DOI: 10.19729/j.cnki.1673-5188.2018.03.007
http://kns.cnki.net/kcms/detail/34.1294.TN.20180710.1529.002.html, published online July 10, 2018

Mechanism of Fast Data Retransmission in CU-DU Split Architecture of 5G NR

HUANG He, LIU Yang, LIU Zhuang, HAN Jiren, and GAO Yin
(ZTE R&D Center, Shanghai 201203, China)

Abstract

The 5G radio access network (RAN) architecture is supposed to be split into the central unit (CU) and the distributed unit (DU) in order to support more flexible transport networks and provide enhanced user experience. However, such functional split may also introduce some new technical issues. In this paper, we study the data fast retransmission issue introduced by this functional split in different scenarios and solutions are provided to handle this issue. With the fast data retransmission mechanism proposed in this paper, the retransmitted data packets could be identified and handled with high priority. In this way, the data delivery between the CU and DU in 5G RAN is assured.

Keywords

5G RAN; central unit; distributed unit; fast retransmission

1 Introduction

The 5G RAN architecture is supposed to be split into the central unit (CU) and the distributed unit (DU) in order to support various types of transport networks and multi-vendor requirement. The latency-tolerant network function resides in the CU entity and the latency-sensitive network function resides in the DU entity [1].

For the higher layer split (HLS) solution, there are eight possible CU-DU split options as shown in [2]. According to the agreements from the 3GPP RAN3#95bis meeting [3], the 3rd Generation Partnership Project (3GPP) has decided to select Option 2 (based on Packet Data Convergence Protocol (PDCP) and decentralized Radio Link Control (RLC)) as HLS solution for normative work in Release 15. As for the lower layer split (LLS) solution, the possible solutions can be found in [4] and the corresponding discussions are still going on. Additionally, the gNB-CU can be further split into the control plane and user plane [5] in order to support more flexible data services. In this paper, we focus on HLS only.

The interface between CU and DU is called F1 interface [6]. One CU can connect to multiple DUs and one DU can support one or more cells. The general principles can be found in [7] and the application protocol for the F1 interface can be found in [8]. Considering the CU-DU split scenario, one newly raised issue is that the data transmission may suffer outage due to different reasons, such as handover and temporary Radio Link Failure (RLF). For example, when user equipment (UE) performs handover from the source DU to the target DU, the packets which are still buffered in the source DU will be lost. Additionally, the air interface between UE and DU may become unstable and the packets which are already in the air may not be confirmed by the UE. Therefore, the DU triggered fast data retransmission needs to be investigated.

The lost PDCP protocol data units (PDUs) need to be identified and retransmitted with high priority in order to facilitate the fast transmission for CU-DU split scenario. The network slicing based handover procedures and mobility management mechanisms are discussed in [9], but the lost data packets during the handover procedure is not mentioned. The Fog Radio Access Network architecture is introduced in [10] and it is claimed that the associated mobility and resource management mechanisms can reduce the signaling cost. However, the lost data packets scenario is not covered. In the following, we will first describe the typical scenarios for this issue and the corresponding solutions for fast retransmission of the lost PDCP PDUs are also provided. In the following, we denote CU/DU as gNB-CU/gNB-DU to make them aligned with the 3GPP specifications.

2 Scenario Description

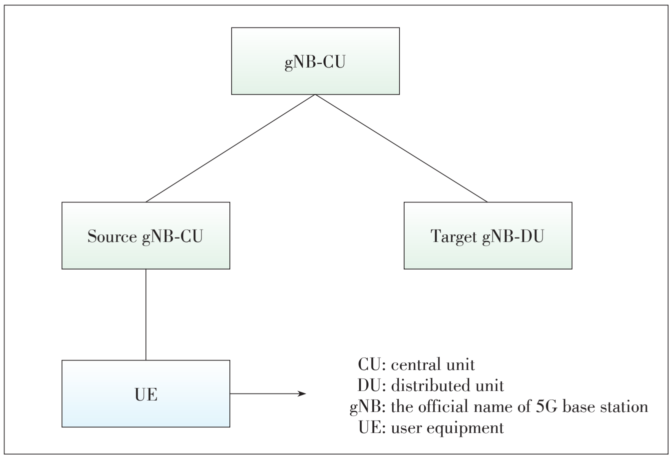
In this section, we will illustrate the PDCP PDU retransmission issue in three typical scenarios: single connectivity, multi-connectivity, and E-UTRAN-NR dual connectivity (EN-DC).

2.1 Single Connectivity Scenario

Single connectivity means that UE is connected to only one gNB-DU at a certain time instance. **Fig. 1** shows the downlink transmission of the intra-gNB-CU inter-gNB-DU handover scenario. When UE moves from the source gNB-DU to the target gNB-DU, one common situation is that there are still remaining data packets buffered in the source gNB-DU waiting for transmission. Since gNB-CU is unaware of the delivery status for these data packets, the current mechanism in LTE cannot assure the fast retransmission of these remaining packets in the source gNB-DU and these data packets will be lost.

2.2 Multi-Connectivity Scenario

The multi-connectivity indicates that UE is served by at



▲Figure 1. Single connectivity scenario.

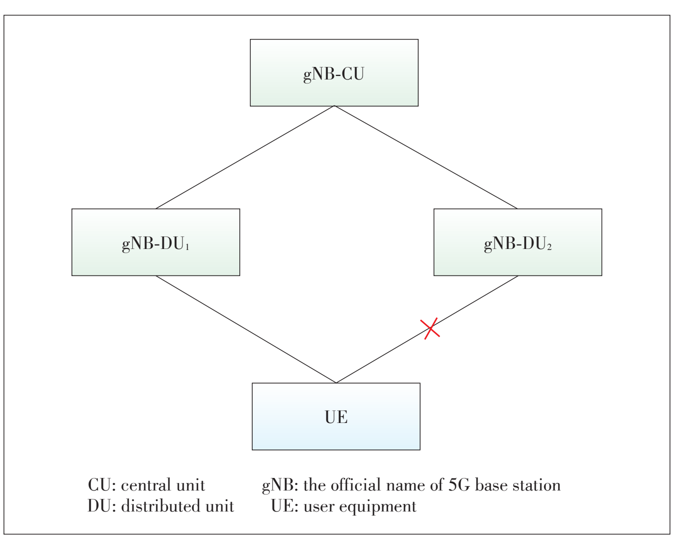
least two gNB-DUs simultaneously. This scenario focuses on the fast retransmission of data packets, during which UE is served by one of the gNB-DUs with a radio link that is subject to RLF. Such a scenario is typically encountered with radio links at high frequencies.

Fig. 2 shows the scenario where UE connects to the gNB-CU via two gNB-DUs simultaneously. The data packets are delivered from gNB-CU to the UE via gNB-DU 1 and gNB-DU 2. At a certain time point, the radio link of gNB-DU 2 encounters RLF and becomes unavailable, thus all the data packets that has not been successfully delivered to the UE via gNB-DU 2 should be retransmitted to the UE via gNB-DU 1. Once the radio link of gNB-DU 2 becomes available again, it is expected that data traffic can be centrally forwarded to the gNB-DU 2 hosting the previously broken link. The traffic transmission in gNB-DU 2 will resume as it was before the RLF.

2.3 EN-DC Scenario

Fast data retransmission in the EN-DC scenario can be considered as another typical case, which indicates that UE may be configured to utilize radio resources provided by two distinct schedulers in two different nodes connected via non-ideal backhaul, one providing E-UTRAN access (eNB) and the other providing NR access (gNB).

As shown in **Fig. 3**, we take the inter-gNB-DU mobility using Master Cell Group (MCG) Signalling Radio Bearer (SRB) as an example, i.e., UE moves from one gNB-DU to another gNB-DU within the same gNB-CU when MCG SRB is available during EN-DC operation.



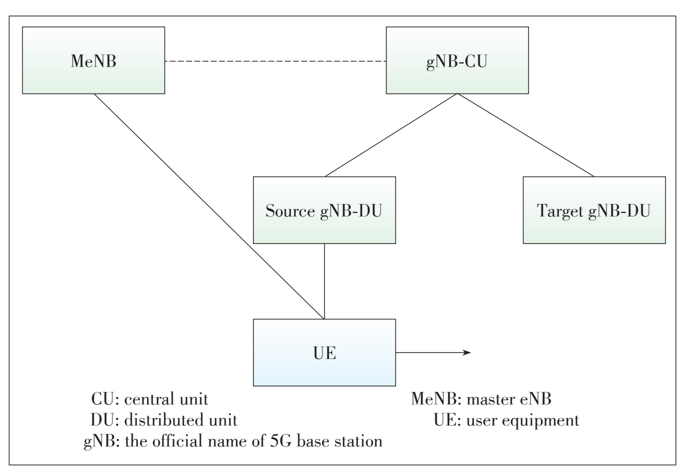
▲Figure 2. Multi-connectivity scenario.

3 Solutions

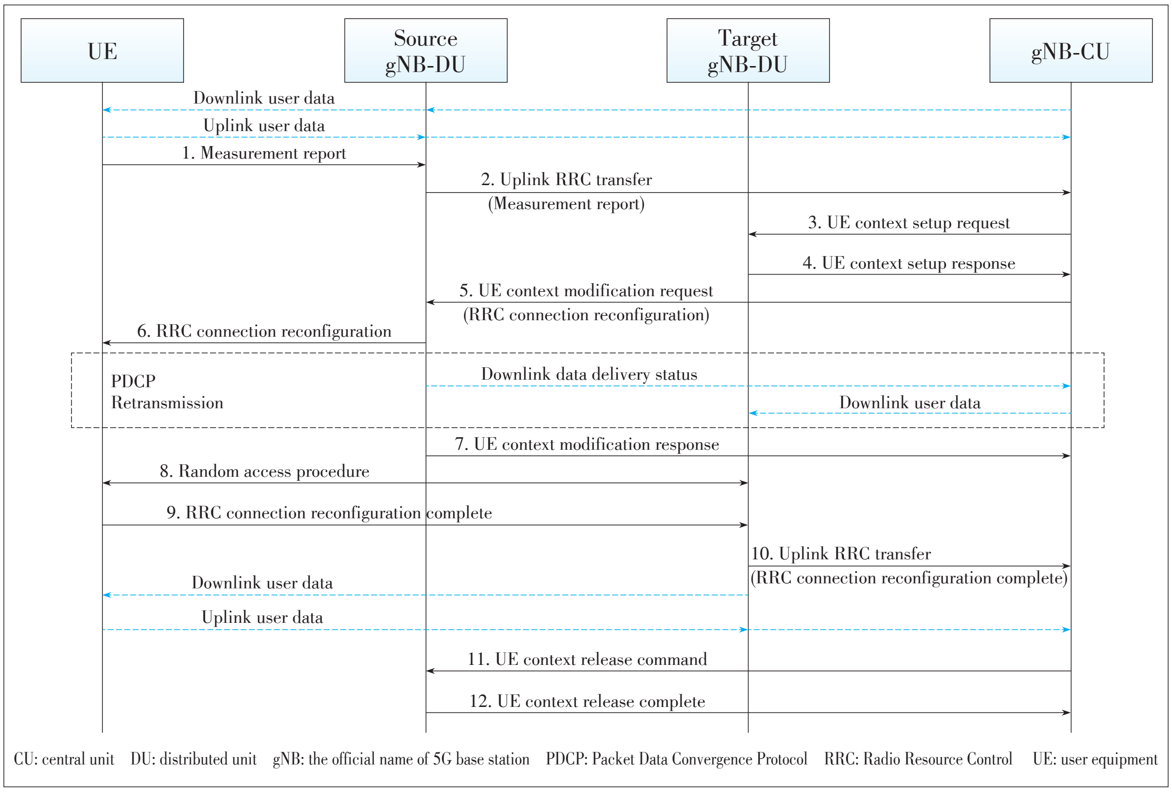
In this section, we will demonstrate that the gNB with CU-DU architecture can be enhanced to enable fast retransmission of PDCP PDUs in a centralized way. The enhanced fast retransmission solutions for above scenarios will be illustrated independently.

3.1 Solution for Single Connectivity Scenario

As shown in **Fig. 4**, the Intra-gNB-DU mobility procedure [5] can be used to demonstrate the retransmission mechanism for single connectivity scenario. It can be seen that at Steps 5 and 6, the gNB-CU sends a UE Context Modification Request message to the source gNB-DU, which includes a generated Radio Resource Control (RRC) Connection Reconfiguration message and indicates to stop the data transmission for UE. The source gNB-DU could respond with a Downlink Data Delivery Status (DDDS) frame via F1-U to inform the gNB-CU about the highest PDCP PDU SN successfully delivered in sequence to UE and the highest NR PDCP PDU sequence number transmitted to the lower layers, which is the key step for the fast retransmission of lost PDCP PDUs (shown in the dashed rectangle in Fig. 4). Based on this information, the gNB-CU is able to identify the unsuccessfully transmitted PDCP PDUs in the source gNB-DU side. Those unsuccessfully delivered PDCP PDUs will be sent from the gNB-CU to the target



▲Figure 3. E-UTRAN new radio-dual connectivity (EN-DC) scenario.



◀Figure 4. Procedure of fast data retransmission in single connectivity scenario (Based on Figure 8.2.1.1-1 shown in TS38.401 [6]).

gNB-DU for retransmission.

3.2 Solution for Multi-Connectivity Scenario

For the multi-connectivity scenario, the centralized retransmission procedure in intra-gNB-CU [5] is shown in **Fig. 5**. This mechanism allows to perform the retransmission of PDCP PDUs that are not delivered by a gNB-DU (gNB-DU 1) because the corresponding radio link toward the UE is subject to outage.

As shown in Fig. 5, UE is receiving data from gNB-DU 1 and gNB-DU 2 simultaneously. At a certain time, gNB-DU 1 realizes that the radio link towards UE is experiencing outage and sends the "Radio Link Outage" notification to the gNB-CU over the F1-U, which is a part of the DDDS frame of the concerned data radio bearer. The message includes the highest PDCP PDU Sequence Number (SN) successfully delivered in sequence to UE and the highest NR PDCP PDU sequence number transmitted to the lower layers in gNB-DU 1. Based on the received information, gNB-CU will retransmit the potentially undelivered PDCP PDUs and the new PDUs via gNB-DU 2. If gNB-DU 1 realizes that the radio link is back in normal operation, it may send a "Radio Link Resume" notification to inform the gNB-CU that the radio link can be used again. Then the gNB-CU may start sending traffic via gNB-DU 1 again.

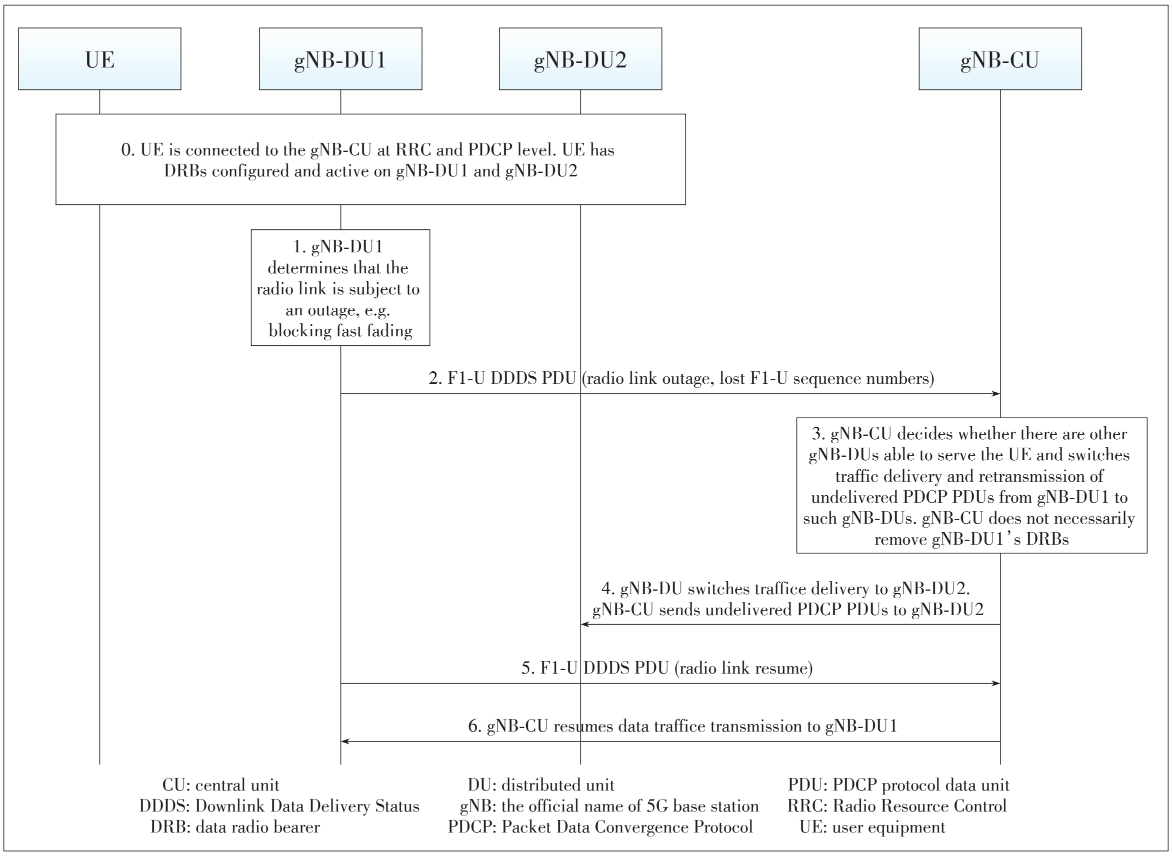
3.3 EN-DC Scenario

This procedure is used for the case where the UE moves from one gNB-DU to another gNB-DU within the same gNB-CU when MCG SRB is available during the EN-DC operation. **Fig. 6** shows the inter-gNB-DU mobility procedure using MCG SRB in EN-DC.

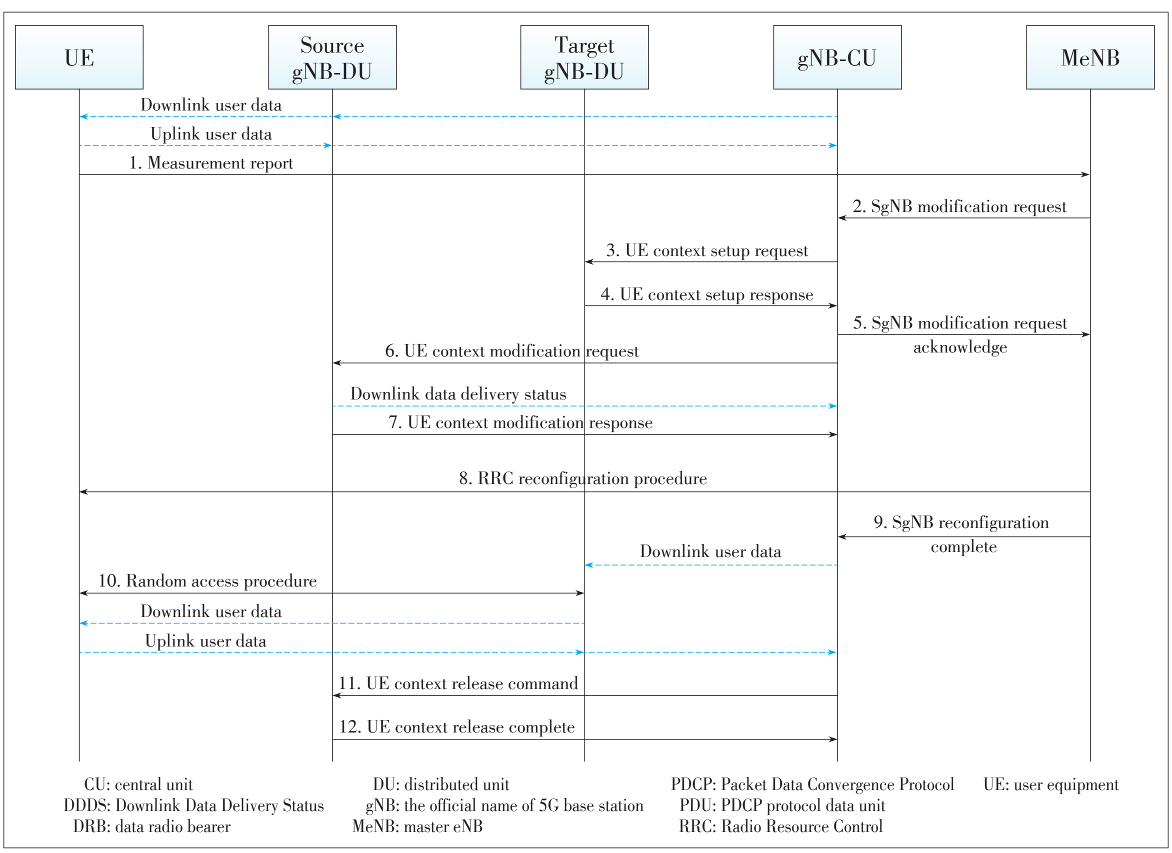
As shown in Fig. 6, UE sends a Measurement Report message to master eNB (MeNB). Then MeNB performs the SgNB Modification procedure and the UE context in the target gNB-DU is created. In the following, the gNB-CU sends a UE Context Modification Request message to the source gNB-DU indicating to stop the data transmission to UE. Correspondingly, the source gNB-DU replies a Downlink Data Delivery Status frame to inform the gNB-CU about the unsuccessfully transmitted downlink data to UE or lower layers. After UE performs RRC Reconfiguration procedure with MeNB, the MeNB sends an SgNB Reconfiguration Complete message to the gNB-CU. Thus the unsuccessfully transmitted PDCP PDUs in the source gNB-DU are sent from the gNB-CU to the target gNB-DU. After the random access procedure is performed between UE and the target gNB-DU, the downlink packets will be sent to the UE.

3.4 Handling Retransmitted Packets with High Priority

The retransmission mechanism still has a problem. The only way for the gNB-CU to acquire the packet delivery status to UE is DDDS [11]. However, DDDS only reports "the highest delivered/transmitted PDCP SN" to gNB-CU, thus the gNB-CU cannot know whether the retransmitted packets are delivered or not. In case the PDCP SN of retransmitted packets is lower than the one already buffered in the gNB-DU, the gNB-CU



◀Figure 5. Procedure for fast data retransmission in the multi-connectivity scenario (Based on Figure 8.3.1-1 in TS38.401 [6]).



◀Figure 6. Procedure for fast data retransmission in the EN-DC scenario (Based on Figure 8.2.2.1-1 in TS38.401 [6]).

may clean the buffer and the retransmitted packets could not be delivered.

In order to facilitate the data transmission procedure, retransmitted data packets need to be identified and handled with high priority at the gNB-DU side. One straightforward way is to introduce an indication in the retransmitted packet. To be specific, a "retransmission flag" is introduced in the spare bit of DL_USER_DATA. The detailed illustration for the DL_USER_DATA frame structure can be found in [12]. When the gNB-CU performs retransmission, it indicates to the gNB-DU whether the packets are retransmitted or not by the value of this flag, which can help the gNB-DU to identify retransmitted packets easily. At the gNB-DU side, once the retransmitted data packets are identified, the gNB-DU will handle the corresponding data frame separately. For example, besides the normal transmission queue, gNB-DU will additionally provide a retransmission queue for the retransmitted packets in order to guarantee the scheduling of retransmitted data with high priority.

4 Conclusions

In this paper, we study the fast data retransmission issue introduced by the functional split of the 5G RAN architecture. Three typical scenarios (the single connectivity scenario, multi-connectivity scenario, and EN-DC scenario) are respectively described in order to illustrate the PDCP PDU data retransmission issue. We also provide the solutions targeting the above scenarios, which have already been agreed and captured by the 3GPP specifications [6]. With the data retransmission mechanism described in this paper, the retransmitted data packets could be identified and handled with high priority. Thus the data delivery between the gNB-CU and the gNB-DU in 5G RAN is assured.

References

[1] NTT DOCOMO, INC., "Revised WID on new radio access technology," 3GPP RP-172109, 2017.
[2] *Study on New Radio Access Technology: Radio Access Architecture and Interfaces*, 3GPP TR38.801, 2016.
[3] I. Toufik, "Report of 3GPP RAN WG3 #95bis," R3-171412, Spokane, USA, Apr. 2017.
[4] *Study of CU-DU Low Layer Split for NR (Release 15)*, 3GPP TS 38.816, 2018.
[5] *Study of Separation of NR Control Plane (CP) and User Plane (UP) for Split Option 2 (Release 15)*, 3GPP TS 38.806, 2018.
[6] *NG-RAN, Architecture Description (Release 15)*, 3GPP TS 38.401, 2018.
[7] *NG-RAN, F1 General Aspects and Principles (Release 15)*, 3GPP TS 38.470, 2018.
[8] NG-RAN, *F1 Application Protocol (Release 15)*, 3GPP TS 38.473, 2018.
[9] H. J. Zhang, N. Liu, X. L. Chu, et al., "Network slicing based 5G and future mobile networks: mobility, resource management, and challenges," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 138–145, Aug. 2017. doi: 10.1109/MCOM.2017.1600940.
[10] H. J. Zhang, Y. Qiu, X. L. Chu, K. P. Long, and V. C. M. Leung, "Fog radio access networks: mobility management, interference mitigation and resource optimization," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 120–127 Dec. 2017. doi: 10.1109/MWC.2017.1700007.
[11] *NG-RAN, NR User Plane Protocol (Release 15)*, 3GPP TS38.425, 2018.
[12] ZTE Corporation, "R3-180131 further discussion on data retransmission indication," 3GPP R3-180610, 2018.

Manuscript received: 2018-02-05

Biographies

HUANG He (huang.he4@zte.com.cn) received the bachelor's degree in computer science and technology from Shanghai Jiao Tong University, China in 2004. He is currently the chief engineer of wireless innovation laboratory of ZTE Corporation and leads the research and standardization work on 5G RAN. He has filed more than 60 patents. He was the rapporteurs of multiple CCSA/3GPP SIs/WIs and the editors of the related protocols.

LIU Yang (liu.yang31@zte.com.cn) received the Ph.D. degree in communication and information systems from Beijing University of Posts and Telecommunications (BUPT), China in 2016. He was a visiting scholar at Department of Electrical and Computer Engineering of North Carolina State University, USA from 2013 to 2015. He is currently a 5G research engineer at ZTE R&D center, Shanghai. His research interests include 5G wireless communications and signal processing.

LIU Zhuang (liu.zhuang2@zte.com.cn) received the master's degree in computer science from Xidian University, China in 2003. He is currently a senior 5G research engineer at ZTE R&D center, Shanghai. His research interests include 5G wireless communications and signal processing.

HAN Jiren (han.jiren@zte.com.cn) received the master's degree in wireless communication systems from University of Sheffield, UK in 2016. He is an advanced research engineer at the Algorithm Department, ZTE Corporation. His research interests include 5G wireless communications and signal processing.

GAO Yin (gao.yin1@zte.com.cn) received the master's degree in circuit and system from Xidian University, China in 2005. Since 2005 she has been with the research center of ZTE Corporation and been engaged in the study of 3G/4G/5G technology. She has authored or co-authored about hundreds of proposals for 3GPP meetings and journal papers in wireless communications and has filed more than 100 patents. She was the rapporteurs of multiple 3GPP WIs. From August 2017, she has been elected as the vice chairman of 3GPP RAN3.