Comparative Analysis of Enhanced Mobile Broadband and Mini-Slot Contributions to High Reliability and Low Latency of 5G Network

Yang Li*

Xi'an Shiyou University, Xi'an, China

ABSTRACT: As one of the three 5G application scenarios, URLLC has the characteristics of ultra-low latency and ultra-high reliability, and becomes an important breakthrough for traditional communication to enter vertical industries. This paper firstly analyzes the key technologies of low delay and high reliability of URLLC, as well as the related technologies enhanced by URLLC. Futhermore, the examples of 3.5 GHz networks of both China Telecom and China Unicom are adopted to analyze the adaptability of URLLC in TDD networks. Research shows that URLLC technology can maintain the characteristics of low latency and high reliability through a variety of paths, and its adaptability in the Internet also presents a stable trend.

1. INTRODUCTION

January 2019, Fuzhou. An operation led by Liu Rong, director of the department of hepatobiliary and pancreatic cancer surgery at Beijing 301 Hospital, is underway. But oddly enough, the target was 50 kilometers away. It turned out that this was an operation using a 5G network to remotely control the robotic arm, Dr. Liu Rong remote liver lobule removal surgery on a pig. The operation lasted nearly an hour and was a success. It is understood that the remote surgery was jointly carried out by Huawei, China Unicom Fujian Branch and Beijing 301 Hospital, and is the world's first 5G remote surgery. Then I believe that many people will have a question, surgery has always been based on accuracy, and under the circumstances of such a long distance, can be accurate? That's what I'm going to talk about today, the application of communication technology to remote control and how to reduce the delay caused by this kind of communication. Therefore, the purpose of this study is to give people a clearer understanding of how URLLC technology is used in the above practical examples of life, and what other fields will be involved in the future life of this technology.

2. THEORETICAL FRAMEWORK

2.1. Design of uRLLC frame structure based on mini-slot

In 3GPP, the proposals of major companies have defined some of the contents of 5GNR protocol, including radio frame length, resource scheduling unit and multi-subcarrier interval, etc. The minimum scheduling unit is changed from subframe in LTE to time slot in NR protocol, which can support more diverse transmission modes, which means that the frame structure of 5G is more flexible and diverse [1]. For the new service uRLLC, a new transmission scheme is also needed to meet its delay and reliability index requirements.

2.2. Coexistence scheme design of uRLLC and eMBB

The uRLLC, which has received much attention in 5G, is often used in vehicle networking and industrial control, so the wireless frame structure of 5G must be sufficiently compatible with such services. In order to guarantee the user-plane delay of uRLLC, license-free transmission is supported in the 5GNR protocol, in which case different uplink data may conflict and interfere with each other thus reducing reliability [2]. In this paper, we select a coexistence scenario of uRLLC and eMBB services between different users in the uplink channel, and focus on the coexistence transmission scheme of the two services to minimize the performance loss of uRLIC and eMBB in terms of delay and reliability.

*Corresponding author. Email: 1783916881@qq.com

3. URLLC FRAME STRUCTURE

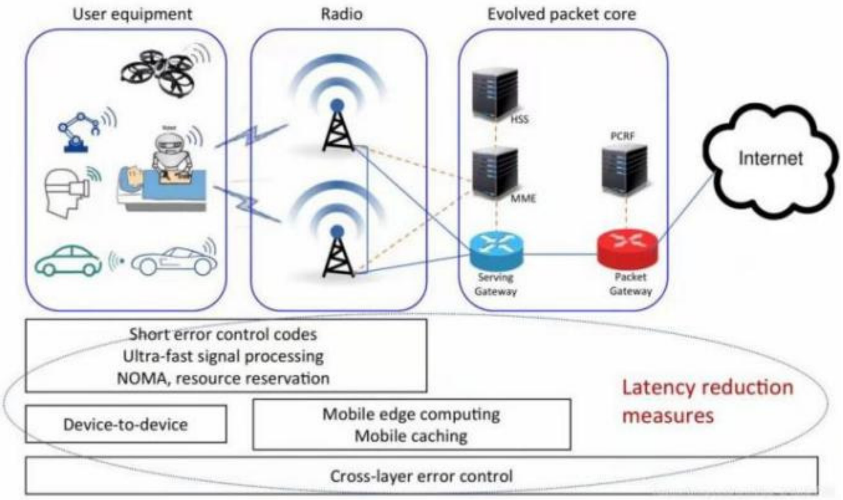

Figure 1 Latency Reduction Measures.

The frame structure has been defined in the 5GNR protocol: a radio frame has a length of 10ms and consists of 10 subframes with a fixed length of 1ms, each of which is composed of several time slots. The normal time slot length is cyclic prefix 14 OFDM symbols and the extended CP has 12 OFDM symbols. 5G NR uses the time slot as the basic scheduling unit, and the time slot length determines the Transmission Time Interval of the packet, which will directly affect the end-to-end delay [3]. To meet some delay-sensitive services, the NR protocol supports a new time slot mini-slot with a minimum length of 1 OFDM symbol. 5G NR supports a subcarrier interval of 15 kHz to the power of two (up to 240 kHz), and a larger subcarrier interval means shorter OFDM symbols [4].

4. URLLC AND EMBB COEXISTENCE SCHEME

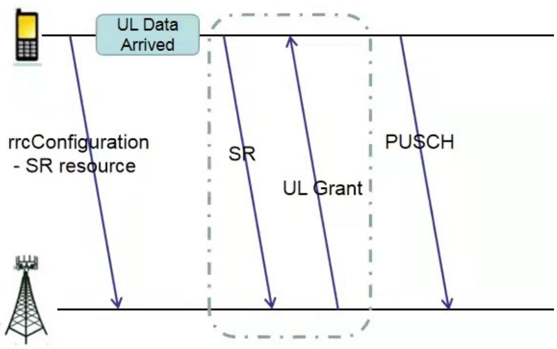

Figure 2 Data.

Because of its service specificity, 3GPP assigns a very high transmission priority to uRLLC. The uRLLC data needs to be sent out immediately upon arrival in order to meet the latency requirements. It may happen that resources are already allocated to eMBB data or eMBB data is being transmitted in the current time period, and uRLLC needs to occupy the resources of the current service for transmission, which will undoubtedly affect the eMBB, resulting in higher BER and lower overall throughput [5].

To minimize the transmission delay of uRLLC, uRLLC has two transmission modes: grant based and grant free. 5G NR introduces three ways of preemption indication, power control and configuration of authorized resources to deal with the coexistence problem in different situations [6].

In the authorization mode, the uRLLC user must send a scheduling request to the base station, wait for the authorization signaling of the base station to allocate time-frequency resources, and finally transmit on the specified resources. If there is no available resource in a short time, the resources of other users must be seized to guarantee the uRLLC delay. That is, the base station tells the preempted users to stop transmission or perform power control on the preempted resources through PI to guarantee the uRLLC transmission [7].

To further reduce the latency, the license-free mode is introduced. In the authorization-free mode, uRLLC users will not need to send scheduling requests and transmit directly. To avoid confusion and reduce inter-user interference, the base station will periodically reserve a portion of dedicated resources for uRLLC. This is the scheme of allocating authorized resources, but the reserved resources are always idle when no uRLLC data arrives, resulting in waste [8]. This aspect of the scheme needs to be further investigated.

Beamforming techniques are classified into Digital Beamforming, Analog Beamforming and Hybrid Beamforming. The structure of pure digital beamforming is that each antenna unit is connected to an RF link. This structure has a large number of RF links, which can fully suppress the interference between different signal streams by beamforming at the baseband digital side, but the number of antenna units in Massive MIMO system is in hundreds and thousands, which makes its hardware implementation difficult [9]. Analog beamforming, on the other hand, allows an RF link to connect multiple antenna units, forming a beam at the analog end of each link to spatially separate the signals, but performance is limited by the number of RF links. Hybrid beamforming is a compromise between hardware complexity and performance, using ABF to reduce the number of RF links and DBF to pre-code between multiple data streams in the baseband to achieve better performance compared to analog beamforming [10]. The two hybrid beamforming structures that have received more attention from developers are the fully connected structure and the fixed subarray structure. The difference between these two structures lies in the way the RF links and antenna units are connected. The fully connected structure connects each RF link to all antenna units in order to fully utilize the gain of each antenna. In contrast, under the fixed sub-array structure, each RF link is connected to some antenna units separately, which greatly reduces the total number of connections. It is a more practical structure in terms of deployment [11].

Although 3GPP has developed a variety of technologies to support URLLC services as described in the previous section, the support of actual networks for URLLC services remains to be studied. In the live network, China Telecom and China Unicom use 2.5ms dual-period frame structure in the 3.5ghz band. If there is data to be sent in the downstream slot, the data can be sent directly. If there is data to be sent in the downstream slot, the data can be sent in the next downstream slot. When K1=1, that is, the delay bias of downstream data sent to ACK feedback is an aggressive value. Without considering retransmission, the minimum delay of downstream data is 1ms, the maximum delay is 3ms, and the average delay is about 1.65ms. At present, the typical value of K1 in the industry is generally 2. At this time, the minimum delay of downlink data is 1.5ms, the maximum delay is 3ms, and the average delay is about 1.95ms. When encountering a downstream slot, the terminal needs to wait for the upstream slot to send the SR and receive the uplink scheduling from the next or two downstream slots. The waiting interval depends on the K5 value. That is, the delay offset from the uplink scheduling request SR to the uplink scheduling authorization. In this case, K5 takes the typical value 2 and waits for the uplink slot to transfer data. Without retransmission, the maximum uplink delay is 2.5 ms, the maximum delay is 5 ms, and the average delay is about 3.75 ms. In this case, the minimum uplink delay is 0.5 ms, the maximum latency is 2.5 ms, and the average latency is about 1.3 ms.

5. URLLC DESIGN UNDER MINI-SLOT

Numerology μ	SCS kHz	Symbol μs	CP μs	Subframes	Slots in a subframe	Slot length μs	*Symbols in a slot*	Subcarriers in a PRB	PRB size MHz
0	15	66.67	~4.8	10	1	1000	14	12	0.18
1	30	33.33	~2.4	10	2	500	14	12	0.36
2	60	16.67	~1.2	10	4	250	14	12	0.72
3	120	8.33	~0.6	10	8	125	14	12	1.44
4	240	4.17	~0.3	10	16	62.5	14	12	2.88

FIGURE 4. The list of 5G NR Numerology.

Figure 3 Numerology.

5.1. Analysis of frame structure and face delay of GNR

Orthogonal Frequency Division Multiplexing (OFDM) technology is used in 5G NR due to its excellent multipath resistance, high spectral efficiency and a series of other advantages [12]. Unlike LTE, which uses DFT-S-OFDM (Discrete Fourier Transform-Spread OFDM) for the uplink, 5GNR uses conventional OFDM for the uplink as well as for the downlink, as it simplifies the design for receivers with spatial multiplexing and unifies the transmission mechanisms for both uplink and downlink. 5G NR still retains DFT-S-OFDM as an auxiliary modulation method for the uplink, as there are scenarios where its advantages of low peak-to-average ratio and high power amplification efficiency are needed [13].

Compared to LTE, 5G NR is not only compatible with the 15 kHz subcarrier spacing, but also supports a wider range of subcarrier spacing configurations, as shown in Table 2-1. The smaller the subcarrier spacing, the higher the spectral efficiency in OFDM systems, while ensuring that the Doppler shift and phase noise do not produce excessive signal errors; and the cyclic prefix (CP) and maximum The relationship between the Cyclic Prefix (CP) and the maximum delay extension limits the upper limit of the subcarrier spacing, so the subcarrier configuration for 5GNR is between 15kHz and 240kHz.

The radio frame length in NR is 10ms, consisting of 10 subframes of 1ms in length, the same length as LTE, which better supports the coexistence of NR and LTE. Such a fixed structure facilitates the synchronisation of time slots and frame structures in LTE and NR co-deployment mode, simplifying cell search and frequency measurement.

It can be seen that the length of the time domain of the radio frame and subframe is fixed for different configurations, the only difference being the number of time slots contained in each subframe. The number of time slots within each subframe is influenced by the subcarrier spacing configuration; for a given subcarrier spacing configuration u, the number of time slots contained in a subframe is 2". There are 14 OFDM symbols per time slot under normal CP and 12 OFDM symbols per time slot under extended CP [14].

Similar to LTE, the channel is represented in a two-dimensional time-frequency physical resource plane as

OFDM symbols in the time domain and as sub-carriers in the frequency domain. An OFDM symbol on a subcarrier is the smallest resource unit (Resource Element, RE), and 12 consecutive subcarriers are a Resource Block (RB), which is the fundamental unit in the frequency domain for resource scheduling, and a series of consecutive RBs with the same configuration is called a BWP (Band Width Part).

The scheduling in 5GNR is based on time slots and supports mini-slot structures with lengths between 1 and 13 OFDM symbols, using smaller scheduling granularity, reducing the number of scheduling symbols, enabling fast transmission and making scheduling more flexible. The minimum number of OFDM symbols in a mini-slot is 1. This transmission mechanism can also be used to change the order of the data transmission queue so that the "mini-slot" transmission data is immediately inserted in front of the regular time slot transmission data already available to a terminal, in order to achieve very low latency[15]. This feature, which eliminates the need to start data transmission at the beginning of each time slot, is particularly useful in scenarios where unlicensed frequency bands are used. In unlicensed bands, the transmitter needs to make sure that the radio channel is not occupied by other transmissions before sending data, using the so-called Listen Before Talk (LBT) strategy. Obviously, data transmission should start as soon as the radio channel is found to be free, rather than waiting until the end of this time slot and the start of the next one.

NR uses both Frequency Division Duplexing (FDD) and Time Division Duplexing (TDD), with different duplexing modes for different frequency bands. Unlike LTE where the subframe is the scheduling unit and the transition point for upstream and downstream, in the NR time slot format, the upstream and downstream services are performed with OFDM symbols as the transition point and there are more types of upstream and downstream symbol configurations. An OFDM symbol in an NR can be classified as "upstream", "downstream" and "flexibly configurable", which makes it possible to switch between upstream and downstream without having to wait until the next time slot starts [16].

The user-plane delay is defined as the one-way transmission time of data packets between the user's Internet Protocol (IP) layer and the P-layer of the Radio Access Network (RAN) edge node, and is one of the main indicators of the user's service data transmission in the physical layer.

5.2. System Design

This study uses MATLAB simulation software to implement the mini-slot frame structure as the frame structure of uRLLC on a link-level simulation platform, and to examine the performance in terms of latency, Bit Error Ratio (BER), Block Error Ratio (BLER) and throughput.

The simulation platform simulates the process of restoring the physical layer data after coding and modulation at the transmitter, passing through the millimetre wave channel, and demodulation and decoding at the receiver.

- Cyclic Redundancy Check (CRC)

A 16-bit check bit is added to the end of the message bits by a specified generating polynomial.

- Channel coding

The channel coding uses Low Density Parity Check (LDPC) with a code rate of 1/2. Since LDPC is a grouping code, the bit sequences are first grouped and the code slice lengths are added, and each group is then encoded in turn according to the generation matrix.

- Rate matching

In order to ensure that the final generated signal rate will match the transmission rate corresponding to the frame structure, the bit stream needs to be repeated and perforated. The scrambling process will add a user-specific pseudo-random sequence to the message sequence on a bit-by-bit basis to encrypt the message and differentiate between users.

- Baseband Modulation

The baseband modulation is Quadrature Amplitude Modulation (QAM), which results in a constellation signal.

- Layer Mapping Precoding

In order to achieve transmission diversity, the modulated symbols are divided into different layers by layer mapping and then mapped to the antenna port by pre-coding.

- Resource mapping

The pre-coded symbols on each antenna port are mapped to the physical time-frequency resource block of that antenna for transmission.

The platform does a Fast Fourier Transform (FFT) of the time domain channel into a frequency domain equivalent channel, simplifying the processing of OFDM modulation, and the symbols at each antenna port are then passed through the channel to the receiver. The received signal at the receiving antenna is sequentially inverse mapped and decoded to obtain the signal sequence. At the receiver end, the DMRS is relied upon to perform channel estimation, and then demodulate and descramble the signal sequence based on the estimated channel. After channel decoding, the signal is converted into a bit stream, which is then used to determine whether retransmission is required based on the check bits.

After a complete send/receive process, the specific error bits of each packet are counted and finally used in the BER and BLER calculations. The retransmission, on the other hand, counts towards the delay consumption.

5.3. Frame structure design

Although the 5GNR changes the minimum scheduling unit from a subframe to a time slot, it is still essentially the time domain length of 14 OFDM symbols, and the latency is not reduced compared to LTE without increasing the subcarrier spacing. In order to ensure the latency metrics of uRLLC in various scenarios, this study uses mini-slot as the physical transmission structure of uRLLC as the main means to reduce latency.

Unlike LTE, 5G NR does not have Cell Specific Reference Signals (CSRS) for transmitting user data, only a Demodulation Reference Signal (DMRS) dedicated to each user for channel estimation. Considering that the mini-slot is relatively short in the time domain, the channel variation is small and a higher subcarrier spacing is possible in the frequency domain, the channel characteristics of one OFDM symbol in the time domain can be considered to approximate the channel characteristics of the whole mini-slot, so the DMRS distribution needs to be designed more tightly.

In order to better demodulate the signal, there are three schemes for the DMRS distribution, all of which occupy the first OFDM symbol of the mini-slot.

model:4 REs are allocated to DMRS in each RB, at the 3rd, 4th, 9th and 10th subcarriers. mode2: Odd number of subcarriers occupying the first OFDM symbol.

mode3: fully occupies all subcarriers of the first OFDM symbol of the mini-slot.

The structure of the designed mini-slot on a single RB is shown in Figure 2-4. The transmission performance of the different DMRS configuration schemes is tested in simulation in the following section.

In terms of latency alone, the shorter the TTI of the mini-slot the better the user-plane latency reduction, but correspondingly the weaker the transmission capability. In order to investigate the variation of delay as the length of the mini-slot changes, and whether the transmission capability can meet the requirements. In this study, 2, 3, 4 and 5 OFDM symbols of mini-slot are selected to study their performance.

6. RESEARCH RELATED TO RESOURCE MANAGEMENT

Traditional wireless resource management techniques mainly include spectrum allocation, power control, user association and other technical means, and the objectives pursued mainly include system capacity, spectrum efficiency, interference management, traffic load balancing and so on. Relevant research on resource management is described as follows:

·Spectrum management

Under the condition of limited spectrum resources, how to manage spectrum resources efficiently is an important research task in 5G heterogeneous networks. As a scarce resource in a heterogeneous network, spectrum utilisation determines the system capacity. In the shared spectrum mode of heterogeneous networks, based on sub-channel sensing techniques, a dynamic spectrum allocation scheme is proposed to improve network spectrum efficiency while suppressing co-channel interference. Under the co-channel deployment of macro base stations and small base stations, the spectrum resource allocation scheme is optimised to improve spectrum utilisation efficiency and maximise system capacity while safeguarding the QoS of users associated with macro base stations. To meet the traffic demand in the network, a traffic-driven spectrum allocation strategy for heterogeneous networks is used to model the traffic and service rate of each base station in the absence and presence of interference, respectively, and to optimise the system spectrum allocation with the target of average system packet sojourn time and throughput[17]. In addition, in order to guarantee the fairness of resource allocation, a proportional fair spectrum allocation scheme is proposed for the two-tier cellular network improvement by considering the fairness between cellular layers under the closed access mode of small base stations.

7. CONCLUSION

Beam training algorithms are divided into two main categories, adaptive beamforming and codebook-based search. Each antenna unit in the adaptive scheme has a phase shifter, and the received signal is used as input to adjust the phase shift dynamically and iteratively according to the adaptive algorithm, and finally converge to get the best beam, such as MUSIC, LMS and other traditional adaptive algorithms. However, this scheme requires multiple iterations, with huge information interaction and computational overhead, and is not very practical in practical application scenarios because the optimal beam is often not calculated in time due to equipment movement and channel changes. In contrast, the codebook search generates a fixed codebook in advance, which consists of several codewords, each of which represents the phase weights of a set of antenna units, i.e., a beam. The beam training process only changes the phase of antenna units according to each code word, without iterative calculation and multiple interactions, which is more practical and is the main research direction in the industry at this stage. Internet technology in constant innovation, cloud computing, fog and edge calculation in the areas of development has built the framework, I hope that through this article, there are more people have a preliminary understanding of the trend of the development of the 5 g at home and abroad, and how should we do in the future to improve under 5 g of related technologies, and in the direction of the continuous optimization for Internet users to reduce costs, Improve efficiency.

REFERENCES

1. Imtiaz, Parvez, Ali, Rahmati, Ismail, & Guvenc, et al. (2018). A survey on low latency towards 5g: ran, core network and caching solutions. IEEE Communications Surveys & Tutorials.
2. Li, C. P. , Jing, J. , Chen, W. , Ji, T. , & Smee, J. , (2017). 5G ultra-reliable and low-latency systems design. 2017 European Conference on Networks and Communications (EuCNC). IEEE.
3. Brunello, D. , Ingemar, J. S. , Ozger, M. , & Cavdar, C. , (2021). Low Latency Low Loss Scalable Throughput in 5G Networks. 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring). IEEE.
4. Wang, C. , Shen, J. , Vijayakumar, P. , Liu, Q. , & Ji, S. , (2021). Ultra-reliable secure data aggregation scheme with low latency for isolated terminals in 5g and beyond defined stins. Computer Standards & Interfaces(6), 103512.
5. Guan, X. , Dube-Demers, R. , Shi, W. , & Rusch, L. A. , (2021). Heterogeneous optical access networks:

enabling low-latency 5g services with a silicon photonic smart edge. Journal of Lightwave Technology, PP(99), 1-1.

6. Ciceri, O. J. , Astudillo, C. A. , Figueiredo, G. B. , Zhu, Z. , & Fonseca, N. , (2021). Passive optical networking for 5g and beyond 5g low-latency mobile fronthauling services.
7. Vu, T. K. , Bennis, M. , Liu, C. , Debbah, M. , & Latvaaho, M. , (2018). Path selection and rate allocation for ultra-reliable and low latency 5g mmwave networks. IEEE.
8. Do, D. T. , Nguyen, T. , Le, C. B. , & Lee, J. W. , (2020). Two-way transmission for low-latency and high-reliability 5g cellular v2x communications. Sensors (Basel, Switzerland), 20(2).
9. Wang, C. , Tang, H. , Wei, Y. , Wang, X. , & Yuan, Q. , (2018). A resource scheduling algorithm with low latency for 5g networks based on effective hybrid genetic algorithm and tabu search. Journal of Xi'an Jiaotong University.
10. Kelvin, Chew, Preetha, & Thulasiramany. Adaptive 5G Low-Latency Communication for Tactile Internet Services.
11. Srinivasulu, O. , Implementation of Low Power, Signal Avilability, Network Coverage, Less Latency, Efficient Bandwidth 5G Technology Communication System.
12. Beyranvand, H. , M Lévesque, Maier, M. , Salehi, J. A. , Verikoukis, C. , & Tipper, D. , (2017). Toward 5g: fiwi enhanced lte-a hetnets with reliable low-latency fiber backhaul sharing and wifi offloading. IEEE/ACM Transactions on Networking, 25(2), 690-707.
13. Mikami, M. , & Yoshino, H. , (2019). Field trial on 5g low latency radio communication system towards application to truck platooning. IEICE Transactions on Communications, E102.B(8).
14. Liu, Shaohan, & Dake. (2019). A high-flexible low-latency memory-based fft processor for 4g, wlan, and future 5g. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 27(3), 511-523.
15. Ohta, A. , (2017). Low latency ARQ scheme for Wireless entrance systems for 5G small cell base stations.
16. Vannithamby, R. , & Talwar, S. , (2016). Towards 5g (applications, requirements and candidate technologies) || low-latency radio-interface perspectives for small-cell 5g networks., 10.1002/9781118979846, 275-302.
17. Zhang, S. , Xu, X. , Wu, Y. , & Lei, L. , (2015). 5G: Towards energy-efficient, low-latency and high-reliable communications networks. 2014 IEEE International Conference on Communication Systems. IEEE.
18. Hashemi, M. , Koksal, C. E. , & Shroff, N. B. , (2017). Hybrid RF-mmWave communications to achieve low latency and high energy efficiency in 5G cellular systems. 2017 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt). IEEE.
19. Esswie, A. A. , & Pedersen, K. I. , (2018). Multi-user preemptive scheduling for critical low latency communications in 5g networks. IEEE.
20. Vardakas, J. S. , Monroy, I. T. , Wosinska, L. , Agapiou, G. , & Verikoukis, C. , (2017). Towards high capacity and low latency backhauling in 5G: The 5G STEP-FWD vision. 2017 19th International Conference on Transparent Optical Networks (ICTON). IEEE.
21. Balevi, E. , & Gitlin, R. D. , (2017). Unsupervised machine learning in 5G networks for low latency communications. IEEE International Performance Computing & Communications Conference (pp.1-2). IEEE Computer Society.
22. Mehrdad, Moradi, Yikai, Lin, Z., & Morley, et al. (2018). Softbox: a customizable, low-latency, and scalable 5g core network architecture. IEEE Journal on Selected Areas in Communications.
23. Jin, R. , Zhong, X. , & Zhou, S. , (2016). The Access Procedure Design for Low Latency in 5G Cellular Network. 2016 IEEE Globecom Workshops (GC Wkshps). IEEE.
24. Nasrallah, A. , Thyagaturu, A. S. , Alharbi, Z. , Wang, C. , Shao, X. , & Reisslein, M. , et al. (2018). Ultra-low latency (ull) networks: the ieee tsn and ietf detnet standards and related 5g ull research. IEEE Communications Surveys & Tutorials.
25. Li, J. , & Chen, J. , (2020). Supporting Low-Latency Service Migration in 5G Transport Networks. Optical Fiber Communication Conference. IEEE.
26. Heinonen, J. , Korja, P. , Partti, T. , Flinck, H. , & Poyhonen, P. , (2016). Mobility management enhancements for 5G low latency services. 2016 ICC - 2016 IEEE International Conference on Communications Workshops (ICC). IEEE.