Ultra-Reliable and Low-Latency Communications in 5G Downlink: Physical Layer Aspects

Hyoungju Ji, Sunho Park, Jeongho Yeo, Younsun Kim, Juho Lee, and Byonghyo Shim

Abstract

URLLC is a new service category in 5G to accommodate emerging services and applications having stringent latency and reliability requirements. In order to support URLLC, there should be both evolutionary and revolutionary changes in the air interface named 5G NR. In this article, we provide an up-to-date overview of URLLC with an emphasis on the physical layer challenges and solutions in 5G NR downlink. We highlight key requirements of URLLC and then elaborate the physical layer issues and enabling technologies including packet and frame structure, scheduling schemes, and reliability improvement techniques, which have been discussed in the 3GPP Release 15 standardization.

Introduction

The new wave of the technology revolution, named the fourth industrial revolution, is changing the way we live, work, and communicate with each other. We are now witnessing the emergence of unprecedented services and applications such as driverless vehicles and drone-based deliveries, smart cities and factories, remote medical diagnosis and surgery, and artificial-intelligence-based personalized assistants (Fig. 1). Communication mechanisms associated with these new applications and services are very different from traditional human-centric communications in terms of latency, energy efficiency, reliability, flexibility, and connection density. Therefore, coexistence of human-centric and machine-type services as well as hybrids of these will render emerging wireless environments more diverse and complex. To address diversified services and applications, the International Telecommunication Union (ITU) has classified 5G services into three categories: ultra-reliable and low latency communication (URLLC), massive machine-type communication (mMTC), and enhanced mobile broadband (eMBB) [1]. To cope with these new service categories, various performance requirements such as massive connectivity, lower latency, higher reliability, and better energy efficiency have been newly introduced. Since the current radio access mechanism cannot support these relentless changes, the Third Generation Partnership Project (3GPP) introduced a new air interface referred to as *New Radio* (NR) [2]. The primary goal of NR is to bring entirely new features and technologies that are not necessarily backward compatible with current fourth generation (4G) LTE systems. Currently, discussions for the 5G NR standardization are underway aiming at commercialization of the first release in 2020.

Among the three categories above, the design of URLLC is perhaps the most challenging. This is because URLLC needs to meet two challenging requirements: *low latency* and *ultra-high reliability*. When we try to improve the reliability, we need to use more resources for signaling, retransmission, redundancy, and parity, resulting in increased latency [3]. In fact, in the current wireless systems, whose sole purpose is to transmit a long packet to maximize the throughput, it is in general very difficult to achieve high reliability and low latency simultaneously. However, when the packet size is small, ultra reliability can be achieved at the expense of the achievable rate reduction [4]. This is because in many URLLC applications where the throughput requirement is not that stringent, reliability can be improved without violating the latency requirement by utilizing resources in the frequency, antenna, and spatial domain. Without doubt, interplay among throughput, latency, and reliability requirements will make physical layer design of URLLC more complicated.

The primary purpose of this article is to present an up-to-date overview of the URLLC communications with an emphasis on the technical challenges and solutions in 3GPP NR downlink. We first describe the URLLC service requirements and then discuss physical layer issues and enabling technologies. These include packet and frame structure, scheduling schemes, and reliability improvement techniques. We also introduce early outcomes in 3GPP NR Release 15.

The rest of this article is organized as follows. In the following section, we explain three service categories in 5G. Then we discuss key requirements of URLLC. Following that, we present physical layer solutions for downlink URLLC services with simulation results and conclude the article in the final section.

Digital Object Identifier: 10.1109/MWC.2018.1700294

Hyoungju Ji, Sunho Park, and Byonghyo Shim are with Seoul National University; Jeongho Yeo, Younsun Kim, and Juho Lee are with Samsung Electronics.

 1536-1284/18/$25.00 © 2018 IEEE

Three Service Categories in 5G

Before we proceed, we provide a brief overview of three service categories in 5G: eMBB, mMTC, and URLLC.

eMBB

eMBB is a service category related to high requirements for bandwidth, such as high-resolution video streaming, virtual reality, and augmented reality. The main challenge in 4G systems is to improve the system throughput (e.g., area, average, peak, perceived, and cell edge throughput). Physical layer technologies introduced to this end include high order modulation transmission, carrier aggregation, cell densification via heterogeneous networks, and multiple-input multiple-output (MIMO) transmission. In essence, the main goal of eMBB is in line with this direction. In order to achieve 100-fold capacity increase over 4G systems, more aggressive physical layer technologies improving the spectral efficiency and exploiting the unexplored spectrum are needed. Technologies under consideration include full-dimension and massive MIMO [5], millimeter-wave communication [6], and spectrally localized waveforms [7].

mMTC

mMTC is a service category to support the access of a large number of machine-type devices. mMTC-based services, such as sensing, tagging, metering, and monitoring, require high connection density and better energy efficiency [8]. Over the years, there have been some trials to support machine-type communications such as narrowband Internet of Things (NB-IoT) in licensed band, and SigFox and LoRa in unlicensed band [9]. These approaches are similar in spirit, but SigFox and LoRa technologies are suited for standalone services, while NB-IoT is a good fit for standards-compatible services. These approaches offer some benefits, such as low power consumption, low operation cost, and improved coverage. However, in the scenario where devices significantly outnumber the resources used for the transmission, an aggressive connection strategy violating the orthogonal transmission principle is required. In recent years, approaches using non-orthogonal spreading sequence or user-specific interleaving have been proposed to accommodate more users than the traditional approach relying on orthogonal multiple access [8].

URLLC

URLLC is a service category to support latency-sensitive services such as remote control, autonomous driving, and tactile Internet [9]. Since the time it takes for human perception or reaction is on the order of tens of milliseconds, packet transmission time for mission-critical applications needs to be on the order of tens to hundreds of microseconds [10]. While the latency of 4G LTE networks was significantly improved from 3G networks, the end-to-end latency is still in the range of 30–100 ms. This is because the backbone network typically uses the best effort delivery mechanism and hence is not optimized for mission-critical service. To reduce the end-to-end latency, therefore, there should be fundamental changes in both wireless link and backbone network. In the backbone link, software defined networking (SDN) and virtual network slicing can be used to construct the private connection to the dedicated URLLC service [9]. Indeed, by using the dedicated network, latency of the backbone link can be reduced significantly. In the wireless link, overhead should be reduced, and the transmission mechanism needs to be streamlined. In fact, since a large portion of the transmit latency is due to the control signaling (e.g., grant and pilot signal), and it takes almost 0.3–0.4 ms per scheduling, it is not very efficient to incorporate a low-latency packet transmission scheme in the current LTE systems. For example, when we design a short packet whose transmission latency is 0.5 ms, more than 60 percent of resources would be wasted for the control overhead. To support URLLC, therefore, many parts of the physical layer should be re-designed.

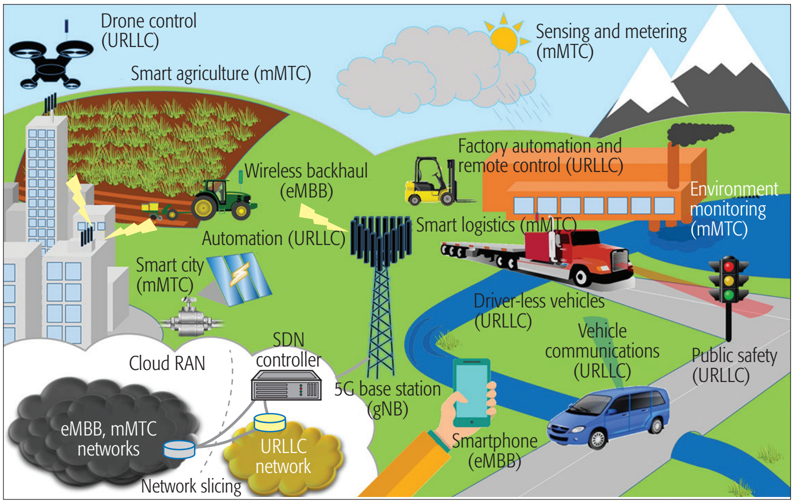


FIGURE 1. Overview of URLLC: deployment with URLLC and other services (eMBB and mMTC) under network slicing.

URLLC Service Requirements

In order to come up with proper solutions to URLLC, it is necessary to understand the key requirements first. In this section, we present the latency and reliability requirements and then discuss the requirement related to the coexistence of URLLC and other services.

Latency Requirement

Physical layer latency T_L can be divided into the following five components (Fig. 2a):

$$T_L = T_{ttt} + T_{prop} + T_{proc} + T_{retx} + T_{sig},$$

- T_{ttt} is the time-to-transmit latency, which corresponds to the time to transmit a packet.
- T_{prop} is the signal propagation time from the transmitter to the receiver.
- T_{proc} is the time to perform the encoding and decoding, and also the channel estimation in the initial transmission.
- T_{retx} is the time taken by retransmission.
- T_{sig} is the pre-processing time for the signaling exchange such as connection request, scheduling grant, channel training and feedback, and queuing delay.

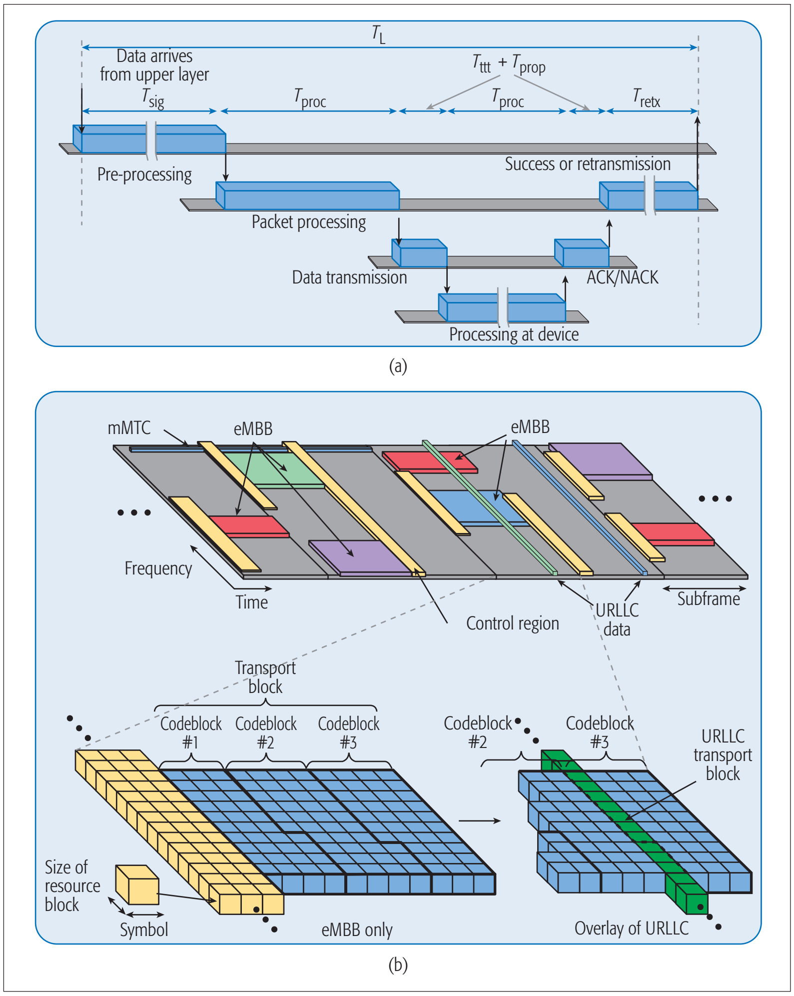


FIGURE 2. Physical layer {downlink} scenarios in URLLC service: a) illustration of latency components; b) transmission of eMBB, mMTC, and URLLC packets at the subframe level, and scheduling of a URLLC packet into an eMBB packet at the symbol level.

In response to the ITU requirements, 3GPP decided that the average latency of URLLC (from layer 2/3, L2/L3, ingression to L2/L3 egression) should be less than 0.5 ms [2]. In order to meet this stringent latency constraint, a packet transmission time T_{ttt} should be on the order of hundreds of microseconds. Since T_{ttt} of the current 4G LTE systems is 1 ms period, a new frame structure reducing T_{ttt} should be introduced. Also, since the latency caused by the channel estimation and feedback would be a bottleneck for the URLLC transmission, a transmission scheme that does not rely on the channel information needs to be considered.

Ultra-High Reliability

In 4G LTE systems, typical reliability for a packet transmission is 0.99. Two key ingredients to achieve this goal are the channel coding (convolution and turbo code) and the partial retransmission of erroneous transport block called hybrid automatic repeat request (HARQ). URLLC services require much better performance, and in fact, the target reliability within 1 ms period should be at least 0.99999 [2]. Further, in the mission-critical applications such as autonomous driving and remote surgery, the reliability should be as high as $1 - 10^{-7}$ [9]. The first thing to do to meet these stringent requirements is to improve the channel estimation accuracy. This is because the channel coding gain is small for the short packet, so the loss, if any, caused by the channel estimation should be prevented as much as possible. This is done by adding more resources to the pilot and using an advanced channel estimation technique. Even in this case, the required URLLC performance might not be satisfied, so additional resources in the frequency, antenna, and spatial domains are required to improve the reliability. Further, an advanced channel coding scheme suitable for the short packet transmission should be employed.[1] If the slot length is very short, a repetitive transmission scheme using time-domain resources can also be a viable option.

Coexistence with eMBB and mMTC

When there is a URLLC service request, whether in the scheduling period or in the middle of eMBB or mMTC transmission, the base station should transmit the URLLC packet immediately [7]. In other words, to support the URLLC packet transmission, ongoing eMBB and mMTC packets should be stopped without notice. As illustrated in Fig. 2b, when a transport block consisting of three codeblocks is transmitted for the eMBB service, each codeblock is mapped sequentially to the scheduled time-frequency resources. Thus, when the URLLC service is initiated in the middle of the eMBB transport block, some of the symbols in the third codeblock are replaced by the symbols of the URLLC packet. Since this interruption is not reported to the mobile devices in use, reception quality of the eMBB and mMTC services will be degraded severely. This problem, referred to as a coexistence problem in the 3GPP NR discussion, is a serious concern to non-URLLC services, so a proper mechanism to protect the ongoing services should be introduced.

URLLC Physical Layer in 5G NR

In contrast to the 4G LTE systems, latency, reliability, and throughput requirements should be jointly considered in 5G NR; hence, there should be a fundamental change in the physical layer architecture (packet, slot, and frame). Specifically, a latency-sensitive packet structure for a fast decoding process and a flexible frame structure to support dynamic change of the resource grid based on the latency requirement are needed. Also, when the URLLC service is initiated, the URLLC packet should be transmitted instantly without delay. To do so, a scheduling scheme minimizing the transmit latency of the URLLC packet should be introduced. Further, since the latency requirement might not be satisfied by the HARQ-based retransmission unless TTI of a packet is very short, a mechanism that significantly reduces the retransmission latency is required. Besides, an approach to use multiple radio interfaces to reduce the latency can be employed. The basic idea of this approach is to choose the radio access technology (RAT) providing the minimum latency among all possible options including 4G LTE, 5G NR, WiFi, and other IEEE 802.x standards. Using this together with device-to-device (D2D) communications, the network layer latency can be reduced substantially.

In this section, we put our emphasis on the physical layer solutions for URLLC including packet

[1] In LTE systems, 12 or 24 resource elements are allocated for demodulation reference signal (DMRS) per resource block (RB).

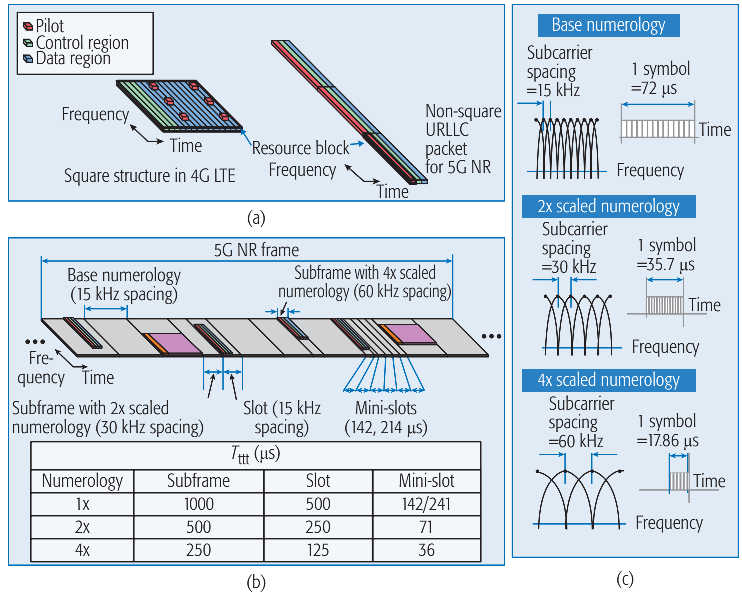
and frame structure to minimize the latency, multiplexing schemes to overlay the URLLC service into eMBB and mMTC services, and approaches to deal with the coexistence problem. We note that the reliability improvement and latency reduction are equally important for the success of URLLC. However, there has been a consensus in the 3GPP NR standard meeting that the latency reduction issue should be considered by priority. This is because the reliability improvement can be achieved by the elaboration of 4G techniques such as channel coding, antenna, space, and frequency diversity schemes, but such is not the case for the latency reduction effort [11, 12].

Packet Structure

The key issue in URLLC packet design is to minimize the processing latency T_{proc} and the time-to-transmit latency T_{ttt}. Note that T_{proc} consists of the time to receive packets, acquire channel information, extract control (scheduling) information, decode data packets, and check errors. In LTE systems, a square-shaped packet structure is popularly used for efficient utilization of the spectrum under channel fading. In 5G NR systems, a non-square packet stretched in the frequency axis is used as a baseline since this structure minimizes the transmission latency T_{ttt} [13]. Furthermore, in order to reduce the latency T_{proc}, three components of a packet (pilot, control, and data parts) should be grouped together to make a pipelined processing of the channel acquisition, control channel decoding, and data detection (Fig. 3a).

Another important issue to be considered is to use an advanced channel coding scheme. In 4G systems, two types of approaches have been employed to ensure the reliability requirement. The first type is the channel coding scheme (turbo and convolution code) with cyclic redundancy check (CRC) attachment for the large packet. The second type is to use a simple code (the repetition and Reed-Muller code) without CRC attachment for the small packet. In 5G NR, polar code and low density parity check (LDPC) code have been adopted for the enhancement of control and data channel, respectively. Over the years, many efforts have been made to improve the decoding performance and computational complexity (and hence processing latency) of these codes such as successive cancellation list decoding of polar code and non-binary LDPC decoding [12].

A recently proposed approach for the second type is sparse vector coding (SVC), a short-packet transmission scheme based on the principle of compressed sensing (CS) [14]. The basic idea of SVC is to map the information into the position of a sparse vector $\mathbf{s}$. Note that when we choose k positions in the n-dimensional vector, $\lfloor \log_2 \binom{n}{k} \rfloor$-bit information can be encoded. For example, if $k = 2$ and $n = 9$, the 5 bits of information can be encoded (e.g., $\mathbf{s} = [0\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0]^T$). By spreading the sparse vector $\mathbf{s}$ using the non-orthogonal spreading sequences $\mathbf{c}_i$, $i = 1, ..., n$, we obtain the transmit vector $\mathbf{x} = [\mathbf{c}_1\ \mathbf{c}_2 \cdots \mathbf{c}_n]\mathbf{s}$ whose dimension m is smaller than n (the dimension of $\mathbf{s}$). Since $\mathbf{s}$ is a sparse vector and k is known in advance, the information vector $\mathbf{s}$ can be recovered via the CS technique [15]. A well-known advantage of the CS technique is that an input vector $\mathbf{s}$ can be recovered using a small number of measurements (resources). Since the computational complexity of the sparse recovery algorithm is proportional to the sparsity k, latency caused by the CS-based algorithm would be very marginal. As shown in Fig. 4a, physical downlink control channel (PDCCH) with SVC is effective in the short packet transmission and outperforms PDCCH using the convolution code in LTE-Advanced systems (3.1 dB at 10^{-5} PER).



T_{ttt} (μs)			
Numerology	Subframe	Slot	Mini-slot
1x	1000	500	142/241
2x	500	250	71
4x	250	125	36

FIGURE 3. Packet and frame structure for URLLC: a) packet structure; b) frame structure; c) supported numerologies for 5G NR.

Frame Structure and Latency-Sensitive Scheduling Schemes

One of the main goals in 5G NR is to design a unified frame structure to cover a wide range of frequency band and various service categories. To this end, a flexible frame structure and a user scheduling mechanism have been introduced.

Flexible Frame for URLLC: One direct option to reduce the time-to-transmit latency T_{ttt} is to reduce the symbol period (Fig. 3b). When the frequency band above 6 GHz (millimeter-wave) is used, due to the path loss, cell radius would be much smaller than that of conventional cellular systems, and thus the channel delay spread will also be small. In this case, by controlling the subcarrier spacing, we can reduce the symbol period (Fig. 3c). For instance, the symbol length can be reduced by half (from 72 μs to 36 μs) by doubling the subcarrier spacing (from 15 kHz to 30 kHz). In doing so, the time to transmit one subframe can be reduced by half (from 1 ms to 0.5 ms). However, when the frequency band below 6 GHz is used, this option might not be desirable due to the large delay spread. In this case, one can alternatively consider reducing the transmission time interval (TTI) of the packet. For example, using mini-slot level (2~3 symbols) and slot level (7 symbols) transmission, T_{ttt} can be reduced to 142, 241, and 500 ms, respectively. In short, by controlling the symbol period and also the number of symbols in a packet, T_{ttt} smaller than 1 ms can be achieved (the table in Fig. 3b). Note that to support this flexible frame

The holy grail of 5G NR is to support diverse service categories and thus how to mitigate the performance degradation of interrupted services is an important issue in the physical layer design. While the flexible frame structure may ease off this problem, due to the implementation complexity and randomness of URLLC packet arrival, a more deliberate solution is required in real deployment scenarios.

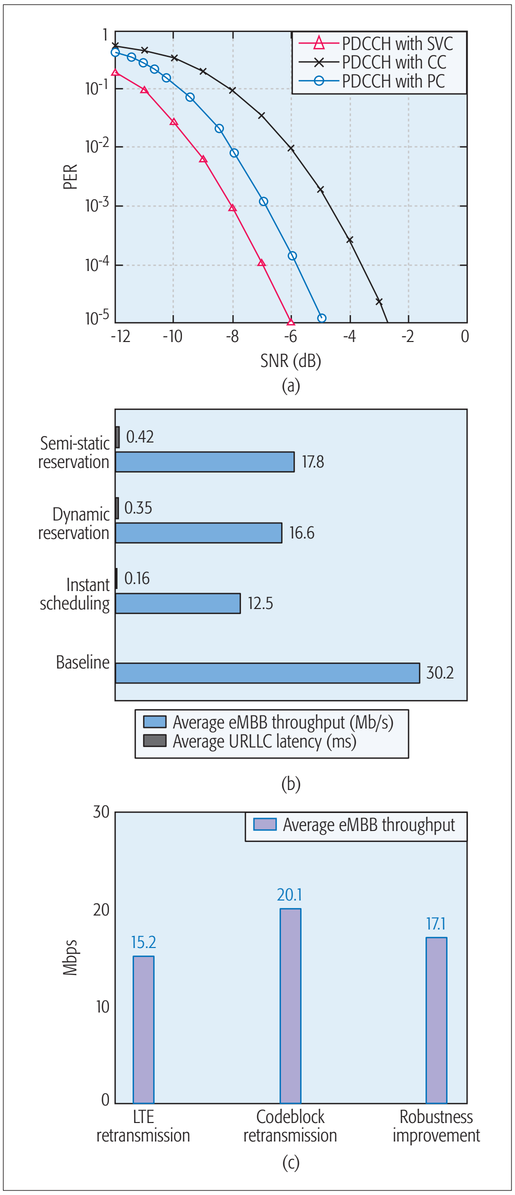


FIGURE 4. Reliability and latency performance for the downlink transmission: a) PER of the short packet transmission (12 bits) for URLLC; b) average latency of URLLC scheduling schemes and impact on eMBB average throughput; c) average throughput performance of coexistence options when the instant URLLC scheduling occurs.

structure, an advanced receiver equipped with fast tracking, quick synchronization, and simultaneous decoding functions is needed.

Scheduling Schemes: Since a URLLC packet is generated abruptly, how to schedule this into existing services is an important issue in system design. Two schemes adopted in the 3GPP NR standard are instant scheduling and reservation-based scheduling [13].

Instant Scheduling: Any ongoing data transmission is interrupted to initiate the URLLC packet. This protocol is effective in reducing the URLLC access time but causes severe performance degradation. Therefore, an approach to mitigate the performance degradation of ongoing services is needed.

Reservation-Based Scheduling: In this scheme, URLLC resources are reserved prior to the data scheduling. Two types of reservation schemes are *semi-static* and *dynamic* reservations. In the semi-static reservation scheme, the base station infrequently broadcasts the configuration of the frame structure such as frequency numerology and service period. In the dynamic reservation scheme, information on the URLLC resource is updated frequently using the control channel of a scheduled user. For example, if an eMBB packet consists of 14 symbols, 10 symbols are used for the eMBB transmission, and the rest are reserved for the URLLC service. The drawback of this approach is that when there is no URLLC transmission in the scheduled period, resources reserved for the URLLC service will be wasted. When compared to semi-static reservation, dynamic reservation requires additional control overhead to indicate the reservation information. Also, overhead to ensure the reliability of the control signaling itself is unavoidable.

In Fig. 4b, we present the simulation results of average latency of URLLC and average throughput of eMBB service. In each scheduling period, the packet size is chosen such that the target packet error rate (PER) of URLLC is 10^{-5}. In the simulation, latency is defined as the time in which a packet, having arrived at the scheduler, is successfully delivered to the mobile. Thus, one can deduce that the throughput is degraded when the latency increases. From Fig. 4b, we see that the instant scheduling strategy outperforms reservation-based scheduling in terms of the average latency but causes a throughput loss of the eMBB service. We also observe that dynamic reservation outperforms semi-static reservation in terms of the latency due to the fast resource adaptation.

Solutions to the Coexistence Problem

As mentioned, the holy grail of 5G NR is to support diverse service categories, and thus how to mitigate the performance degradation of interrupted services is an important issue in the physical layer design. While the flexible frame structure may ease this problem, due to the implementation complexity and randomness of URLLC packet arrival, a more deliberate solution is required in real deployment scenarios. Two approaches discussed in the 5G NR standard meetings are *reactive* and *proactive* strategies.

Reactive Strategy: The main idea is to give priority to the URLLC packet while ensuring the reliability of the other channel interrupted by URLLC. Two approaches adopted in the 3GPP NR are as follows [7].

- Preemption indicator transmission: In this approach, the base station indicates which resources are used for the URLLC transmission. Recalling that the URLLC packet is stretched in the frequency axis (Fig. 3a), URLLC transmission will interrupt the whole system bandwidth and thus degrade all data channels in use. To report this event to the scheduled users, the base station broadcasts a *preemption indicator* consisting of time and/or frequency information of the interruption. This indicator helps users identify the reason for packet errors and what part of the packet is safe from the interruption.

- Retransmission of selected codeblocks: When the ongoing service is interrupted by the URLLC transmission, part of the codebook affected by URLLC is retransmitted. By transmitting a combining indicator or a flush-out indicator, the receiver can perform soft symbol combining of the transmitted and retransmitted codeblocks. One can further achieve better coding gain by lowering the code rate of the retransmitted codeblock.

Proactive Strategy: If the URLLC transmission occurs frequently, the efficiency of the reactive approach will be reduced due to the frequent retransmissions. The main idea of the proactive strategy is to ensure the reliability of ongoing services while supporting the URLLC transmission. Specific schemes to support the proactive strategy include robustness improvement and service sharing.

- Robustness improvement: To reduce the initial packet error of non-URLLC packets, the base station intentionally lowers the code rate by adding extra parity bits or employing outer error correction code to the non-URLLC packets [13]. Since the URLLC data transmission interferes non-URLLC packets, this approach can help reduce the packet error of non-URLLC packets.
- Resource sharing: This strategy supports the ongoing data channel and the URLLC data channel simultaneously. Multiple-antenna or beam-domain techniques are employed for this purpose. Basically, the spatial layer (rank) of the channel is divided in two, and then one part is used for eMBB and the other for URLLC. If there is no extra spatial layer, power-domain non-orthogonal transmission can be applied [8].

In Fig. 4c, we plot the average throughput of 4G LTE HARQ, reactive strategy (retransmission of selected codeblock), and proactive strategy (robust channel coding). We first observe that the performance degradation caused by the instant access is mitigated by the coexistence solutions. We also observe from Fig. 4c that the retransmission of selected codeblocks outperforms the LTE HARQ scheme by a large margin, achieving 32 percent gain in throughput. However, due to the substantial parity overhead, we observe that the robustness improvement scheme suffers from throughput loss (about 17 percent) over the reactive strategy.

Conclusion

URLLC is one of the key services in 5G communications having wide applications including automated controls, tactile Internet, remote operations, and intelligent transportation systems. Despite its importance, the physical layer technologies to seamlessly integrate URLLC into 5G NR are in their infancy. In this article, we have discussed the key requirements for URLLC and presented the physical layer enabling technologies. In order to satisfy stringent latency and reliability requirements, many parts of the physical layer should be re-designed, and the techniques presented in this article can serve as a starting point. Other than the solutions we have discussed, there are many interesting issues worth exploring such as the beamforming strategies for control and data part and the reconfigurable URLLC protocol. Also, study of advanced transceiver architecture to support dynamic numerology adaptation and simultaneous decoding is needed. In this article, we put our focus on the URLLC transmission in the downlink, but there are many open issues for the uplink direction such as one-shot access, active user detection, and grant-free transmission.

Parameter	Value and description
Coding scheme (Fig. 4a)	PDCCH with SVC: $n = 92$, $m = 42$, $k = 2$ [14], PDCCH with convolution code (CC): 1/3 rate, PDCCH with polar code (PC): 1/4 rate [12]
Simulation method	Link level with 3 km/h extended pedestrian-A (EPA) channel
System model	20 MHz bandwidth with 3.5 GHz center frequency, 15 kHz spacing, and 1 slot = 1 ms
Scheduler	Frequency-selective (6 RBs) for eMBB and full-bandwidth (100 RBs) for URLLC.
Receiver	Minimum mean square error (MMSE) using DMRS
URLLC transmission	• Instant scheduling: Randomly selects one symbol within a slot • Reservation: 4 symbols are reserved (fixed in semi-static and adapted in dynamic) • TTI = 1/14 ms, and maximum number of re-transmission = 2
eMBB transmission	• LTE re-transmission: set target block error rate (BLER) = 10^{-2} • Retransmission of codeblock: set target BLER = 10^{-2}, per codeblock retransmission • Robustness improvement: set target BLER = 10^{-3}, per transport block retransmission • TTI = 1ms, and maximum number of retransmission = 4

TABLE 1. Simulation assumptions.

Acknowledgment

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (2014R1A5A1011478) and a Cross-Ministry Giga KOREA Project grant funded by the Korea government (MSIT) (No. GK17P0501, Development of Ultra Low-Latency Radio Access Technologies for 5G URLLC Service).

References

[1] ITU-R Rec. M.2083-0, “IMT Vision — Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond,” Sept. 2015.
[2] 3GPP Tech. Rep. 38.913, “Study on Scenarios and Requirements for Next Generation Access Technologies (Release 14),” v14.2.0, 2017.
[3] B. Soret *et al.*, “Fundamental Tradeoffs among Reliability, Latency and Throughput in Cellular Networks,” *Proc. GLOBECOM*, Dec. 2014, pp. 1391–96.
[4] Y. Polyanskiy, H. V. Poor, and S. Verdú, “Channel Coding Rate in the Finite Blocklength Regime,” *IEEE Trans. Info. Theory*, vol. 56, no. 5, May 2010, pp. 2307–59.
[5] H. Ji *et al.*, “Overview of Full-Dimension MIMO in LTE-Advanced Pro,” *IEEE Commun. Mag.*, vol. 55, no. 2, Feb. 2017, pp. 176–84.
[6] S. Han *et al.*, “Large-Scale Antenna Systems with Hybrid Analog and Digital Beamforming for Millimeter Wave 5G,” *IEEE Commun. Mag.*, vol. 53, no. 1, Jan. 2015, pp. 186–93.
[7] 3GPP Tech. Rep. 38.802, “Study on New Radio Access Technology Physical Layer Aspects (Release 14),” v14.1.0, 2017.
[8] C. Bockelmann *et al.*, “Massive Machine-Type Communications in 5G: Physical and MAC-layer Solutions,” *IEEE Commun. Mag.*, vol. 54, no. 9, Sept. 2016, pp. 59–65.
[9] P. Schulz *et al.*, “Latency Critical IoT Applications in 5G: Perspective on the Design of Radio Interface and Network Architecture,” *IEEE Commun. Mag.*, vol. 55, no. 2, Feb. 2017, pp. 70–78.
[10] ITU-T Watch Report, “Tactile Internet,” Aug. 2014.
[11] N. A. Johansson *et al.*, “Radio Access for Ultra-Reliable and Low-Latency 5G Communications,” *Proc. IEEE ICC Wksp.*, June 2015, pp. 1185–89.

[12] M. Sybis *et al.*, "Channel Coding for Ultra-Reliable Low-Latency Communication in 5G Systems," *Proc. IEEE VTC-Fall*, Sept. 2016, pp.1–5.

[13] 3GPP Tech. Spec. 38.211, "Technical Specification Group Radio Access Network, NR (Release 15)," v15.0.0, 2017.

[14] H. Ji, S. Park, and B. Shim, "Sparse Vector Coding of Control Channel for Ultra-Reliable and Low Latency Communications," to appear, *IEEE Trans. Wireless Commun*., 2018.

[15] J. W. Choi *et al*., "Compressed Sensing for Wireless Communications: Useful Tips and Tricks," *IEEE Commun. Surveys & and Tutorials*, vol. 19, Feb. 2017, pp. 1527–50.

Biographies

Hyoungju Ji is currently working toward a Ph.D. degree at the School of Electrical and Computer Engineering, Seoul National University, Korea. He joined Samsung Electronics in 2007, and has been involved in 3GPP RAN1 LTE, LTE-Adv., LTE-A Pro, and NR technology developments and standardizations. His current interests include multi-antenna techniques, massive connectivity, machine-type communications, and IoT communications.

Sunho Park received his B.S., M.S., and Ph.D. degrees from the School of Information and Communication, Korea University, Seoul, in 2008, 2010, and 2015, respectively. From 2015 to 2017, he was with the Institute of New Media and Communications, Seoul National University, as a research assistant professor. He is currently a senior engineer with Samsung Electronics. His research interests include wireless communications and signal processing.

Jeongho Yeo received his B.S. and Ph.D. degrees in electrical engineering from Pohang University of Science and Technology, Korea, in 2006 and 2014, respectively. Since September 2014, he has been with Samsung Electronics and has been working for 4G LTE Advanced Pro and 5G NR standardization in 3GPP.

Younun Kim received B.S. and M.S. degrees in electronic engineering from Yonsei University, and his Ph.D. degree in electrical engineering from the University of Washington, in 1996, 1999, and 2009, respectively. He joined Samsung Electronics in 1999 and has since worked on the physical later standardization of cdma2000, HRPD, LTE, and recently NR. Currently, he is serving as the Vice Chairman of the 3GPP RAN1 (physical layer) working group.

Juho Lee is currently a Master (technical VP) with Samsung Electronics, where he is leading research and Samsung's activity for standardization of 5G in 3GPP. He also actively worked on 3G and 4G since he joined Samsung Electronics in 2000. He was a Vice Chairman of 3GPP RAN1 from February 2003 to August 2009. He received his Ph.D. degree in electrical engineering from Korea Advanced Institute of Science and Technology in 2000.

Byonghyo Shim is a professor in the Department of Electrical and Computer Engineering at Seoul National University. He received B.S. and M.S. degrees in control and instrumentation engineering from Seoul National University in 1995 and 1997, respectively, and an M.S. degree in mathematics and a Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign in 2004 and 2005, respectively. From 2005 to 2007 he worked for Qualcomm Inc., and from 2007 to 2014 he was with Korea University. His current research focuses on 5G wireless communications, machine learning, and information theory.