiving this request, the gNB can then allocate a grant to the device, as illustrated in Figure 1.7.

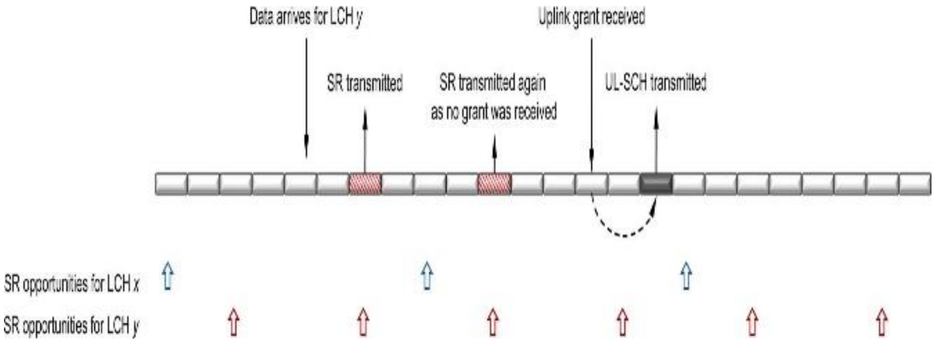

FIGURE 1.7 Example of scheduling request operation.

Each device has the flexibility to be allocated dedicated PUCCH scheduling request resources. The frequency of these assignments can vary, ranging from every second OFDM symbol, catering to ultra-low latency services, to as infrequent as every 80 ms, to ensure low overhead [19]. To avoid contention, only one scheduling request can be transmitted at any given time. In scenarios where multiple logical channels have data ready for transmission, a sensible approach is to prioritize and trigger the scheduling request corresponding to the highest-priority logical channel. To increase the chances of successful scheduling, a scheduling request is replicated in subsequent resources up to a customizable limit. This repetition continues until the gNB (gNodeB) responds with a grant for transmission. Additionally, there's an option to configure a prohibit timer, which regulates the frequency at which a scheduling request can be sent**.**

2.3 Buffer Status Reports

Devices with a valid grant are not required to request uplink resources. Nevertheless, in order to enable the scheduler to allocate appropriate resources to each device in the future, this data is conveyed to the scheduler during the uplink transmission through MAC control [19]. The presence of a buffer status report is indicated by setting the LCID field in one of the MAC subheaders to a reserved value, as shown in Fig. 1.8

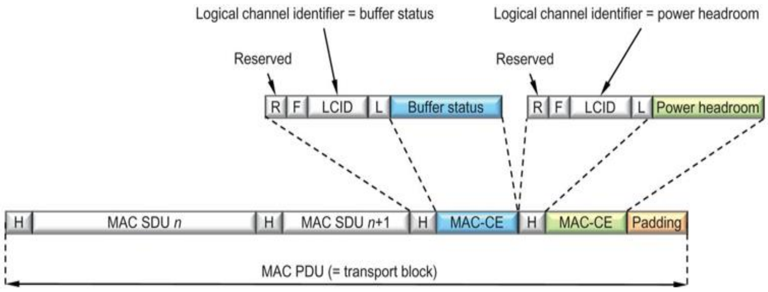

FIGURE 14.8 MAC control elements for buffer status reporting and power headroom reports.

In terms of scheduling, having buffer information for each logical channel is advantageous, even though it may lead to notable overhead. As a solution, the logical channels are organized into a maximum of eight logical-channel groups, and the reporting is performed on a per-group basis. The buffer-size field in a buffer-status report denotes the total data awaiting transmission across all logical channels within a logical-channel group.

2.4 Power Headroom Reports

Additionally, the uplink scheduler takes into consideration not only the buffer status but also the available transmission power in each device. It is essential to avoid scheduling a data rate higher than what the device's transmission power can support. In the downlink, the scheduler easily knows the available power as the power amplifier is located in the same node. But in the uplink, the power availability, or power headroom, must be communicated to the gNB (gNodeB). This is achieved through power headroom reports, which are sent from the device to the gNB when the device is scheduled to transmit on the UL-SCH (Uplink Shared Channel) refer [23] for further details.

The 3 defined types of power headroom reports in NR (New Radio) are Type 1, Type 2, and Type 3. Type 1 power headroom can be reported even when there is no actual PUSCH (Physical Uplink Shared Channel) transmission. It assumes a default transmission configuration corresponding to the minimum possible resource assignment.

Type 2 power headroom reporting is similar to Type 1 but assumes simultaneous PUSCH and PUCCH (Physical Uplink Control Channel) reporting. This feature is not fully supported in the first release of the NR specifications but is planned for finalization in later releases.

Type 3 power headroom reporting is specifically used to handle SRS (Sounding Reference Signal) switching. It comes into play when there are SRS transmissions on an uplink carrier where the device is not configured to transmit PUSCH. The purpose of this report is to evaluate the uplink quality of alternative uplink carriers and, if found beneficial, to (re)configure the device to use this carrier for uplink transmission instead.

Conclusion

The New Radio (NR) system operates with precision through dynamic and semi-static scheduling mechanisms, orchestrating the allocation of resources for efficient communication. The dynamic downlink scheduling within NR involves intricate decisions on transmission assignments, utilizing various techniques such as CSI reports and spatial correlation to optimize channel quality. The concept of bandwidth adaptation is a notable feature, allowing devices to seamlessly switch between wide and narrow bandwidths to optimize resource usage based on data demands. Additionally, downlink preemption handling ensures urgent data transmission takes precedence when necessary, employing techniques like hybrid-ARQ retransmissions and preemption indicators to maintain effective communication. Moving to dynamic uplink scheduling, the NR system continues to exhibit its flexibility by distributing power among devices and effectively managing transmission parameters. The uplink scheduler takes charge of transport format determination, leveraging uplink grants to allocate resources for devices' transmission. Uplink priority handling is crucial for multiplexing different logical channels, and scheduling requests play a pivotal role in requesting and granting uplink resources. Buffer status reports provide valuable information about data awaiting transmission, while power headroom reports ensure that transmission power is optimized and well-suited for each device's capabilities.

In essence, the New Radio system's dynamic scheduling mechanisms, coupled with bandwidth adaptation and efficient handling of preemption and priority, contribute to a highly adaptable and responsive communication framework. This adaptability enables NR to accommodate varying data demands and ensure seamless, low-latency communication for a wide range of applications.

Reference

1. K. Takeda, H. Xu, T. Kim, K. Schober and X. Lin, "Understanding the Heart of the 5G Air Interface: An Overview of Physical Downlink Control Channel for 5G New Radio," in IEEE Communications Standards Magazine, vol. 4, no. 3, pp. 22-29, September 2020, doi: 10.1109/MCOMSTD.001.1900048.
2. H. Kim, J. Kim and D. Hong, "Dynamic TDD Systems for 5G and Beyond: A Survey of Cross-Link Interference Mitigation," in IEEE Communications Surveys & Tutorials, vol. 22, no. 4, pp. 2315-2348, Fourthquarter 2020, doi: 10.1109/COMST.2020.3008765.
3. E. j. Shin, O. -S. Park, S. C. Cho and J. Shin, "PDCCH implementation based on 5G NR system and comparison with LTE PDCCH," 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2020, pp. 962-964, doi: 10.1109/ICTC49870.2020.9289336.
4. A. Mamane, M. Fattah, M. E. Ghazi, M. E. Bekkali, Y. Balboul and S. Mazer, "Scheduling Algorithms for 5G Networks and Beyond: Classification and Survey," in IEEE Access, vol. 10, pp. 51643-51661, 2022, doi: 10.1109/ACCESS.2022.3174579.
5. C. Wang, S. Xu and J. Xin, "The Cross-Carrier Scheduling in LTE-NR Dynamic Spectrum Sharing," 2022 14th International Conference on Computer Research and Development (ICCRD), Shenzhen, China, 2022, pp. 411-414, doi: 10.1109/ICCRD54409.2022.9730277.
6. Z. Liu, S. Sun, Q. Gao and H. Li, "CSI Feedback Based on Spatial and Frequency Domains Compression for 5G Multi-User Massive MIMO Systems," 2019 IEEE/CIC International Conference on Communications in China (ICCC), Changchun, China, 2019, pp. 834-839, doi: 10.1109/ICCChina.2019.8855979.
7. V. Ramaswamy, J. T. Correia and D. Swain-Walsh, "Analytical Evaluation of Bandwidth Part Adaptation in 5G New Radio," 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Helsinki, Finland, 2021, pp. 985-990, doi: 10.1109/PIMRC50174.2021.9569671.
8. W. Yang, C. -P. Li, A. Fakoorian, K. Hosseini and W. Chen, "Dynamic URLLC and eMBB Multiplexing Design in 5G New Radio," 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2020, pp. 1-5, doi: 10.1109/CCNC46108.2020.9045687.
9. J. Rischke, C. Vielhaus, P. Sossalla, S. Itting, G. T. Nguyen and F. H. P. Fitzek, "Empirical Study of 5G Downlink & Uplink Scheduling and its Effects on Latency," 2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), Belfast, United Kingdom, 2022, pp. 11-19, doi: 10.1109/WoWMoM54355.2022.00017.
10. S. Sun, S. Moon and J. -K. Fwu, "Practical Link Adaptation Algorithm With Power Density Offsets for 5G Uplink Channels," in IEEE Wireless Communications Letters, vol. 9, no. 6, pp. 851-855, June 2020, doi: 10.1109/LWC.2020.2973152.
11. M. W. Nomeir, Y. Gadallah and K. G. Seddik, "Uplink Scheduling for Mixed Grant-Based eMBB and Grant-Free URLLC Traffic in 5G Networks," 2021 17th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Bologna, Italy, 2021, pp. 187-192, doi: 10.1109/WiMob52687.2021.9606298.

12. V. Braun, K. Schober and E. Tiirola, "5G NR Physical Downlink Control Channel: Design, Performance and Enhancements," 2019 IEEE Wireless Communications and Networking Conference (WCNC), Marrakesh, Morocco, 2019, pp. 1-6, doi: 10.1109/WCNC.2019.8885990.
13. D. Gonzalez G, M. Garcia-Lozano, S. Ruiz and J. Olmos, "On the Role of Downlink Control Information in the Provision of QoS for NRT Services in LTE," 2012 IEEE 75th Vehicular Technology Conference (VTC Spring), Yokohama, Japan, 2012, pp. 1-6, doi: 10.1109/VETECS.2012.6240095.
14. Ye N, Han H, Zhao L, Wang AH. Uplink nonorthogonal multiple access technologies toward 5G: A survey. Wireless Communications and Mobile Computing. 2018 Jun 12;2018.
15. N. H. Mahmood, R. Abreu, R. Böhnke, M. Schubert, G. Berardinelli and T. H. Jacobsen, "Uplink Grant-Free Access Solutions for URLLC services in 5G New Radio," 2019 16th International Symposium on Wireless Communication Systems (ISWCS), Oulu, Finland, 2019, pp. 607-612, doi: 10.1109/ISWCS.2019.8877253.
16. A. Sengupta, B. Mondal, V. Sergeev, A. Davydov and A. Papathanassiou, "Contiguous Multi-User Scheduling and Power Control for 5G-NR Uplink," 2021 55th Annual Conference on Information Sciences and Systems (CISS), Baltimore, MD, USA, 2021, pp. 1-6, doi: 10.1109/CISS50987.2021.9400277.
17. K. R. G., S. Kumawat, S. Amuru and K. Kuchi, "Enhanced Transport Block Processing for 5G NR PUSCH Coverage Enhancement," 2021 National Conference on Communications (NCC), Kanpur, India, 2021, pp. 1-6, doi: 10.1109/NCC52529.2021.9530122.
18. L. Kundu, G. Xiong and J. Cho, "Physical Uplink Control Channel Design for 5G New Radio," 2018 IEEE 5G World Forum (5GWF), Silicon Valley, CA, USA, 2018, pp. 233-238, doi: 10.1109/5GWF.2018.8517042.
19. Y. -H. Kim, H. Ju, C. B. Jeong and M. -S. Lee, "Performance comparison of DTX detection schemes for 5G NR PUCCH," 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2020, pp. 1391-1394, doi: 10.1109/ICTC49870.2020.9289538.
20. Zhang Q, Nikou A, Daoutis M. Predicting Buffer Status Report (BSR) for 6G Scheduling using Machine Learning Models. In2022 IEEE Wireless Communications and Networking Conference (WCNC) 2022 Apr 10 (pp. 632-637). IEEE.
21. Abraham AK. An optimized method of buffer status reporting for uplink data in lte. In2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT) 2015 Mar 5 (pp. 1-4). IEEE.
22. N. Salhab, R. Rahim, R. Langar and R. Boutaba, "Deep Neural Networks approach for Power Head-Room Predictions in 5G Networks and Beyond," 2020 IFIP Networking Conference (Networking), Paris, France, 2020, pp. 579-583.
23. T. Ranstrom, P. Pietraski and S. Pattar, "5G PUSCH Channel Estimation and Decoding Subject to High-Power Pulse Radar Interference," MILCOM 2022 - 2022 IEEE Military Communications Conference (MILCOM), Rockville, MD, USA, 2022, pp. 437-441, doi: 10.1109/MILCOM55135.2022.10017787.
24. F. Abinader et al., "Impact of Bandwidth Part (BWP) Switching on 5G NR System Performance," 2019 IEEE 2nd 5G World Forum (5GWF), Dresden, Germany, 2019, pp. 161-166, doi: 10.1109/5GWF.2019.8911626.