Blockage-Aware Flow Control in E-UTRA-NR Dual Connectivity for QoS Enhancement

JUNGJUN KIM¹,², (Graduate Student Member, IEEE), AND SAEWOONG BAHK¹,², (Senior Member, IEEE)
¹Department of Electrical and Computer Engineering (ECE), Seoul National University, Seoul 08826, South Korea
²Institute of New Media and Communications (INMC), Seoul National University, Seoul 08826, South Korea

Corresponding author: Saewoong Bahk (e-mail: sbahk@snu.ac.kr).

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2021-0-02048) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation).

ABSTRACT

Millimeter-wave (mmWave) communication included in new radio (NR) provides high throughput to meet the stringent requirements of modern applications. However, its performance is highly dependent on whether obstacles are blocking line-of-sight (LOS) paths. Blockages on the LOS paths jeopardize reliable connections. E-UTRA-NR dual connectivity (EN-DC) is one option for managing unstable connectivity using robust long-term evolution (LTE) as well as mmWave NR. EN-DC aggregates and utilizes multiple links to provide a better quality of service (QoS) to users. However, the performance of EN-DC is strongly affected by flow control due to the possibility of out-of-order packet delivery at the packet data convergence protocol (PDCP) layer. In this paper, we propose a blockage-aware flow control scheme (BAFC) that appropriately responds to the dynamics of mmWave links in the presence of mobile blockages and improves users’ QoS. To this end, we also develop PDCP scheduling schemes that satisfy various users’ QoS requirements. Through extensive simulations, we verify that BAFC exploits over 99% of the sum capacity of links on average under rapidly changing channels due to mobile blockages. Moreover, BAFC combined with PDCP scheduling achieves higher throughput or reduces latency by up to 43.9% with minimal throughput degradation.

INDEX TERMS

Blockage, dual connectivity, EN-DC, flow control, millimeter wave, 5G

I. INTRODUCTION

The demand for higher throughput in wireless communications is increasing due to the development and prevalence of applications such as virtual reality (VR), augmented reality (AR), and 4K video streaming. Moreover, advanced applications such as mixed reality (MR), hologram streaming, and 3D precision maps for autonomous driving will appear in the near future. These applications not only require much higher data rates but also have other stringent quality of service (QoS) requirements compared to conventional applications. For example, full-view advanced VR requires 11.94 Gbps data rate, with a typical round-trip time of 5 ms [1]. Such requirements cannot be satisfied using the only sub-6 GHz band due to the lack of radio resources.

Accordingly, the 3rd generation partnership project (3GPP) includes the millimeter-wave (mmWave) band in fifth generation (5G) mobile communication systems. Due to the abundance of available resources, 5G communication systems operating in the mmWave band can provide the throughput of multi-gigabits per second [2]. Although the mmWave band enables high throughput wireless communications, it suffers from inherent challenges due to the high frequency characteristics of 30–300 GHz. MmWave band has a large propagation attenuation with distance and a low degree of diffraction, making it vulnerable to blockage [3].

The effect of blockage on mmWave band communications has been investigated in several studies in the literature. The authors in [4] measure signal attenuation due to pedestrian blocking at the 28 GHz band, showing the degree of attenuation varying from 6.3 dB to 15.6 dB. In [5], when a line-of-sight (LOS) path is blocked by a person, the corresponding channel suffers from 7.6 dB to 17.5 dB loss depending on the
beamwidth. Measurements are also performed at mmWave frequencies other than 28 GHz. Shadowing by a pedestrian in the 60 GHz band or 73 GHz band has an effect (i.e., 5 – 18 dB) similar to that in the 28 GHz band [6], [7]. The effect of blockage by vehicles is studied in [8]–[10]. The degree of shadowing depends on measurement scenarios, and the peak attenuation is around 20 dB, as shown in [8]–[10]. The blockage effect by pedestrians or vehicles can lead to reduced throughput as well as connection failures.

To cope with possible connection failures at the mmWave band, multi-radio dual connectivity (MR-DC) offered by the 3GPP [11] is a promising solution. Among various MR-DC configurations, we pay attention to the use of evolved universal terrestrial radio access (E-UTRA) and new radio (NR) together as E-UTRA-NR dual connectivity (EN-DC). In EN-DC, a long-term evolution (LTE) evolved node B (eNB) operating in the sub-6 GHz band works as a control plane anchor to provide reliable connectivity to a user equipment (UE). The UE can receive/transmit data from/to eNB and NR next generation node B (gNB). For example, downlink traffic towards the UE is split in the middle, and part of the traffic is delivered to the UE over the LTE link and the rest over the mmWave NR link. The UE can achieve high throughput, the sum throughput of both links optimally, and the network can satisfy the UE’s stringent QoS requirements through link diversity. However, achieving optimal throughput in EN-DC is not straightforward.

Flow control, which decides the traffic splitting ratio on each link, has a significant impact on the performance of EN-DC. In-order delivery of packet data convergence protocol (PDCP) layer accounts for this effect. With in-order delivery, the throughput heavily depends on the order of arrival of PDCP protocol data units (PDUs). Specifically, if previous PDUs (i.e., PDUs with smaller sequence numbers) are not received, subsequent PDUs are buffered at the PDCP layer without being sent to the upper layer. To this end, flow control needs to avoid sending too many PDUs to one link and consider the order of arrival of PDUs at the destination. Furthermore, considering the dynamic nature of mmWave channels, flow control in EN-DC with mmWave NR link needs fast responsiveness and recovery.

Utilizing extra PDCP scheduling along with flow control helps to achieve fast recovery. For example, avoid allocating traffic to a previously overloaded link helps to rapidly resolve the problem of out-of-order PDUs. Also, duplicate transmission of PDUs over the both links in EN-DC provides link diversity, helping to reduce latency due to out-of-order PDUs. Duplicate PDUs arrive at the UE via the low-latency link, helping to resolve the out-of-order problem rapidly. PDCP scheduling should be done carefully since prioritizing PDUs can overload or underutilize a link. Moreover, careless duplication wastes resources and degrades throughput performance.

In this paper, we propose a blockage-aware flow control scheme (BAFC) in EN-DC where the NR link operates in the mmWave band, aiming to reduce PDCP reordering and maximize throughput. BAFC matches the transfer completion times of the traffic allocated to the two links in EN-DC. Moreover, BAFC limits the maximum transfer completion time over the two links by controlling the amount of allocated traffic to achieve fast responsiveness under highly dynamic mmWave channels.

We consider two types of PDCP scheduling: priority scheduling and packet re-injection scheduling. Priority scheduling (P) prefers a link with low expected delay, and packet re-injection (R) scheduling allows duplicate transmission of packets to exploit link diversity. Then we integrate BAFC with each scheduling to further improve throughput and latency performance of BAFC; they are named BAFC-P and BAFC-R, respectively.

The main contributions of this paper are three-folded:

- We propose BAFC, a novel flow control scheme that rapidly adapts to fast changing wireless channels under an EN-DC environment including mmWave links. BAFC decides the best traffic split ratio over the two links and controls the maximum amount of traffic allocated to each link to maximize throughput and achieve high responsiveness.

- For better ability to recover when link quality suddenly changes, we consider two types of PDCP scheduling and combine each with BAFC, i.e., BAFC-P and BAFC-R. BAFC-P resolves out-of-order packets as quickly as possible without sacrificing throughput, i.e., maximizing short-term throughput. On the other hand, BAFC-R deals with out-of-order packets using both links with minimal throughput degradation to minimize the buffering time of packet at the UE’s PDCP layer.

- Through extensive simulations using 5G-LENA [12] and ns-3 [13], we demonstrate that BAFC shows near-optimal throughput performance even during a blockage event. Moreover, we show that BAFC-P improves throughput, and BAFC-R significantly reduces the buffering time of packets compared to BAFC.

To the best of our knowledge, this is the first work that considers blockage effects in flow control design for an EN-DC environment. The rest of this paper is organized as follows. We describe related work and preliminaries in Section II. Section III presents the detailed design of BAFC and considered PDCP scheduling. In Section IV, we evaluate the performance of our schemes. Finally, we conclude the paper in Section V.

II. RELATED WORK AND PRELIMINARIES

A. RELATED WORK

There are many studies on the flow control for dual connectivity. In [14], the authors discuss traffic splitting with a fixed split ratio. Using a fixed split ratio is not desirable for EN-DC with mmWave NR links since mmWave link quality fluctuates frequently. Adaptation to highly dynamic mmWave channels requires dynamic flow control rather than using a fixed split ratio. In [15], the secondary node (SN) periodically requests data from the master node (MN) based on
information including the previous average throughput and the SN’s buffer status. The SN determines the amount of data requested to maintain a certain buffering time at its buffer, so the system can avoid both underutilization and overload. However, the operation needs the flow control period to be shorter than the round-trip delay of backhaul, resulting in high control overhead. Moreover, in-order delivery of PDCP is not considered in [15], which is critical to the throughput of EN-DC. In [16], the authors formulate a mixed integer linear programming problem to maximize network throughput without a low-complexity algorithm. The algorithm should be run whenever the channel states change. Due to this high level of complexity, the algorithm is difficult to apply in practice.

Some studies are conducted with different goals (i.e., [17], [18]) and scenarios (i.e., [19], [20]). In [17], the authors propose a flow control scheme that prioritizes user association with SNs to maximize the number of satisfied users in terms of QoS. In [18], the authors use cascade control to maintain delay skew at predetermined time intervals. In [19], a UE assists the MN by reporting the UE-received split ratio when its difference from the MN-transmitted split ratio exceeds a certain threshold. Even though the proposed flow control scheme in [19] can fine-tune the traffic split ratio faster and finer, it needs a change on the UE side. As a result, for the proposed flow control scheme in [19] to be deployed, all the UEs need to be modified to transmit certain control signals, which is not scalable. In [20], the authors consider the DC of LTE and non-3GPP technology. LTE-wireless local area network aggregation (LWA) integrates LTE and wireless local area network (WLAN) for downlink traffic. In [20], a path for packet transmission is determined by comparing the delays of LTE and WLAN paths when using LWA. Since the authors in [20] only consider a backhaul with non-zero latency, their proposed scheme cannot be applied when both backhaul links (i.e., X2-U and S1-U interfaces) have greater than zero latency. Moreover, a sudden drop in mmWave link capacity is not considered in [20].

Multi-path transport schedulers share some similarities with DC flow control since they are enabled to use multiple data paths and ensure in-order delivery. The authors in [21] consider a Wi-Fi station associated with two access points (APs) operating at the 5 GHz band and the 60 GHz band and the 60 GHz band. They work under the assumption that data paths and ensure in-order delivery. Moreover, in-order delivery of PDCP is not considered in [15], which is critical to the throughput of EN-DC. In [16], the authors formulate a mixed integer linear programming problem to maximize network throughput without a low-complexity algorithm. The algorithm should be run whenever the channel states change. Due to this high level of complexity, the algorithm is difficult to apply in practice.

In [22], the authors try to ensure fast responsiveness by maintaining a certain number of backlogged packets in each subflow of MPTCP. The authors in [23] improve the quality of experience (QoE) for videos in multi-path QUIC (MPQUIC) environments. For QoE improvement, XLINK re-injects lost packets to the path that has not been used for the original transmission and quickly recovers from packet losses. However, despite the similarities in architecture and roles, multi-path schedulers (i.e., [21]–[23]) work with feedback on end-to-end delivery, so they are difficult to use for DC.

B. BLOCKAGE MODEL

We use 3GPP blockage model B [24] to demonstrate the effect of mobile blockages on the signal quality. This model uses a geometric method of capturing the blockage effect, so it is possible to follow varying amounts of signal attenuation as blockages move. It uses a simple knife edge diffraction model to calculate attenuation caused by blockages. As a result, it does not take into account the penetration or reflection of the beam, which compromises the accuracy of modeling.

Before using the model B as our blockage model, we verify whether it successfully represents the blockage effect. To this end, we reproduce real-world measurements from [5] and [10] using this model and compare the results. Fig. 1 shows simulation results under recreated scenarios. Although a simpler and smoother shape than the real-world measurement results, model B succeeds in approximately mimicking the blockage effect of mobile blockages in terms of blocking duration and maximum attenuation.

The smoothness of model B, the key difference between simulation and measurement, has little effect when evaluating the link quality of the system. The reason is that the success or failure of a transmitted symbol is determined by the quantized signal quality level. For scalability, we use model B rather than real-world measurements to evaluate the performance of our proposed scheme in various configurations.

C. E-UTRA-NR DUAL CONNECTIVITY

In EN-DC, each UE is connected to one eNB acting as MN and one gNB acting as SN [11]. Focusing on the user plane, the eNB and gNB are connected via X2-U interface, and both base stations (BSs) are connected to the evolved packet core (EPC) via S1-U interface in our considered scenario. There are three types of bearers in EN-DC: master cell group (MCG) bearer, secondary cell group (SCG) bearer, and split bearer.
bearer. The network can configure either LTE PDCP or NR PDCP for MCG bearer while the others always use NR PDCP.

Specifically, data allocated to the split bearer are split at NR PDCP and forwarded to LTE RLC or NR RLC. NR PDCP is located at the gNB central unit (CU) whereas the radio link control (RLC) and lower layers are located at the gNB distributed unit (DU). The interface that connects the CU and DU of the gNB is called F1-U. For downlink traffic, which we will consider throughout this paper, traffic from split bearers arriving at the gNB CU is split into the eNB and the gNB DU via X2-U and F1-U interfaces, respectively. Afterward, the split traffic will be delivered to the UE through the corresponding wireless channel of each BS.

Data split at the NR PDCP should be merged at the UE’s PDCP layer that takes charge of in-order delivery to the upper layer. To this end, out-of-order PDUs are buffered until all PDUs necessary to resolve the out-of-order situation have been received. After the cavities are filled, the PDCP layer reorders the packets and delivers them to the upper layer. On the other hand, if the out-of-order situation has not been resolved within a certain time, the UE’s PDCP layer sends out-of-order packets to the upper layer instead of initiating the retransmission procedure.

The reason for this behavior is that there is no ACK mechanism in the PDCP layer, whereas the other layers that ensure in-order delivery (i.e., TCP and RLC) have ACK defined, and retransmission of necessary packets is not possible without ACK. The reordering procedure of PDCP results in an increase in buffered packets as well as reduced throughput and increased latency. Moreover, when TCP is used as the transport layer protocol, TCP may treat PDCP reordering as packet loss due to congestion and initiate slow start, resulting in severe throughput degradation.

In this paper, we consider an eNB operating in the sub-6 GHz band and a gNB operating in the mmWave band (i.e., 28 GHz band) that construct an EN-DC architecture. To eliminate the effect of multi-user medium access control (MAC) scheduling and solely investigate the effect of blockage on flow control, we connect an EN-DC UE to both BSs as shown in Fig. 2 in our scenario. Both X2-U and F1-U interfaces are assumed to have infinite capacity and fixed non-negative latency.

### III. PROPOSED SCHEME

In this section, we propose BAFC that aims to reduce PDCP reordering and maximize throughput. Unlike conventional flow control algorithms, BAFC determines the traffic split ratio, the portion of DL traffic allocated to each link, and the maximum amount of traffic allocated to each link when needed. We also propose BAFC-P and BAFC-R by adding the PDCP scheduling of priority and re-injection, respectively, to BAFC for further performance improvement in terms of short-term throughput and buffering time at the PDCP layer. The buffering time is the amount of time a packet stays in the UE’s PDCP buffer, defined as the time difference between arriving at the PDCP layer and leaving for the upper layer. To ensure in-order delivery, received packets need to be buffered if any previous packets have not yet been received.

### TABLE 1. Definitions of variables

| Variables | Definitions |
|-----------|-------------|
| $T$ (s)   | Period of flow control |
| $d_l$ (s) | Latency of X2-U interface |
| $d_m$ (s) | Latency of F1-U interface |
| $S_l$ (bps) | Estimated throughput of the LTE link during a flow control period |
| $S_m$ (bps) | Estimated throughput of the mmWave NR link during a flow control period |
| $C_l$ | Reported CQI value from the eNB |
| $C_m$ | Reported CQI value from the gNB DU |
| $E_{l,C_l}$ (bps/Hz) | Spectral efficiency of the LTE link regarding $C_l$ |
| $E_{m,C_m}$ (bps/Hz) | Spectral efficiency of the mmWave NR link regarding $C_m$ |
| $R_l$ (Hz s) | Time-frequency resource for data transmission of the eNB |
| $R_m$ (Hz s) | Time-frequency resource for data transmission of the gNB DU |
| $q_l$ (bits) | Estimated queue length of the eNB |
| $q_m$ (bits) | Estimated queue length of the gNB DU |
| $g_l$ (bits) | Reported queue length of the eNB |
| $g_m$ (bits) | Reported queue length of the gNB DU |
| $D$ (bps) | Traffic demand of the UE |
| $B$ (bits) | Amount of backlogged traffic in the gNB CU |
| $\tau_l$ (s) | Completion time of the LTE link |
| $\tau_m$ (s) | Completion time of the mmWave NR link |
| $M_l$ (bits) | Total amount of traffic allocated to the LTE link during a flow control period |
| $M_m$ (bits) | Total amount of traffic allocated to the mmWave NR link during a flow control period |
| $I_s$ (s) | Subframe length in LTE |
| $L_l$ (bits) | Total amount of traffic allocated to the LTE link during $I_s$ |
| $L_m$ (bits) | Total amount of traffic allocated to the mmWave NR link during $I_s$ |
| $I_l$ (bits) | Total amount of traffic re-injected to the LTE link |
| $I_m$ (bits) | Total amount of traffic re-injected to the mmWave NR link |
A. BLOCKAGE-AWARE FLOW CONTROL SCHEME (BAFC)

To conduct flow control, the eNB and gNB DU report needed information to the gNB CU every \( T \) seconds in our scenario. The reported information is about channel quality indicator (CQI) value and queue length at each BS, and experiences backhaul latency when delivered to the gNB CU. The latencies of X2-U interface and F1-U interface are denoted as \( d_l \) and \( d_m \), respectively. We assume the eNB and the gNB DU know \( d_l \) and \( d_m \), respectively. With this assumption, the BSs match the arrival time of the information at the gNB CU at the same time. That is, the eNB and gNB DU transmit the information \( d_l \) and \( d_m \) to the gNB CU, respectively, before the target arrival time.

The first thing the gNB CU does is to estimate the throughput of each link over a flow control period, denoted as \( S_l \) and \( S_m \), using \( C_l \) and \( C_m \) which are the reported CQI values from the eNB and gNB DU. Throughput can be estimated by multiplying the average amount of available resources per unit time by the spectral efficiency of each link. Given the reported CQI values, the gNB CU calculates the spectral efficiency of each link, denoted as \( E_{l,C_l} \) and \( E_{m,C_m} \), respectively. We denote time-frequency resources that can be used for each transmission of each link during the flow control period \( T \) as \( R_l \) and \( R_m \), respectively. Then throughput estimates for LTE link and mmWave NR link are calculated as

\[
S_l = \frac{E_{l,C_l} R_l}{T},
\]

\[
S_m = \frac{E_{m,C_m} R_m}{T},
\]

respectively, where the spectral efficiencies corresponding to the reported CQI values for LTE and NR links are obtained from [25] and [26], respectively. \( R_l \) and \( R_m \), vary according to the system configurations such as the length of physical downlink control channel (PDCCH) and the number of reference signals. For example, \( R_l \) decreases as a length of LTE PDCCH increases from 1 to 3 symbols.

This means that the gNB CU should know every configuration of the eNB and gNB DU to estimate throughput using (1) and (2), which is not practical. As an alternative, we measure the maximum throughput of each link, and estimate throughput by multiplying the ratio of spectral efficiency corresponding to the reported CQI value to the maximum spectral efficiency. Denoting the CQIs with the maximum spectral efficiency as \( C_{l,\text{max}} \) and \( C_{m,\text{max}} \) from [25] and [26], respectively, we estimate the throughput of each link as

\[
S_l = S_{l,\text{max}} \frac{E_{l,C_l}}{E_{l,C_{l,\text{max}}}},
\]

\[
S_m = S_{m,\text{max}} \frac{E_{m,C_m}}{E_{m,C_{m,\text{max}}}},
\]

respectively, where \( S_{l,\text{max}} \) and \( S_{m,\text{max}} \) represent the maximum throughput of LTE link and mmWave NR link, respectively.

After the throughput estimate, the gNB CU performs queue length estimation of the eNB and gNB DU. Queue length estimation is necessary due to the difference between the reported queue length and the actual queue length of each BS at the first arrival of traffic allocated according to the newly decided split ratio. The difference is caused by the packet arrivals and departures after the queue length measurement of the eNB and gNB DU and before the first packet with the newly performed flow control.

Packets transmitted from the gNB CU to the eNB during \( T - 2d_l \) to \( T \) are not included in queue length reporting. However, these packets affect the queue length of the eNB at \( T + d_l \) which is important for BAFC. Similarly, packets allocated to the gNB DU during \( T - 2d_m \) to \( T \) should be considered for the queue length estimation. In Fig. 3, the packets that make the difference are transmitted from the gNB CU to each BS in the colored intervals. Let us denote the amount of traffic allocated to each link for the colored intervals in Fig. 3 as \( A_l \) and \( A_m \), respectively.

Packets transmitted from the eNB to the UE during \( T - d_l \) to \( T + d_l \) should also be considered to estimate the queue length of the eNB. Regarding the throughput of the LTE link, \( S_l \), the packet transmission amount is \( 2S_l d_l \). Likewise, the gNB DU transmits \( 2S_m d_m \) bits packets to the UE, which causes the queue length report value to differ from the actual queue length upon reception of the first packet of new flow control. Considering the packet arrivals and departures, we can express the estimated queue length as

\[
Q_l = \max \left( 0, q_l + A_l - 2S_l d_l \right),
\]

\[
Q_m = \max \left( 0, q_m + A_m - 2S_m d_m \right),
\]

respectively, where \( q_l \) and \( q_m \) indicate the reported queue length from each BS, respectively.

Based on the estimates of throughput and queue length at each BS, the gNB CU performs BAFC with a period of \( T \). The role of BAFC is to decide the portion of traffic allocated to mmWave NR link, denoted as \( \eta \), then the remaining portion of traffic, \( 1 - \eta \), is allocated to LTE link. In some cases, BAFC not only outputs the traffic split ratio, but also limits the maximum amount of traffic allocated to each link.

Firstly, we concentrate on deciding the split ratio \( \eta \) without considering the maximum amount of traffic allocated when the UE demands constant bit rate (CBR) traffic with rate \( D \). To maximize the throughput of the EN-DC UE, the
maximum completion time of traffic allocated to each link should be minimized. The completion time is defined as the time between the start of traffic allocation at the gNB CU and the last packet arrival of the allocation at the UE for each period. Let the total amount of traffic allocated to each link during a traffic allocation period as \( M_l \) and \( M_m \), respectively. If all the packets arrived at the gNB CU are allocated to one of the two links, we have \( M_l \) and \( M_m \) as
\[
M_l = (DT + B)(1 - \eta), \quad M_m = (DT + B)\eta,
\]
respectively, where \( DT \) is the amount of total traffic arriving at the gNB CU according to the traffic demand of the UE, and \( B \) represents the amount of the backlogged traffic at the gNB CU. Adding the estimated queue length of each BS to the traffic amount that will be allocated, each BS should transmit \( M_l + Q_l \) and \( M_m + Q_m \), respectively, toward the UE for a given period. Assuming both links have non-zero throughput and considering the throughput and backhaul latency, we can express the completion time of each link, given the split ratio, as
\[
\tau_l(\eta) = \frac{M_l + Q_l}{S_l} + d_l, \quad \tau_m(\eta) = \frac{M_m + Q_m}{S_m} + d_m,
\]
respectively. The last term adds backhaul latency to the completion time because split traffic towards the gNB DU experiences backhaul latency to arrive at the BSs after the allocation. Optimal \( \eta \), denoted as \( \eta^* \), that minimizes the larger value between \( \tau_l(\eta) \) and \( \tau_m(\eta) \) is given as
\[
\eta^* = \arg\min_\eta \max(\tau_l(\eta), \tau_m(\eta)). \quad (11)
\]
To obtain \( \eta^* \), we first express the first derivative of \( \tau_l(\eta) \) and \( \tau_m(\eta) \) in terms of \( \eta \) as
\[
\frac{\partial \tau_l}{\partial \eta} = -\frac{DT + B}{S_l}, \quad \frac{\partial \tau_m}{\partial \eta} = \frac{DT + B}{S_m}, \quad (12) \quad (13)
\]
respectively. Since the numerators in (12) and (13) are positive, and the denominators are assumed to be positive, \( \partial \tau_l/\partial \eta \) is non-positive and \( \partial \tau_m/\partial \eta \) is non-negative. Note that solving the min-max problem for a non-decreasing function and a non-increasing function is equivalent to finding a point where two functions are equal. As a result, \( \eta^* \) is the split ratio that makes \( \tau_l(\eta) \) and \( \tau_m(\eta) \) equal. Substituting (7) and (8) into (9) and (10), respectively, we get optimal split ratio \( \eta^* \) as
\[
\eta^* = \frac{S_m(DT + B) + S_lS_m(d_l - d_m) + S_mQ_l - S_lQ_m}{(S_l + S_m)(DT + B)}.
\]
Ignoring queue length and backhaul latency of the both links, Eq. (14) shows the throughput ratio of the two links, which corresponds to the idea in [21].

**Algorithm 1 BAFC**

1. repeat
2. For each flow control period
3. Collect \( C_l, C_m, q_l, \) and \( q_m \)
4. Estimate \( S_l, S_m, Q_l, \) and \( Q_m \)
5. Calculate \( \eta_p \)
6. if \( \tau_l(\eta_p) > T + \max(d_l, d_m) \) or \( \tau_m(\eta_p) > T + \max(d_l, d_m) \) or \( S_l = 0 \) or \( S_m = 0 \) then
7. \( M_l \leftarrow \max(0, S_l(T + \max(d_l, d_m) - d_l) - Q_l) \)
8. \( M_m \leftarrow \max(0, S_m(T + \max(d_l, d_m) - d_m) - Q_m) \)
9. else
10. \( M_l \leftarrow (DT + B)(1 - \eta_p), \) \( M_m \leftarrow (DT + B)\eta_p \)
11. end if
12. repeat
13. For each split period
14. \( L_l \leftarrow \frac{M_l}{T}, L_m \leftarrow \frac{M_m}{T} \)
15. while \( L_m > 0 \) do
16. Assign incoming traffic to gNB DU
17. \( L_m \leftarrow L_m - \text{packet size} \)
18. end while
19. while \( L_l > 0 \) do
20. Assign incoming traffic to eNB
21. \( L_l \leftarrow L_l - \text{packet size} \)
22. end while
23. until End of the flow control period
24. until System terminates

Practically, \( \eta^* \) should lie between 0 and 1, but with extremely long queue length at the gNB DU or large F1-U latency, it may be negative. On the other hand, long queue length at the eNB or large X2-U latency might make \( \eta^* \) greater than 1. In the former case, allocating traffic to mmWave NR link is undesirable because such operations incur out-of-order packet delivery at the PDCP layer. Similarly, in the latter case, all the traffic should be allocated to mmWave NR link. Therefore, we practically set the optimal split ratio as
\[
\eta_p = \min(1, \max(0, \eta^*)). \quad (15)
\]

Using \( \eta_p \) at the gNB CU is desirable if the channel qualities of both links are stable. However, splitting traffic with \( \eta_p \) leads to poor responsiveness to the channel dynamics. For example, when a blockage blocks the LOS path between the gNB DU and UE, \( S_l + S_m \) may become smaller than \( D \). In such a case, allocating all the traffic increases \( Q_l, Q_m \), or both. When the channel quality is resumed, the system has a new split ratio. Nevertheless, traffic split with the new ratio has no effect on system performance until all the backlogged traffic at the BSs are transmitted to the UE. To sum up, inadvertent traffic allocation without considering the increase in the queue length of the BSs degrades the responsiveness.

We limit \( \tau_l \) and \( \tau_m \) by controlling \( M_l \) and \( M_m \) to achieve high responsiveness. In detail, if \( \tau_l(\eta_p) > T + \max(d_l, d_m) \)
or \( \tau_m(n_p) > T + \max(d_l, d_m) \), we adjust \( M_l \) and \( M_m \) to satisfy
\[
M_l = \arg\min_{x \geq 0} \{ \tau_l(M_l = x) - \{ T + \max(d_l, d_m) \} \}, \tag{16}
\]
\[
M_m = \arg\min_{x \geq 0} \{ \tau_m(M_m = x) - \{ T + \max(d_l, d_m) \} \}, \tag{17}
\]
respectively, where \(| x |\) takes the absolute value of \( x \). Solving (16) and (17), we get
\[
M_l = \max(0, S_l(T + \max(d_l, d_m) - d_l) - Q_l), \tag{18}
\]
\[
M_m = \max(0, S_m(T + \max(d_l, d_m) - d_m) - Q_m). \tag{19}
\]
Note that both (18) and (19) work even when \( S_l = 0 \) or \( S_m = 0 \) while we assume positive throughput to obtain \( n_p \).

Lastly, BAFC splits traffic according to \( M_l \) and \( M_m \) every \( T_s \) seconds. We set \( T_s \) as a subframe length in LTE. During every split period, the gNB CU assigns \( M_l T_s / T \) and \( M_m T_s / T \) of traffic to the gNB DU and eNB, respectively. Algorithm 1 summarizes BAFC, where \( L_l \) and \( L_m \) denote split traffic allocated to each link during \( T_s \), respectively.

**B. PDCP SCHEDULING**

To alleviate the PDCP reordering problem further, we design two types of PDCP scheduling on top of BAFC: priority scheduling (BAFC-P) and packet re-injection (BAFC-R). For traffic allocation, BAFC-P gives high priority to a link with lower expected latency over the other link. BAFC-R transmits duplicate packets over a link with low expected delay when the other link has a large number of packets in the queue. Both PDCP scheduling schemes help to resolve the PDCP reordering problem fast. Let us consider using the baseline BAFC when blockage degrades the quality of the mmWave link. Packets will continue to be assigned to the gNB DU while packets queued at the gNB DU during the previous flow control periods are being transmitted to the UE. As a result, the newly assigned packets to the gNB DU will incur the PDCP reordering problem even when all the previously queued packets are delivered.

1) Priority scheduling

To avoid the PDCP reordering problem as much as possible, BAFC-P performs priority scheduling on top of BAFC. The amount of data to be prioritized is determined by the difference in time needed to empty the queues on the two links. Revisiting (9) and (10), the time needed to empty the queues of the eNB and gNB DU can be calculated as \( \tau_1(1) \) and \( \tau_m(0) \)\(^2\), respectively. If \( \tau_1(1) > \tau_m(0) \), i.e., LTE link takes more time to empty its queue, traffic should be preferentially allocated to mmWave NR link. To make mmWave NR link transmit preferentially allocated traffic just until LTE link finish draining the queue, we have the maximum amount of prioritized traffic as
\[
P_{\text{max}, m} = S_m(\tau_1(1) - \tau_m(0)). \tag{20}
\]
Similarly, if \( \tau_m(0) > \tau_1(1) \), we obtain the maximum amount of prioritized traffic that the gNB CU allocates to the eNB as
\[
P_{\text{max}, l} = S_l(\tau_m(0) - \tau_1(1)). \tag{21}
\]

The results of (20) and (21) represent the maximum amount because a smaller amount should be preferentially allocated when priority scheduling is delayed. For example, when traffic should be allocated to the eNB with higher priority, traffic allocated \( \tau_m(0) - d_l \) after flow control will arrive at the eNB when the gNB DU has already cleared its queue. The allocation is not effectively prioritized. Hence, we perform priority scheduling for traffic to be transmitted to the UE before the queue of the other link becomes empty.

For ease of explanation, let us assume a situation where the queue of the mmWave NR link is long and the LTE link is preferentially used. We consider a split period starting \( t \) after flow control. If\(^3\) \( t + T_s < \frac{Q_l + P_l}{S_l} \), where \( P_l \) denotes the amount of traffic allocated to the LTE link through priority scheduling, the allocated traffic arrives before the eNB empties its queue. In this case, priority scheduling is performed until \( P_l + M_l T_s / T_s \geq P_{\text{max}, l} \).

For the traffic that would arrive after the eNB clears its queue, priority scheduling is performed when \( t + T_s + d_l + \frac{M_l T_s}{S_l} < \frac{Q_l}{S_l} + d_m \). With this criterion, the allocated traffic will be delivered to the UE before the gNB DU drains its

\(^2\)\( S_l \) and \( S_m \) are assumed to be positive. Otherwise, from (18) and (19), one of the two links has no allocated data, making priority scheduling meaningless.

\(^3\)We consider all the allocated traffic transmitted from the gNB CU at the end of the split period to be conservative and avoid underutilization of the mmWave NR link.
Algorithm 3 *BAFC-R*

1: repeat
2: Run line 3–11 of Algorithm 1
3: Perform re-injection
4: if $\tau_m(0) > \tau_l(1)$ then
5: \hspace{1em} $P_l \leftarrow I_l$
6: \hspace{1em} $M_l \leftarrow \max(0, M_l - I_l)$
7: else
8: \hspace{1em} $P_m \leftarrow I_m$
9: \hspace{1em} $M_m \leftarrow \max(0, M_m - I_m)$
10: end if
11: Run line 4–19 of Algorithm 2
12: until System terminates

queue. Here, also the gNB CU performs scheduling until $P_l + M_l T_k \geq P_{max,j}$. After priority scheduling is performed, the gNB CU starts allocating traffic to the gNB DU. Since, some split periods are used for priority scheduling, traffic amount $M_m$ should be evenly assigned for the rest of the period. We can similarly obtain the conditions for priority scheduling when the mmWave NR link is preferentially used. The procedures for *BAFC-P* are summarized in Algorithm 2.

2) Re-injection

We propose *BAFC-R* that aims to resolve the out-of-order delivery problem as fast as possible. *BAFC-R* performs as follows. When one link has a long queue, packets later in the queue are re-injected on the other link at the start of the next flow control period. For example, when the gNB DU has five packets in the queue and the eNB has a clear queue, the gNB CU re-injects the fourth and fifth packets to the eNB. While the gNB DU transmits three packets from the head of the queue, the eNB transmits the two re-injected packets to the UE. As a result, the out-of-order delivery problem has been resolved within 60% of the time compared to the baseline BAFC.

To obtain the optimal amount of re-injected traffic, denoted as $I_l$, let us assume the gNB DU has a long queue. In this case, $I_l$ should satisfy

$$Q_m - I_l \leq S_m \leq Q_l + I_l + d_l. \quad (22)$$

As a result, we have

$$I_l = \frac{S_l S_m}{S_l + S_m} \left( \frac{Q_m}{S_m} - \frac{Q_l}{S_l} + d_m - d_l \right). \quad (23)$$

Note that link selection for re-injection is the same as that for priority scheduling. If $\tau_m(0) > \tau_l(1)$, re-injection should be performed to the LTE link. For $\tau_m(0) < \tau_l(1)$, traffic should be re-injected to the gNB DU. Then, the optimal amount of traffic re-injected to the gNB DU is

$$I_m = \frac{S_l S_m}{S_l + S_m} \left( \frac{Q_l}{S_l} - \frac{Q_m}{S_m} + d_l - d_m \right). \quad (24)$$

Traffic re-injection of the amount $I_l$ or $I_m$ is performed immediately upon the start of a flow control period since the corresponding traffic is already buffered at the gNB CU. Traffic split after re-injection should consider the amount of re-injected traffic. Furthermore, the gNB CU performs priority scheduling after re-injection, assuming the amount $I_l$ or $I_m$ has already been allocated with high priority. Algorithm 3 summarizes the overall procedures of *BAFC-R*.

Although *BAFC-R* alleviates the out-of-order delivery problem, it has small overhead. The first overhead is the transmission of redundant packets via both links of the EN-DC UE. To avoid meaningless re-injection, we prevent the gNB CU from re-injecting, due to the backhaul latency difference by setting $I_l$ or $I_m$ to 0 when $Q_m$ or $Q_l$ is 0, respectively. The second overhead is buffering of traffic sent during the past flow control periods at the gNB CU for re-injection. Considering large computing power and memory of the gNB, we claim that this is negligible.

**IV. PERFORMANCE EVALUATION**

In this section, we evaluate the performance of our proposed schemes through simulation. To this end, we first implement the 3GPP blockage model B on LTE module of ns-3 [13] and the mmWave NR module of 5G-LENA simulator [12]. After implementing model B in the simulators, we collect PDCP layer packet reception traces and CQI traces under saturated conditions. We perform our simulation experiments using the 3GPP blockage model B on LTE module of ns-3 [13] and the mmWave NR module of 5G-LENA simulator [12]. After implementing model B in the simulators, we collect PDCP layer packet reception traces and CQI traces under saturated conditions.

**TABLE 2.** Simulation Settings

| Parameters                        | Value |
|-----------------------------------|-------|
| gNB DU center frequency           | 28 GHz|
| eNB center frequency              | 2100 MHz|
| gNB DU bandwidth                  | 400 MHz|
| eNB bandwidth                     | 100 MHz|
| gNB DU Tx power                   | 33 dBm|
| eNB Tx power                      | 44 dBm|
| gNB DU height                     | 10 m  |
| eNB height                        | 25 m  |
| UE height                         | 1.5 m |
| MmWave NR subframe length         | 0.25 ms|
| LTE subframe length               | 1 ms  |
| PDCP packet size                  | 1252 bytes|
| Bus size                          | 10.5 m × 2.5 m × 3.5 m |

**FIGURE 4.** Capacity of mmWave links during blockage events according to the UE location.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
traffic. PDCP traces show the maximum number of PDCP packets received by the UE for each link, considering the channel quality. Then we construct the EN-DU architecture in Python. Our EN-DC simulator splits traffic at the PDCP of the gNB CU at the packet level. Split traffic is delivered to either the eNB or gNB DU after a certain backhaul latency, and finally arrives at the UE.

To examine our proposed schemes, we consider an urban grid scenario [27]. We install the eNB and gNB DU in the same location as in the bird’s-eye-view, on the walls of a building, but with different heights. A static UE is located across the street and a bus goes to the lane closest to the UE. We vary the position of the UE to generate various blockage effects in terms of duration, degree, and shape. In detail, we examine 41 locations of the UE along a street with adjacent points 5 m apart and the center point just across the BSs. Fig. 4 shows examples of capacity fluctuation in mmWave NR links when a bus traverses the LOS path between the UE and the gNB DU depending on the UE location. Even for the same blockage event source, the impact of blockage is different according to the UE location. We summarize the simulation parameters in Table 2 including the configurations in [28].

A. IDEAL BACKHAUL
We first evaluate our proposed schemes under an ideal backhaul scenario where both X2-U interface and F1-U interface have zero latency. This case eliminates the error in queue length estimation. We compare the throughput of BAFC, BAFC-P, and BAFC-R with that of the best fixed ratio scheme during blockage events. The best fixed scheme is a decentralized scheme in [14] that achieves the highest throughput when the LOS paths are secured for both links, found via exhaustive search with 1% granularity. We compare the performance of the proposed scheme with that of the best fixed scheme because it is the only work that matches our goal and is applicable to all scenarios under consideration. Furthermore, we normalize the throughput using the sum capacity of both links. Since the sum capacity is the maximum achievable throughput in EN-DC, normalized throughput represents the optimality of the proposed schemes.

Fig. 5 shows the normalized throughput of the best fixed scheme and the proposed schemes, collected in different UE locations. Because link capacity changes during blockage events, the best fixed scheme suffers from throughput
degradation. In the worst case, the best fixed scheme only provides 65.4% of the sum capacity to users. BAFC, BAFC-P, and BAFC-R achieve 99.4%, 99.5%, and 99.3% of the normalized throughput with the flow control period of 10 ms on average, respectively. This near-optimal performance comes from the rapid adaptation by controlling queue lengths and precisely determining the amount of traffic allocated to each link to avoid PDCP reordering.

By alleviating the out-of-order delivery problem fast, BAFC-P achieves the best throughput among our proposed schemes. BAFC-R performs the worst in terms of throughput among our proposed schemes because of the redundant packet transmission over the two links, losing the opportunity to transmit new packets. Nevertheless, BAFC-R achieves 98.4% of the normalized throughput in the worst case.

We observe a similar tendency for the 50 ms flow control period. BAFC-P provides the highest throughput to the user while BAFC-R provides the lowest throughput. However, the 50 ms control period leads to throughput decrease for all the proposed schemes. Performance degradation comes from slow adaptation due to the long flow control period since our schemes adjust their policy every period. Despite the performance degradation, all the proposed schemes show 98.5% of the optimal throughput on average, even for the worst case, BAFC-R shows 99.9% of the normalized throughput.

We investigate the maximum buffering time of packets during a blockage event. To meet stringent QoS requirements of delay sensitive applications even with mobile blockages, the maximum buffering time should stay low even with dramatic fluctuations of mmWave channels. Fig. 6 shows the distribution of the maximum buffering time at each UE location. BAFC-R reduces the maximum buffering time by up to 43.2% compared to BAFC and BAFC-P with the 10 ms flow control period. In the case of the 50 ms flow control period, BAFC-R reduces the maximum buffering time by up to 43.9%. Note that a longer flow control period also increases the maximum buffering time.

Slow adaptation due to a longer flow control period increases the number of out-of-order packets when channel quality becomes bad in the middle, resulting in a longer buffering time. The difference in throughput and delay performance between BAFC, BAFC-P, and BAFC-R can be explained by the throughput over time of each scheme, as shown in Fig. 7. In Fig. 7, the blue line indicates throughput performance of BAFC, the red line represents performance of BAFC combined with priority scheduling (i.e., BAFC-P), and the yellow line shows performance of BAFC-R. The purple line shows performance of the best fixed scheme. Without blockages, the best fixed scheme shows the same performance as the proposed schemes. On the other hand, the proposed schemes outperform the best fixed scheme significantly during blockage events. During 1 s to 1.5 s, the proposed schemes show about 45% throughput gain compared to the best fixed scheme. Moreover, the gain increases as the quality of mmWave NR links degrades more severely.

Focusing on 800 ms – 1 s, the throughput of each of the proposed schemes spikes once. Such spikes appear when cavities are filled, and out-of-order packets are resolved and delivered to the upper layer. The higher the number of out-of-order packets, the larger the area under each spike. Note that no spike appears with the best fixed scheme since it does not resolve the problem of out-of-order packets after link quality is degraded. A spike in BAFC is thicker than the spikes in BAFC-P and BAFC-R. This means, BAFC requires more time than BAFC-P and BAFC-R to resolve the out-of-order problem. As a result, BAFC-P shows better throughput performance than BAFC. The first spike in BAFC-R appears earlier than the spikes in the others, and after that, the throughput of BAFC-R is lower than that of the others. Packet re-injection in BAFC-R helps to alleviate the out-of-order problem faster but causes a negative impact on throughput performance later on.

**B. ASYMMETRIC BACKHAUL LATENCY**

We now consider a scenario where the gNB DU is co-located with the gNB CU but the eNB is not. In this case, F1-U
Nonetheless, BAFC-R still significantly reduces the maximum buffering time, as shown in Fig. 8. It reduces the maximum buffering time by up to 27.5% and 31.5% compared to BAFC and BAFC-P, respectively, when the flow control period is set to 10 ms. When the flow control period is 50 ms, BAFC-R reduces the maximum buffering time by up to 39.1% and 39.8%, compared to BAFC and BAFC-P, respectively. Since a longer flow control period is more likely to incur a longer queue length at the mmWave link when a blockage event occurs, $I_l$ will become larger, and so does the gain of BAFC-R.

C. SYMMETRIC NON-ZERO BACKHAUL LATENCY

The last scenario we consider is a symmetric non-zero backhaul latency case. The gNB CU takes charge of multiple gNB DUs to reduce the deployment cost for it. As a result, F1-U interface has non-zero latency. We set the latency of X2-U interface and F1-U interface to be equal to 3 ms. Despite the estimation error due to the backhaul latency, our proposed schemes successfully achieve near-optimal throughput. The average normalized throughput during blockage events is over 98.5% in all cases. BAFC-R with the 50 ms flow control period shows the worst case performance, achieving 95.3% of the normalized throughput. BAFC-R always reduces the maximum buffering time by up to 40% in the symmetric non-zero backhaul latency scenario regardless of length of the flow control period.

D. SUMMARY AND DISCUSSION

To sum up, our proposed schemes, namely, BAFC, BAFC-P, and BAFC-R successfully achieve near-optimal throughput even during blockage events in all the simulated scenarios. Among the proposed schemes, BAFC-P shows the highest throughput by fast resolving the problem of out-of-order packet delivery, and BAFC-R reduces the buffering time through the cooperation of the two links. However, there remain several items to discuss.

1) We have not considered multi-user systems to eliminate ambiguity in performance analysis, caused by multi-user MAC scheduling. For the multi-user case, the throughput estimate we used may be inaccurate. We can use throughput estimation algorithms, such as those presented in [29].

2) We assumed that the gNB CU knows the user’s traffic demand, which is impractical. The user’s demand can be predicted by utilizing prediction algorithms, such as those shown in [30].

3) We evaluated the performance of our proposed schemes using only CBR traffic. However, they are not based on any specific type of traffic. Derivation of the amount of traffic split to each link, priority scheduling, and packet re-injection are performed assuming all the traffic arrives at the beginning of the flow control period. Hence, even with traffic patterns other than CBR, our proposed schemes may show slightly degraded performance, but they are expected to work well.
4) Handover may be induced when a blockage event occurs, which may cause the proposed schemes not to perform properly. As discussed in [31], frequent handovers due to short-term blockage are inappropriate. We may consider handover skipping when short-term blockage events occur.

V. CONCLUSION

In this paper, we proposed BAFC, a flow control scheme for EN-DC with LTE link and mmWave NR link. To cope with very dynamic channel characteristics of mmWave links, BAFC controls the queue length of each BS. It matches the data transfer completion times of the two links to maximize throughput while mitigating the PDCP reordering problem at the UE. To improve throughput performance and decrease buffering time, we additionally developed two PDCP scheduling schemes: priority scheduling (BAFC-P) and packet re-injection (BAFC-R). Through extensive simulations using traces from well-known simulators, we showed that BAFC achieves over 99% optimal throughput regardless of the degree of channel fluctuations due to mobile blockages. Moreover, BAFC-P shows better throughput performance compared to BAFC, and BAFC-R significantly reduces buffering time with minimal throughput decrease.

REFERENCES

[1] F. Hu, Y. Deng, W. Saad, M. Bennis, and A. H. Aghvami, “Cellular-connected wireless virtual reality: Requirements, challenges, and solutions,” IEEE Communications Magazine, vol. 58, no. 5, pp. 105–111, May. 2020.

[2] S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter-wave cellular wireless networks: Potentials and challenges,” Proc. IEEE, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[3] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, “A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges,” Wireless networks, vol. 21, no. 8, pp. 2657–2676, Nov. 2015.

[4] X. Chen, L. Tian, P. Tang, and J. Zhang, “Modeling of human body shadowing based on 28 GHz indoor measurement results,” in Proc. IEEE VTC-Fall, 2016.

[5] T. Choi, C. U. Bas, R. Wang, S. Hur, J. Park, J. Zhang, and A. F. Molisch, “Measurement based directional modeling of dynamic human body shadowing at 28 GHz,” in Proc. IEEE GLOBECOM, 2018.

[6] Y. Koda, K. Yamamoto, T. Nishio, and M. Morikura, “Time series measurement of IEEE 802.11 ad signal power involving human blockage with HMM-based state estimation,” in Proc. IEEE VTC-Fall, 2017.

[7] G. R. MacCartney, T. S. Rappaport, and S. Rangan, “Rapid fading due to human blockage in pedestrian crowds at 5G millimeter-wave frequencies,” in Proc. IEEE GLOBECOM, 2017.

[8] V. Semkin, U. Virk, A. Karttunen, K. Haneda, and A. V. Räisänen, “8-band propagation channel measurements in an urban street canyon,” in Proc. IEEE EuCAP, 2015.

[9] J.-J. Park, J. Lee, J. Liang, K.-W. Kim, K.-c. Lee, and M.-D. Kim, “Millimeter wave vehicular blockage characteristics based on 28 GHz measurements,” in Proc. IEEE VTC-Fall, 2017.

[10] C. U. Bas, R. Wang, S. Sangodoyin, S. Hur, K. Whang, J. Park, J. Zhang, and A. F. Molisch, “Dynamic double directional propagation channel measurements at 28 GHz-invited paper,” in Proc. IEEE VTC-Spring, 2018.

[11] 3GPP TS 37.340, “E-UTRA and NR: Multi-connectivity; Stage-2,” V16.10.0, Mar. 2020.

[12] CTTC. 5G LENA Project. [Online]. Available: https://5g- lena.cttc.es/download/

[13] The Network Simulator-3. [Online]. Available: https://www.isi.edu/nsnam/ns/

[14] 3GPP R2-132859, “Throughput evaluation and comparison of with and without UP bearer split,” Aug. 2013.

[15] H. Wang, C. Rosa, and K. I. Pedersen, “Dual connectivity for LTE-advanced heterogeneous network,” Wireless networks, vol. 22, no. 4, pp. 1315–1328, May. 2016.

[16] M.-S. Pan, T.-M. Lin, C.-Y. Chiu, and C.-Y. Wang, “Downlink traffic scheduling for LTE-A small cell networks with dual connectivity enhancement,” IEEE Communications Letters, vol. 20, no. 4, pp. 796–799, Apr. 2016.

[17] R. P. Antonioli, E. B. Rodrigues, D. A. Sousa, I. M. Guerreiro, C. F. e Silva, and F. R. Cavalcanti, “Adaptive bearer split control for 5G multi-RAT scenarios with dual connectivity,” Computer Networks, vol. 161, pp. 183–196, Oct. 2019.

[18] T. Wigren, K. Lau, R. Delgado, and R. H. Middleton, “Delay skew packet flow control in wireless systems with dual connectivity,” IEEE Trans. on Vehicular Technology, vol. 67, no. 6, pp. 5357–5371, Jun. 2018.

[19] R. P. Antonioli, I. M. Guerreiro, D. A. Sousa, E. B. Rodrigues, T. F. Maciel, F. R. P. Cavalcanti et al., “User-assisted bearer split control for dual connectivity in multi-RAT 5G networks,” Wireless Networks, vol. 26, no. 5, pp. 3675–3685, Mar. 2020.

[20] D. López-Pérez, D. Laselva, E. Wallmeier, P. Purovesi, P. Lundén, E. Virtej, P. Lechowicz, E. Malkamaki, and M. Ding, “Long term evolution-wireless local area network aggregation flow control,” IEEE Access, vol. 4, pp. 9860–9869, Jan. 2016.

[21] S. K. Saha, S. Aggarwal, R. Pathak, D. Koutsoukias, and J. Widmer, “MuShier: an agile multipath-TCP scheduler for dual-band 802.11 ad/ac wireless LANs,” in Proc. ACM MobiCom, 2019.

[22] C. Lee, J. Jung, and J.-M. Chung, “DEFT: Multipath TCP for high speed low latency communications in 5G networks,” IEEE Trans. on Mobile Computing, vol. 20, no. 12, pp. 3311–3323, Dec. 2020.

[23] Z. Zheng, Y. Ma, Y. Liu, F. Yang, Z. Li, Y. Zhang, J. Zhang, W. Shi, W. Chen, D. Li et al., “Xlink: Qoe-driven multi-path quic transport in large-scale video services,” in Proc. ACM SIGCOMM, 2021.

[24] 3GPP TR 38.901, “Study on channel model for frequencies from 0.5 to 100 GHz,” V16.1.0, Dec. 2019.

[25] 3GPP TS 36.213, “Physical layer procedures,” V16.8.0, Dec. 2021.

[26] 3GPP TS 38.214, “Physical layer procedures for data,” V16.5.0, Mar. 2021.

[27] 3GPP TR 38.913, “Study on scenarios and requirements for next generation access technologies,” V16.0.0, Jul. 2020.

[28] 3GPP TR 38.802, “Study on new radio access technology physical layer aspects,” V14.2.0, Sep. 2017.

[29] D. Raca, A. H. Z rahran, C. J. Sreenan, R. K. Sinha, E. Halepovic, R. Jana, and V. Gopalakrishnan, “On leveraging machine and deep learning for throughput prediction in cellular networks: Design, performance, and challenges,” IEEE Communications Magazine, vol. 58, no. 3, pp. 11–17, Mar. 2020.

[30] A. Azari, P. Papapetrou, S. Denic, and G. Peters, “User traffic prediction for proactive resource management: Learning-powered approaches,” in Proc. IEEE GLOBECOM, 2019.

[31] L. Sun, J. Hou, and T. Shu, “Spatial and temporal contextual multi-armed bandit handovers in ultra-dense mmWave cellular networks,” IEEE Trans. on Mobile Computing, vol. 20, no. 12, pp. 3423–3438, Jun. 2020.
SAEWOONG BAHK (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from Seoul National University (SNU), in 1984 and 1986, respectively, and the Ph.D. degree from the University of Pennsylvania, in 1991. He was with AT&T Bell Laboratories as a Member of Technical Staff, from 1991 to 1994, where he had worked on network management. From 2009 to 2011, he served as the Director of the Institute of New Media and Communications. He is currently a Professor at SNU. He has been leading many industrial projects on 3G/4G/5G and the IoT connectivity supported by Korean industry. He has published more than 300 technical articles and holds more than 100 patents. He is a member of the National Academy of Engineering of Korea (NAEK). He was a recipient of the KICS Haedong Scholar Award, in 2012. He was President of the Korean Institute of Communications and Information Sciences (KICS). He has been serving as Chief Information Officer (CIO) of SNU. He was General Chair of the IEEE WCNC 2020 (Wireless Communication and Networking Conference). He was Director of the Asia-Pacific Region of the IEEE ComSoc. He is an Editor of the IEEE Network Magazine and IEEE Transactions on Vehicular Technology. He was TPC Chair of the IEEE VTC-Spring 2014, and General Chair of JCCI 2015, Co-Editor-in-Chief of the Journal of Communications and Networks (JCN), and on the Editorial Board of Computer Networks Journal and the IEEE Tran. on Wireless Communications.

***