Feedback Enhancements for Semi-Persistent Downlink Transmissions in Ultra-Reliable Low-Latency Communication

Trung-Kien Le†, Umer Salim⋆, Florian Kaltenberger†
†EURECOM, Sophia-Antipolis, France
⋆TCL Mobile, Sophia-Antipolis, France
Trung-Kien.Le@eurecom.fr

***Abstract*—Ultra-reliable low-latency communication (URLLC) is one of the new service categories specified by The 3rd Generation Partnership Project (3GPP) that will be used for the applications with the strict reliability and latency requirements. To satisfy these requirements in Release 16, URLLC downlink (DL) transmission can be scheduled in the semi-persistent scheduling (SPS) resources to reduce control overhead and a retransmission if necessary is triggered by hybrid automatic repeat request (HARQ) feedback transmitted in uplink (UL) resources. However, in time division duplex (TDD) mode, HARQ feedback is dropped when there are no available UL resources for physical uplink control channel (PUCCH) in the indicated slot/sub-slot because that slot/sub-slot contains DL symbols in semi-static TDD configuration or flexible symbols updated to DL symbols in dynamic TDD configuration. Dropping the feedback results in a degradation of DL SPS quality of service (QoS) and may not be acceptable for URLLC based services and applications. This paper provides a scheme to guarantee the transmission of HARQ feedback and potential data retransmission when there is UL-DL slot/sub-slot conflict at the indicated feedback resource. The scheme comprises a dynamic indication of feedback resource without using an associated downlink control information (DCI). An acknowledgement (ACK)-only feedback protocol is proposed to best suit the scenario in question. The combination of dynamic indication of feedback resource and ACK-only feedback structure guarantees higher reliability and brings flexibility to the transmission of HARQ feedback as confirmed by the simulations.**

***Index Terms*—5G, URLLC, downlink semi-persistent scheduling transmission, feedback, time division duplex**

I. Introduction

A. URLLC in 5G New Radio (NR)

3GPP specified three service categories in 5G NR consisting of enhanced mobile broadband (eMBB), massive machine-type communication (mMTC) and URLLC to meet the requirements of the technology revolution with the advent of new applications such as augmented reality, remote surgery, industrial automation, to name but a few. URLLC aims at the application with the stringent requirements of reliability and latency and has started to become one of the main focuses of 5G physical layer design.

The requirements of URLLC are defined by 3GPP in [1]: 'A general URLLC reliability requirement for one transmission of a packet is 10^{-5} for 32 bytes with a user plane latency of 1 ms". The definitions of latency and reliability are clarified in [2]: "Reliability can be evaluated by the success probability of transmitting X bytes within a certain delay, which is the time it takes to deliver a small data packet from the radio protocol layer 2/3 service data unit (SDU) ingress point to the radio protocol layer 2/3 SDU egress point of the radio interface, at a certain channel quality (e.g., coverage-edge)". Even in some scenarios, URLLC is requested to achieve higher latency and reliability requirements as noted in [3]: "Higher reliability (up to 10^{-6}), short latency in the order of 0.5 to 1 ms, depending on the use cases (factory automation, transport industry and electrical power distribution)."

B. Techniques standardized by 3GPP for 5G NR to improve URLLC performance

To make the transmission achieve URLLC requirements, 3GPP standardized some techniques in physical layer design.

3GPP specified new numerology for 5G frame structure. Subcarrier spacing (SCS) has flexible values: 15 kHz, 30 kHz, 60 kHz, 120 kHz and 240 kHz [4]. A higher value of SCS reduces time length of Orthogonal Frequency Division Multiplexing (OFDM) symbols.

Physical downlink shared channel (PDSCH)/physical uplink shared channel (PUSCH) mapping Type B where one transmission time interval (TTI) contains 2, 4 or 7 OFDM symbols specified as sub-slot is standardized to supplement to PDSCH/PUSCH mapping Type A where one TTI contains one slot (14 symbols) [4].

Two new downlink control information (DCI) formats with a smaller payload size are supported to schedule UL and DL data transmission to increase DCI reliability [5].

The base station (gNB) can configure the configured grant (CG) resources to the user equipment (UE) so that the UEs can transmit UL data in these CG resources without sending a scheduling request (SR) and waiting for an UL grant [6]. The UEs are also configured to transmit automatically a number of repetitions without feedback from the gNB [6], [7]. These techniques lead to a decrease of transmission and processing time due to SR, UL grant and feedback.

Based on the foundation of the standardized techniques, this paper focuses on the enhancements of feedback for the transmission in DL SPS resources that is similar to the transmission in UL CG resources. The problem of feedback cancellation due to slot/sub-slot conflict in TDD configuration and the prior art are presented in Section II. The proposed scheme that allows dynamic K1 indication and an ACK-only feedback structure is introduced in Section III. This scheme eliminates HARQ feedback's drop and ensures the performance of DL SPS transmission without increasing payload of the activation DCI or using the associated DCI. Section IV shows the result of the proposed scheme. Some remarks of this work are provided in Section V.

II. HARQ FEEDBACK IN DL SPS TRANSMISSION IN TDD CONFIGURATION

A. Feedback cancellation in DL SPS transmission in TDD configuration

To reduce overhead, alignment time, transmission time and processing time of physical downlink control channel (PDCCH), the gNB configures DL SPS resources which occur periodically for a set of UEs. In DL SPS transmission, the gNB configures partially the parameters of DL SPS configurations using higher layer signaling such as RRC signaling. After having made the configurations, the gNB activates the desired DL SPS configuration by sending a DCI using a single cell-radio network temporary identifier (SC-RNTI). Thereby, the UEs are aware of the locations and the parameters of the potential PDSCH transmission such that the gNB can transmit PDSCH in the DL SPS resources to the UE without an associated PDCCH scheduling the particular transmission.

Similar to the DL dynamic transmissions, when the gNB configures and activates the DL SPS resources by a DCI, the activation DCI contains PDSCH-to-HARQ-feedback indicator K1 indicating the number of slots/sub-slots from the end of PDSCH to the beginning of feedback on PUCCH. This parameter indicating the PDSCH-to-HARQ-feedback timing is then continually used throughout the active time of SPS configuration. If the UE receives PDSCH in slot/sub-slot n, the UE will transmit HARQ feedback in slot/sub-slot n+K1. However, in TDD configuration, if the slot/sub-slot n+K1 is a DL slot/sub-slot, the UE cannot transmit PUCCH carrying HARQ feedback in that slot/sub-slot and cancels that PUCCH transmission.

In 3GPP Release 15, the periodicity of DL SPS is bigger than 10 ms so DL SPS transmission does not happen frequently. However, in 3GPP Release 16, the periodicity may be set as low as 0.5 ms which is equivalent to 1 slot with SCS of 30kHz [8], thus making them suitable for URLLC transmission. Nevertheless, the value K1 is only indicated once at activation of DL SPS configuration. This is different from dynamic DL scheduling where the scheduling DCI will indicate the feedback timing for the scheduled transmission. Given a lower SPS resources' periodicity and a flexible format design in 5G where dynamic slot/sub-slot format update using slot-based DCI is supported to transmit low latency transmission [9], the collision probability between the UL resources of HARQ feedback of DL SPS PDSCH and DL/flexible slot/sub-slot in semi-static and dynamic TDD configuration increases that results in an increase of feedback cancellation. Therefore, when the DL SPS configurations with short periodicity are used for URLLC services/applications, the HARQ relevant conflicts can become a serious harm to URLLC QoS.

In Fig. 1, the DL SPS resources are configured to a UE with a periodicity of P in TDD configuration. The K1 value indicated by the activation DCI is 3. A PDSCH is sent to the UE in the first SPS resources. Based on K1 equal to 3, a HARQ feedback is expected to be transmitted in the fourth slot in Fig. 1 that is 3 slots after the first slot containing SPS PDSCH. However, this slot is a DL slot and cannot be used for an UL transmission of HARQ feedback. Following Release 15, this HARQ feedback is dropped without being resumed to be transmitted in the next PUCCH.

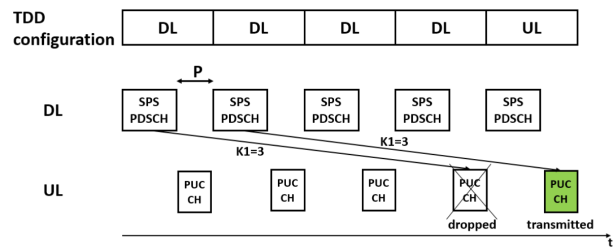

Fig. 1. HARQ feedback cancellation in DL SPS transmission in TDD.

If the UEs drop HARQ feedback due to UL-DL slot/sub-slot conflict as shown in Figure 1, the reliability of PDSCH degrades due to missing HARQ feedback and might not achieve the requirement because the gNB does not have the information to carry out the retransmissions in case of the first PDSCH's failure. On the other hand, if the gNB retransmits transport block (TB) in case of feedback's absence, it causes a poor spectrum utilization due to unnecessary retransmissions when the UE decodes correctly the first TB but an ACK feedback is dropped due to TDD conflict.

B. Prior art

In [5], two options are supported. The first one is to defer HARQ-ACK until the first available valid UL slot (or PUCCH resource) but only semi-static TDD configuration is considered. The second one is to indicating K1 value for each SPS transmission but no scheme is mentioned.

In [10] and [11], two options are proposed to ensure the transmission of HARQ feedback. In the first option, multiple K1 values are indicated for one SPS configuration to guarantee SPS HARQ feedback to be pointed to an UL slot/sub-slot. However, the activation DCI has to carry multiple values of K1 so it causes an increase of overhead and a decrease of PDCCH reliability. In the second option, the UE postpones the collided SPS HARQ-ACK to the UL slot/sub-slot available for HARQ transmission. The UL slots/sub-slots are the semi-static UL slots/sub-slots that the gNB and the UE have an aligned understanding. This option is also proposed in [12]. However, this option is only for semi-static TDD configuration and does not discuss the UE behaviour in dynamic TDD configuration where the flexible slots/sub-slots can be updated to the UL slots/sub-slot by slot format indication (SFI).

In [13], it is argued that deferring HARQ feedback to the first available UL slot/sub-slot is a simple option but causes the imbalanced HARQ feedback because many HARQ feedback postponed in the previous slots might be transmitted in the same UL slot/sub-slot. Therefore, they propose that K1 is indicated to the UE in each DL SPS transmission. However, no scheme of K1 indication is proposed.

In [14], two options are proposed. The first one is to defer HARQ feedback to the next available UL slot/sub-slot within a higher-layer configured time window. There is no discussion about using the flexible slots/sub-slots switched to the UL slots/sub-slots by SFI. In the second option, HARQ slots/sub-slots for DL SPS are reserved periodically. When the HARQ occasion is not available UL occasion, HARQ feedback is postponed to the latest periodically configured HARQ slot/sub-slot. Nevertheless, it consumes much resources for the reserved slots/sub-slots.

III. Enhancements for HARQ feedback in DL SPS transmission in TDD

A. Dynamically indication of K1 value for each DL SPS transmission

As discussed above, PDSCH-to-HARQ-feedback timing indicator is indicated to the UE by the activation DCI for DL SPS configuration transmitted from the gNB. The bits in the PDSCH-to-HARQ feedback timing indicator field specify a row index in the table configured by RRC parameter: dl-DatatoUL-ACK in PUCCH-Config. Based on the index from the activation DCI, the UE knows the value of K1 in the row corresponding to that index. In DCI format 1_0, 3 bits in the PDSCH-to-HARQ feedback timing indicator are mapped to K1 values from 1 to 8. In DCI format 1_1, there are from 0 bit to 3 bits used to indicate the row index of K1 from the mapping table. The number of bits is determined by $log_2 I$ where I is the number of rows in the table.

The value of K1 stays fixed throughout the SPS active time once the gNB activates DL SPS resources. DL SPS PDSCH is transmitted without an associated PDCCH so the gNB cannot change K1 at the time of PDSCH transmission to avoid the UL-DL slot/sub-slot conflict.

To overcome such shortcoming in DL SPS feedback mechanism, the SPS-PDSCH-to-HARQ-feedback timing K1 is indicated dynamically with each SPS transmission. In this case, the activation DCI may include a default value which may then be overridden by a later specific value. To avoid confusion, this initial value of K1 can be even removed from the activation DCI. Thus, a non-numerical value of K1 is used in the activation DCI. It is a signal to tell the UE that K1 value is indicated dynamically so it should expect and decode a specific numerical K1 value in the next data transmission. A row in the table specified in the PUCCH-Config.dl-DatatoUL-ACK has a non-numerical value. When the gNB chooses to transmit K1 dynamically in each SPS transmission, it sends the activation DCI with the sequence in the PDSCH-to-HARQ feedback timing indicator mapped to the row containing a non-numerical value. For example, as shown in Table I, the gNB transmits a sequence 111 as PDSCH-to-HARQ-feedback timing indicator in the activation DCI to indicate a non-numerical value of K1 which is interpreted by the UE to instruct it to wait for a specific K1 value for each transmission. With the other sequences, the UE interprets the K1 value appropriately and applies that value for the relevant DL SPS resources.

TABLE I
Mapping table of PDSCH-to-HARQ-fedback timing indicator with a non-numerical K1 value

Row index	**PDSCH-to-HARQ-fedback timing indicator**	**K1 value**
0	000	1
1	001	2
2	010	3
3	011	4
4	100	5
5	101	6
6	110	7
7	111	FFFF

Because there is no PDCCH associated with SPS PDSCH, dynamic K1 value utilized for a specific transmission is specified by PDSCH itself. Some bits (1-3 bits) in PDSCH are used to indicate K1 value. These bits are mapped to K1 value in a table as when K1 is indicated by DCI. Thus, the SPS-PDSCH-to-feedback-timing indication is transmitted with data on the PDSCH where the data packet is added the bits of K1 indication then they are encoded as one TB to be transmitted on the PDSCH. The embedded K1-indication bits exploit cyclic redundancy check (CRC) and redundant bits of the encoded data packet so the reliability of K1 detection is guaranteed without an increase of the separate CRC and redundant bits for K1 indication.

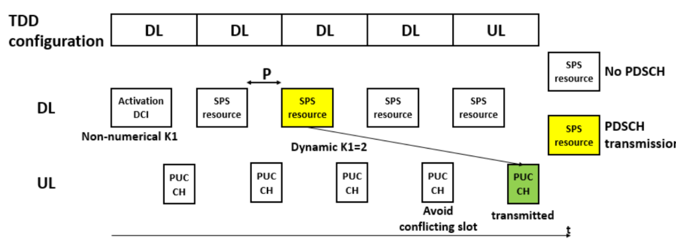

Fig. 2. Dynamic K1 indication in DL SPS transmission.

The operation of DL SPS transmission with dynamic K1 indication is shown in Fig. 2. At the beginning, an activation DCI specifying a non-numerical K1 is transmitted from the gNB to the UE to active DL SPS configuration. When the gNB transmits PDSCH in DL SPS resource, it also adds dynamic K1 indication in PDSCH to indicate K1 value. In this case, with the latest information of slot configuration, the gNB chooses K1 to be 2 in order to avoid a DL slot and points the feedback to the closest available UL slot.

In TDD configuration, if the UE receives several SPS PDSCHs in the consecutive DL slots/sub-slots, the UE might have to transmit all the feedback of the previous SPS PDSCHs in the same PUCCH resource. However, multiplexing many feedback in the same PUCCH resource degrades reliability of feedback. Therefore, based on the size of PUCCH resources and the payload of feedback, the gNB should choose the dynamic values of K1 to avoid the multiplexing of feedback surpassing the capacity of PUCCH resources. As in Fig. 3, the feedback of SPS PDSCH 3 and 4 is dynamically indicated to be transmitted in PUCCH2 instead of PUCCH1 - the closest UL slot after the DL slots to avoid an overload in PUCCH1.

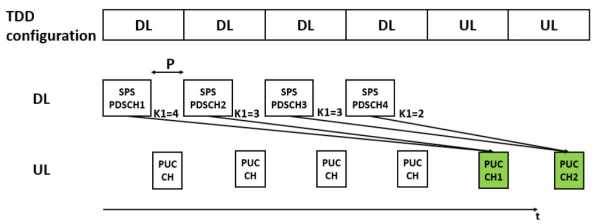

Fig. 3. Dynamic K1 indication in the multiplexing of UL feedback for DL SPS transmission.

B. ACK-only feedback structure

A UE is required to transmits HARQ feedback for every DL SPS occasion. Even if there is no TB transmitted in a DL SPS occasion, the UE still transmits negative acknowledgement (NACK) feedback to the gNB. However, this scheme results in very high overhead with shorter periodicity of DL SPS occasions. A further drawback arises when K1 is specified dynamically in PDSCH since the UE does not have a K1 value for an occasion with no transmission. To reduce the feedback burden and support dynamic K1 indication, an ACK-only feedback structure is used. The UE is configured to transmit only the ACK feedback if it decodes correctly PDSCH in the DL SPS occasions and obtains K1 value in PDSCH. In all other cases, no feedback transmission is sent. The gNB is aware of which occasions it has made a transmission in and the K1 values set for those transmissions it makes. The gNB is thus aware when feedback should be received, and if no feedback is received at that time, can assume the transmission failed and make a retransmission.

The operation of ACK-only feedback structure in the dynamic-K1-indication-by-PDSCH scheme is shown in Table II.

ACK-only feedback structure is used when K1 is indicated dynamically. Dynamic-K1-indication scheme is activated if a non-numerical K1 value is indicated in the activation DCI. Therefore, a non-numerical K1 value in the activation DCI is interpreted by the UE as an indication to activate ACK-only feedback structure in DL SPS transmission.

IV. Numerical results

The first simulation is done to compare the reliability of the transmissions with an initial PDSCH and a potential retransmission in two schemes. In the conventional scheme, the feedback is cancelled if it is scheduled in a DL slot/sub-slot of TDD configuration. Therefore, if the first PDSCH is not decoded correctly by the UE, a retransmission is not done and the packet is lost. In the proposed scheme, the feedback's resources is indicated to avoid DL-UL slot/sub-slot conflict and an ACK-only feedback structure is used so a retransmission is carried out if necessary. A packet of 160 bits

TABLE II
ACK-only feedback structure in DL SPS transmission

Case	UE behavior	gNB behavior	Note
No TB transmitted	No feedback	No action	No K1 is dynamically indicated to the UE
TB transmitted but not decoded	No feedback	Retransmit TB in SPS or dynamic resources (PDSCH associated with PDCCH)	K1 value is dynamically indicated to the UE in PDSCH but the gNB does not receive any feedback
TB transmitted and decoded correctly but ACK not decodable at the base station	ACK transmission	Retransmit TB in SPS or dynamic resources (PDSCH associated with PDCCH)	The UE transmits an ACK but the gNB fails to decode (due to bad channel conditions, etc). The gNB will retransmit but the UE will detect retransmission due to new data indicator (NDI) setting
TB transmitted and decoded correctly	ACK transmission	No action	K1 value is dynamically indicated to the UE in PDSCH, decoded correctly and used to determine the UL slot for ACK transmission

encoded by low-density parity-check (LDPC) code and modulated in quadrature phase shift keying (QPSK) is transmitted with MCS2 in the first transmission and with MCS1 in the retransmission following the failure of the first transmission in additive white Gaussian noise (AWGN) channel.

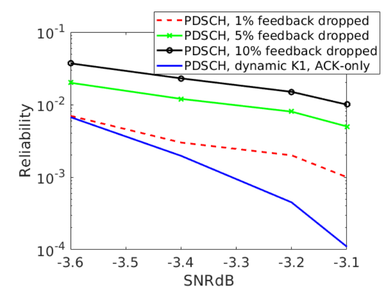

Fig. 4. Performance of PDSCH with retransmission ensured in the proposed scheme and PDSCH with potential feedback cancellation.

In the conventional scheme, the percentages of the indicated resources for feedback in DL slot/sub-slot of TDD configuration are 1%, 5% and 10% corresponding to the rate of feedback cancellation. As be shown in Fig. 4, because of feedback cancellation, no retransmission is carried out in case of the first PDSCH's failure so PDSCH transmission's error (packet loss) increases remarkably. In contrast, PDSCH transmission with the proposed scheme achieves a much lower error rate. The benefits of the proposed scheme become more important with a high probability of slot/sub-slot conflict.

The use of ACK-only feedback structure not only makes dynamic K1 indication feasible but also reduces significantly feedback overhead. Fig. 5 shows the percentage of feedback reduction of the ACK-only feedback structure compared to the conventional feedback structure. Two cases where 1% and 10% SPS resources are used for PDSCH transmission are

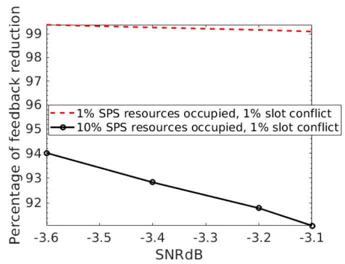

Fig. 5. The reduction of feedback in ACK-only feedback structure.

considered. The number of feedback is counted after the first transmissions of the packet with MCS2 are done.

The second simulation is carried out to compare the performance (measured by block error rate (BLER)) of the conventional PDSCH without dynamic K1 indication embedded and the proposed PDSCH with dynamic K1 indication embedded in Section III-A. In the first scenario of PDSCH without K1 indication, a packet of 160 bits is encoded by LDPC code and transmitted on PDSCH. In the second scenario of PDSCH with K1 indication, a packet of 160 bits is added 3 indication bits to create a TB with 163 bits. The TB is encoded by LDPC code and transmitted on PDSCH. In both scenarios, modulation technique is QPSK and channel is AWGN.

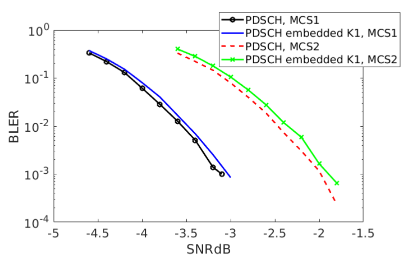

Fig. 6. Performance of normal PDSCH and PDSCH embedded dynamic K1 indication.

Fig. 6 shows the performance of PDSCH with MCS1 and MCS2. In the both two MCSs, a decrease of K1-embedded PDSCH's reliability is insignificant compared to a full PDSCH's reliability. The difference is negligible. Therefore, an use of dynamic K1 indication multiplexing with PDSCH does not affect the performance of PDSCH.

The reliability of K1 detection is also the reliability of K1-embedded PDSCH as shown in Fig. 6. The scheme with PDSCH embedded with K1-indication bits makes K1 indication's detection achieve high reliability as PDSCH without an increase of separate CRC and redundant bits to encode the indication bits. The successful detection of K1 avoids an unnecessary PDSCH transmission and helps reduce resource consumption as well as increase spectrum utilization.

Fig. 7 compares BLER of a standard activation DCI containing one K1 (3 bits) with 24 data bits and 24 CRC bits used in the proposed method and BLER of an activation DCI containing multiple K1 with three bits for each K1 as proposed in [10] and [11]. 384 resource elements are used to transmit the DCI encoded by Polar code and modulated in QPSK in AWGN channel. The UE decoder is min-sum Successive cancellation list (SCL) decoder with list size 8.

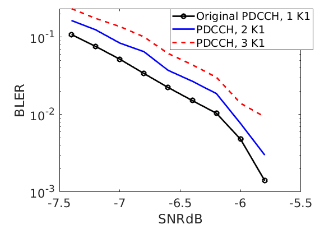

Fig. 7. Performance of PDCCH with multiple K1 value.

The more K1 values DCI contains, the higher BLER PDCCH suffers from that makes the URLLC transmission unable to achieve the strict requirements and low spectrum efficiency.

Moreover, the method with multiple K1 indications in DCI of [10] and [11] only reduces the probability of the feedback conflict but cannot guarantee the transmission of feedback and the data retransmission. A set of K1 values cannot cover all the possibilities of UL slots/sub-slots' positions, especially if the slot format is updated periodically by DCI.

V. Conclusion

The work enhances HARQ feedback transmission for DL SPS PDSCH transmission. A scheme consisting of dynamic K1 indication by PDSCH and ACK-only feedback structure is presented. Dynamic K1 indication embedded in PDSCH guarantees high reliability of K1 detection without an increase of CRC and redundant bits as well as ensures the feedback transmission. ACK-only feedback structure assures the operation of dynamic K1 indication, reduces feedback overhead and ensures the necessary retransmission.

Acknowledgments

The work has been partially supported by TCL and the H2020 project 5GENESIS (5genesis.eu).

References

[1] 3GPP TR 38.913 v15.0.0, "Study on scenarios and requirements for next generation access technologies."
[2] 3GPP TR 38.802 v14.2.0, "Study on New Radio Access Technology Physical Layer Aspects."
[3] Huawei, HiSilicon, Nokia, Nokia Shanghai Bell, "New SID on Physical Layer Enhancements for NR URLLC". 3GPP RP-182089, TSG-RAN#81, Gold Coast, Australia, Sept 10–13, 2018.
[4] 3GPP TR 38.211 v15.7.0, "Physical channels and modulation."
[5] Chairman's notes RAN1 #98b, Chongquing, China, October 14–20,2019.
[6] 3GPP TS 38.214 v15.7.0, "Physical layer procedures for data."
[7] 3GPP TS 38.331 v15.7.0, "Radio Resource Control (RRC) protocol specification."
[8] Chairman's notes RAN1 #98, Pargue, CZ, August 26–30,2019.
[9] 3GPP TS 38.213 v15.7.0, "Physical layer procedures for control."
[10] Huawei, HiSilicon, SIA,"Other aspects for URLLC/IIOT enhancements", 3GPP R1-1909317, Prague, Czech Republic, August 26–30, 2019.
[11] WILUS Inc.,"On SPS PDSCH for NR URLLC", 3GPP R1-1911320, Chongqing, China, October 14– 20, 2019.
[12] Vivo,"Other issues for URLLC", 3GPP R1-1908164, Prague, Czech Republic, August 26–30, 2019.
[13] ZTE,"Other enhancements for Rel-16 URLLC", 3GPP R1-1910106, Chongqing, China, October 14– 20, 2019.
[14] CMCC,"Discussion on DL SPS enhancements", 3GPP R1- 1910169, Chongqing, China, October 14– 20, 2019.