Contents lists available at ScienceDirect

Optical Fiber Technology

journal homepage: www.elsevier.com/locate/yofte

Enhancing coordinated multi-point joint reception with dynamic bandwidth allocation

Guojun Zhu^{*}, Yunfeng Peng, Tonghui Ji

School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing, 100083, China

ARTICLE INFO

Dynamic bandwidth allocation

Keywords:

Joint reception

CoMP

C-RAN

TDM-PON

ABSTRACT

Coordinated Multipoint (CoMP) Joint Reception (JR) is a vital technology to mitigate uplink inter-cell interference (ICI) and can be facilitated by the centralized radio access network (C-RAN). In C-RAN architecture, base stations (BSs) are decoupled into the central unit (CU) and the distributed units (DUs). Meanwhile, a fronthaul (FH) is employed to connect CU and DUs. However, FH constructed by the time division multiplexing passive optical network (TDM-PON) significantly affects the performance of JR. The inherent polling mechanism of TDM-PON results in the asynchronous forwarding of JR data, which leads to high JR latency and a large amount of storage consumption. To alleviate the problem, we presented a novel dynamic bandwidth allocation (DBA) scheme, which classifies network data as JR and non-JR data and optimizes the data forwarding order of different types and DUs. A heuristic algorithm called Rearrange Sub-timeslot DBA (RS-DBA) was proposed to implement the DBA scheme. Simulation results indicate the proposed RS-DBA algorithm can reduce the average latency of JR data and storage resources consumption. In addition, the RS-DBA can make a better tradeoff between latency and efficiency.

1. Introduction

In the 5G network, a large number of small cells are deployed densely to cope with the explosion of mobile traffic [1]. The ultra-dense network (UDN) could raise system capacity and bandwidth efficiency, but it will incur severe inter-cell interference (ICI) which degrades the quality of service (QoS) of the cell edge user [2,3]. Coordinated Multi-Point (CoMP) is a fundamental solution to mitigate ICI by reconciling the data transmissions of multiple adjacent cells [4,5]. There are two kinds of CoMP techniques: coordinated scheduling (CS) which needs only the exchange of control information between base stations (BSs), and joint transmission/reception (JT/JR) which requires the BSs to share all the user data [6,7]. In addition, the emerging centralized radio access network (C-RAN) is believed to be a promising technology to enhance network collaboration [8-10]. By splitting the functionality of BSs, C-RAN assembles their baseband processing capabilities into a central unit (CU), which can achieve localized and quick information exchange [11]. Although the C-RAN facilitates data sharing in CoMP, the data transmission from distributed units (DUs) to CU is still a heavy burden for the fronthaul (FH) [12,13]. In particular, JT/JR demands to deliver multiple copies of data for a user, which immensely increases the transmission overhead. For the downlink, only one copy of JT data is needed to broadcast to DUs, thus scarcely affecting the load of FH. In the uplink, however, all copies of JR data must be transmitted to

CU through the FH, which will inevitably lead to massive bandwidth consumption and high latency.

To promote the uplink CoMP performance, some efforts have been made toward reducing the FH bandwidth consumption owing to JR [14-19]. By compressing the data in DUs or selecting an appropriate functional split, the FH traffic for JR can be decreased effectively. Nevertheless, little attention has been paid to the JR latency significantly influenced by the access method used in FH. The typical FH is constructed from the passive optical network (PON), either based on time division multiplexing (TDM) or wavelength division multiplexing (WDM) [12,20]. The WDM-PON assigns each DU a wavelength and thus can simultaneously transmit the JR data of a user in different DUs to CU. Instead, different DUs must forward the JR data of a user in sequence in the TDM-PON. The copy of JR data of a user that reaches CU early should be held in the cache to wait for those later ones [19]. Because the joint processing will not be executed until all copies of the JR data of a user are received, the JR has to consume considerable storage resources and experience high latency.

In this work, we consider reducing the storage consumption and JR latency by optimizing the bandwidth allocation of the TDM-PON-based FH. We present a novel dynamic bandwidth allocation (DBA) scheme to subdivide the upstream timeslot of each DU into two sub-timeslots for

https://doi.org/10.1016/j.yofte.2022.103207

Received 11 August 2022; Received in revised form 6 November 2022; Accepted 18 December 2022



Regular Articles





Corresponding author. E-mail address: zhgjuon@163.com (G. Zhu).

Available online 28 December 2022 1068-5200/© 2022 Elsevier Inc. All rights reserved.

transmitting JR and non-JR data, respectively. By reordering the subtimeslots of all DUs, the proposal will diminish the transmission interval between different copies of JR data. We conducted simulations in two cases, where the varying numbers of DUs and edge user equipment (UE) ratios were set. The results show that the proposed DBA scheme can decrease the use of storage resources and the average latency of JR data.

The remainder of the paper is organized as follows. Section 2 presents related works. Section 3 describes the network model, where JR in TDM-PON-based C-RAN is illustrated and the storage consumption and latency problems are analyzed. In Section 4, we propose the novel DBA scheme as well as the corresponding heuristic algorithm. Simulations and results are shown in Section 5. Finally, Section 6 concludes this paper.

2. Related works

Due to the problem of multi-copy transmission, achieving an efficient JR with limited FH has received much attention. Some researchers have striven to reduce the FH bandwidth occupied by JR. In [14], Bai et al. considered a scenario where two users transmit signals to two DUs and proposed a hybrid decode compress-and-forward relay scheme to improve the available sum data rate of the users. The numerical results show that the hybrid scheme outperforms the whole and partial decode-and-forward methods, especially when the data rate of one user is low. In [15], Qi et al. focused on the concrete compression method and optimized the distributed Wyner-Ziv compression using an iterative algorithm to achieve minimal fronthaul traffic for a target user rate. In addition, the authors analyzed a more complex scenario where the single-user case is extended to a multi-user case. Unlike the compression-based scheme which could result in data loss, reconfiguring the functional split has been studied to implement low-rate and lossless FH transmission for JR [16-19]. In [16], Miyamoto et al. proposed a split physical (PHY) layer processing (SPP) architecture, which splits the BS functions between demodulation and channel decoding and enables uplink JR through the log likelihood ratio (LLR) combining. Because the number of LLR quantization bits is usually set to a pretty small value, the SPP can dramatically reduce the FH bandwidth with a negligible signal-noise ratio degradation. The authors also demonstrated the robust performance of SPP under different modulation and coding schemes (MCSs) [17]. Besides, an enhanced SSP was proposed to control the LLR quantization threshold adaptively, thereby further reducing the FH bandwidth of JR [18]. However, the split in the PHY layer will restrict the cooperation gain of C-RAN. In [19], Boviz et al. presented the dynamic functional split conception, in which the low-PHY split is used for uplink JR while the MAC-PHY split is used for regular uplink and downlink. Its main advantage is to achieve a tradeoff between the FH data rate and RAN centralization. Although the above works improve the JR in FH bandwidth utilization, they rarely consider how the FH affects the latency performance of JR. Therefore, in this paper, we investigate the JR latency in the TDM-PON-based FH.

The uplink latency in TDM-PON is usually closely related to the used bandwidth allocation method. In [21], Kramer et al. proposed Interleaved Polling with Adaptive Cycle Time (IPACT), which is a classical dynamic protocol for the Ethernet PON and is also the basis of most DBA schemes. Based on the IPACT protocol, the authors devised five DBA algorithms. Another pioneering protocol is Bandwidth Guarantee Polling (BGP), which divides upstream bandwidth into small equal bandwidth units called entries and maps each entry to an ONU [22]. The BGP scheme enables OLT to poll an ONU many times during an allocation cycle and thus can achieve differentiated services according to the service level agreement (SLA). For the multiservice provisioning over TDM-PON, Luo et al. provided an overview of DBA and mainly introduced the inter-ONU and intra-ONU allocation algorithms [23,24]. In [25], Mirahmadi et al. considered the DBA in hybrid optical wireless networks, where bandwidth requests can be obtained by predicting



Fig. 1. Schematic of JR in TDM-PON-based C-RAN.

the incoming traffic to ONUs. With the help of mobile scheduling information, the cooperative DBA (CO-DBA) scheme was proposed to reduce the latency in TDM-PON-based FH [26,27]. In [28], Uzawa et al. extended the usage of CO-DBA to a convergence TDM-PON, on which mobile fronthaul and Internet-of-Things sub-network coexist. To improve the number of ONUs accommodated in the CO-DBA scheme, Hisano et al. studied the issue of forwarding order control and proposed a two-stage optimization algorithm [29]. Note that these previous proposals designed for the general purpose do not consider the latency characteristic of a specific application. In this work, we focus on developing an appropriate DBA algorithm to improve the latency performance of JR.

3. Network model

3.1. JR in TDM-PON-based C-RAN

C-RAN is a typical 5G access network architecture, as shown in Fig. 1. With specified functional split options, C-RAN divides BSs into a CU and several DUs. The CU usually consists of some storage and computing resource entities and executes partial or whole baseband processing. DUs at cell sites implement the radio frequency (RF) function and the rest of baseband processing. Besides, a TDM-PON is used as FH to connect CU with DUs, where the optical line terminal (OLT) and the optical network units (ONUs) are attached to CU and DUs, respectively. Because ONUs share the bandwidth of the upstream channel, TDM-PON takes a pooling mechanism to evade uplink transmission collision [21]. When ONUs forward data to OLT, the transmission request information, either "REPORT" messages from ONUs or indirect mobile-scheduling information from DUs [26,30], is gathered in advance. OLT performs the DBA algorithm and then attains transmission time windows for ONUs. Once receiving the allocation result through transmission grant information such as "GATE" messages, each ONU forwards its uplink data in the assigned timeslot.

For the JR, the cell edge user transmits the same uplink data to CU through multiple paths, which improves transmission quality. An example of the procedure of JR is illustrated in Fig. 1. It is supposed that UE #k is a cell edge user, which can be recognized by measuring the received signal reference power (RSRP). When getting the permission for uplink transmission, UE #k sends its data to DU #i and DU #j by utilizing the same radio resources blocks (RBs) of these two DUs.

Table 1 Definitions of symbols

Symbol	Description
М	Number of DUs (ONUs)
Ν	Number of cell edge UEs
i(j)	Identifier of DUs (ONUs)
k	Identifier of cell edge UEs
Р	$M \times N$ connections matrix
р	Element in P, equals to one or zero
$p_{ik}(p_{jk})$	Condition of connection between UE $#k$ and DU
	#i(#j)
$UL_{ik}(UL_{jk})$	Uplink of UE $\#k$ through DU $\#i(\#j)$
Т	Uplink latency
T_R	Latency from UE to DU
T_{O}	Latency from DU to CU
$T_{ik}(T_{jk})$	T of $UL_{ik}(UL_{jk})$
$T_{R,ik}(T_{R,jk})$	T_R of $UL_{ik}(UL_{jk})$
$T_{O,ik}(T_{O,jk})$	T_O of $UL_{ik}(UL_{jk})$
ΔT_k	Latency difference between UL_{ik} and UL_{jk}
$t_{0,k}$	Average arrival time when the data of UE $\#k$
	reach ONU
$t_{ik}(t_{jk})$	Average arrival time when the data of UE $\#k$
	reach OLT through ONU #i(#j)
D_k	Data amount of UE $\#k$
ij	Identifier of virtualized single edge cell UE
$t_i(t_j)$	Average arrival time when the data in ONU $\#i(\#j)$
	reach OLT; the center time of timeslot allocated to
	ONU $\#i(\#j)$
ΔT_{ij}	JR latency difference for UE #ij
С	Storage consumption for JR data during a
_	transmission cycle
R_k	JR data latency for UE $\#k$
A_k	Latency constant for UE $\#k$
R_{ij}	JR data latency for UE #1
A_{ij}	Latency constant for UE $\#ij$
L	Average latency of all the JR data during a
4	transmission cycle
a _{ij}	Latency weight coefficient of UE #1
л	Constant term of transmission avala
start	Start time of transmission cycle
3 _i	Size of timesiot anocated to ONU #1

After receiving the raw data of UE #k, each DU performs some low layer processing functions determined by the chosen Next Generation Fronthaul Interface (NGFI) [31]. Here, we consider the 3GPP split option 7-2 [32]. Subsequently, the processed data are encapsulated as Ethernet frames and stored in attached ONUs to wait for their transmission time windows. If the timeslot of ONU #i is before that of ONU #j, data in ONU #i will be forwarded to OLT first. These earlier data are buffered in the storage entities of CU until the latter ones arrive. Once all copies of data are received, the CU executes related joint processing functions [33].

3.2. Storage and latency problems in JR

As mentioned in Section 1, TDM-PON brings negative effects on JR due to its channel sharing. Here, we investigate how the bandwidth allocation affects the JR performance, in terms of storage consumption and data latency. Table 1 shows all the symbols used in this section and their definitions.

It is assumed that *M* DUs and *N* cell edge UEs are in C-RAN. The identifier of DUs is $\{i(j) \in \mathbb{N}^+ \mid 1 \le i(j) \le M\}$ and that of cell edge UEs is $\{k \in \mathbb{N}^+ \mid 1 \le k \le N\}$. In addition, we define an $M \times N$ matrix *P*, which indicates the condition of connections between the UEs and the DUs. Each element *p* in matrix *P* is equal to one if its corresponding connection exists, and otherwise zero. In this work, we assumed that each cell edge UE connects only to two DUs with the highest estimated RSRP, because more serving DUs will occupy much FH bandwidth but yield little performance improvement [33–36]. Namely, $\sum_{i=1}^{M} p_{ik} = 2$.

As shown in Fig. 2, if cell edge UE #k participates in JR with the help of DU #i and DU #j, its data will be transmitted to CU through



Fig. 2. Analysis of uplink latency in JR.

uplinks UL_{ik} and UL_{jk} . Each uplink consists of a wireless link between UE and DU as well as an optical link between DU and CU. Therefore, the uplink delay can be expressed as

$$T = T_R + T_O,\tag{1}$$

where T_R is the transmission time from UE to DU and T_O is the transmission time from DU to CU. More narrowly, T_R comprises the air interface latency and data processing time in the DU. T_O includes DU-ONU interface latency, TDM-PON latency, and OLT-CU interface latency [37]. According to Eq. (1), latencies of UL_{ik} and UL_{jk} are given by

$$T_{ik} = T_{R,ik} + T_{O,ik},\tag{2}$$

$$T_{jk} = T_{R,jk} + T_{O,jk}.$$

3.2.1. Storage consumption

7

The latency difference between UL_{ik} and UL_{jk} is denoted as follow:

(3)

$$\Delta T_k = \left| T_{ik} - T_{jk} \right| \,. \tag{4}$$

Due to the fact that distances between cell edge UE and cooperative DUs are nearly equal, we assume the difference of air interface latency is negligible [38]. Moreover, we consider the data processing time is a constant for all DUs [37]. It means that

$$T_{R,ik} = T_{R,jk}.$$
(5)

Therefore, Eq. (4) is simplified to

$$\Delta T_k = \left| T_{O,ik} - T_{O,jk} \right|. \tag{6}$$

As the interface latencies are equal for all uplinks [37], we assume they are zeros so that the latency analysis can be simplified. Consequently, the value of ΔT_k will depend on TDM-PON latencies, which are decided by the DBA results.

Fig. 3 illustrates the TDM-PON latencies of different JR uplinks. The symbol $t_{0,k}$ denotes the average arrival time when the data of UE #k reach ONU #i and ONU #j. The data amount of UE #k is D_k . In addition, t_{ik} (t_{jk}) is the average arrival time when the data of UE #k reach OLT through ONU #i(#j). To facilitate the discussion, we use the center time of the data arrival period as the average data arrival time. Therefore, TDM-PON latencies of UL_{ik} and UL_{ik} can be expressed as

$$T_{O,ik} = t_{ik} - t_0, (7)$$

$$T_{O,ik} = t_{ik} - t_0. (8)$$

Fig. 3. JR upstream transmission in TDM-PON.

According to Eqs. (6), (7), and (8), the JR latency difference for UE $\#k \Delta T_k$ is also given by

$$\Delta T_k = \left| t_{ik} - t_{jk} \right|. \tag{9}$$

The above discussion analyzes the JR latency difference for a single cell edge UE. In fact, a pair of cooperative DUs usually support many edge UEs, which transmit their JR data to CU with the help of the same TDM-PON timeslots. Nevertheless, it is difficult to define the position of data of each UE in timeslots. Moreover, TDM-PON allocates timeslots to ONUs rather than the UEs. Therefore, to simplify the latency analysis, all the cell edge UEs of a cooperative DUs pair is treated as a single user.

Let UE #*ij* is the virtualized single user served by DU #*i* and DU #*j*. We assume D_{ij} is the data amount of UE #*ij*, and thus

$$D_{ij} = \begin{cases} \sum_{k \in \{k \mid p_{ik} p_{jk} = 1\}} D_k, & i \neq j \\ 0, & i = j \end{cases}$$
(10)

The symbol $t_i(t_j)$ is the average arrival time when the data in ONU #i(#j) reach the OLT. Therefore, being similar to Eq. (9), the JR latency difference for UE #ij is expressed as

$$\Delta T_{ij} = \left| t_i - t_j \right|. \tag{11}$$

We define the storage consumption for JR data during a TDM-PON transmission cycle as

$$C = \sum_{i=1}^{M} \sum_{j=i+1}^{M} D_{ij} \Delta T_{ij},$$
(12)

namely,

$$C = \sum_{i=1}^{M} \sum_{j=i+1}^{M} D_{ij} \left| t_i - t_j \right|.$$
(13)

3.2.2. Average latency

As mentioned in Section 1, because the earlier copies of JR data have to wait for later ones, JR data latency is determined by the time when the latest duplicate reaches CU. Therefore, the JR data latency of UE #k can be expressed as

$$R_k = \max(T_{ik}, T_{jk}). \tag{14}$$

Considering Eqs. (2), (3), (5), (7) and (8), the Eq. (14) is transformed into

$$R_k = \max(t_{ik}, t_{ik}) + A_k, \tag{15}$$

where A_k is a constant and can be expressed as

$$A_k = T_{R,ik} - t_{0,k}.$$
 (16)

Similarly, the JR data latency of UE #ij is given by

$$R_{ij} = \max(t_i, t_j) + A_{ij}.$$
(17)

We define the average latency of all JR data during a TDM-PON transmission cycle as

$$L = \sum_{i=1}^{M} \sum_{j=i+1}^{M} d_{ij} R_{ij},$$
(18)

where d_{ij} is the latency weight coefficient of UE #*ij* and can be expressed as

$$d_{ij} = D_{ij} / \sum_{i=1}^{M} \sum_{j=i+1}^{M} D_{ij}.$$
(19)

Namely,

$$L = \sum_{i=1}^{M} \sum_{j=i+1}^{M} d_{ij} \max(t_i, t_j) + \sum_{i=1}^{M} \sum_{j=i+1}^{M} d_{ij} A_{ij}.$$
 (20)

We use A to represent the constant term in Eq. (20),

$$A = \sum_{i=1}^{M} \sum_{j=i+1}^{M} d_{ij} A_{ij}.$$
 (21)

Therefore,

$$L = \sum_{i=1}^{M} \sum_{j=i+1}^{M} d_{ij} \max(t_i, t_j) + A.$$
 (22)

3.2.3. Optimization problem

In this work, we aim to minimize the storage consumption and latency of JR data. Eqs. (13) and (22) are objective functions and the optimization problem can be expressed as min $C\& \min L$.

Next, we discuss the constraints for the variables $\{t_i \mid i = 1, 2, ..., M\}$. We assume that the start time of the TDM-PON transmission cycle is t_{start} and the size of the timeslot allocated to ONU #i is S_i . For convenience in analyzing, the guard time between two consecutive timeslots is not considered here. Firstly, each variable t_i should be in the TDM-PON transmission cycle, thus

$$t_{start} < t_i < \sum_{x=1}^M S_x.$$
⁽²³⁾

Also, each t_i is a discrete variable, and its value is constrained by

$$t_i \in \left\{ t \middle| t = t_{start} + \frac{1}{2} S_i + \sum_{x \in X} S_x, X \subseteq \left\{ y \in \mathbf{N}^+ \mid 1 \le y \le M \land y \ne i \right\} \right\}.$$
(24)

To avoid the timeslot collision, any two different variables t_i and t_j are constrained by

$$\left| t_{i} - t_{j} \right| \in \left\{ \Delta t \left| \Delta t = \frac{1}{2} S_{i} + \frac{1}{2} S_{j} + \sum_{x \in X} S_{x}, X \subseteq \right. \right. \\ \left\{ y \in \mathbf{N}^{+} \mid 1 \le y \le M \land y \ne i \land y \ne j \right\} \right\}.$$

$$(25)$$

It is clear that each feasible solution to the optimization problem is a set of times. In a transmission cycle, the set of times indicates a timeslot arrangement, which corresponds to an ONUs sequence. Therefore, to solve this optimization problem, we need to find the optimum ONUs sequence. It is similar to the traveling salesman problem (TSP) [39], in which an optimum citie sequence is sought to make a minimal-cost tour. The difference between them is that the objective function of TSP is a sum function of Euclidean distance, while the objective functions of our optimization problem are mathematical operations like Eqs. (13) and (22). Therefore, without loss of generality, the proposed optimization problem is an NP-Hard problem.

Fig. 4. Comparison between (a) the conventional DBA scheme and (b) the proposed DBA scheme.

4. DBA scheme and algorithm

4.1. Rearrange sub-timeslot dynamic bandwidth allocation

To solve the above problem, a novel DBA scheme is presented as shown in Fig. 4(b). The uplink data in ONUs are classified as JR data generated by cell edge UEs and non-JR data generated by cell center UEs. OLT allocates bandwidth to each ONU and each type of data depending on the following strategies: (1) neighboring timeslots should be assigned to coordinated ONUs as much as possible; (2) the timeslot of each ONU should be divided into two sub-timeslots, one is for JR data and the other is for non-JR data; (3) sub-timeslots used for a JR pair should be as close as possible; (4) all sub-timeslots for JR should be in a transmission cycle and as far forward as possible. The aim of Rule (1) is to reduce storage consumption by decreasing the JR latency difference. Rule (2) is used in each timeslot for defining two isolated sub-timeslots to match with different types of data. With the help of Rule (3), a further reduction of JR latency difference can be achieved. Rule (4) is aimed at forwarding JR data as earlier as possible so that the average latency of JR data decreases. A simple example is illustrated in Fig. 4, where the proposed DBA scheme obtains significant gains compared to the conventional DBA scheme.

Here, we propose a heuristic algorithm called Rearrange Sub-timeslot DBA (RS-DBA) to realize above DBA scheme. The RS-DBA algorithm is shown in Table 2 and consists of two parts. In Part A, we calculate the size of sub-timeslot allocated to each type of data in each ONU. Considering the fairness among ONUs, we compute bandwidth allocated to each ONU using the max-min fairness algorithm (step 4). In each ONU, when allocated bandwidth satisfies its requirement, the allocation for each type of data is equal to demand (step 8). Otherwise, the requirement of JR data is met first (step 10). According to this bandwidth allocation, the size of each sub-timeslot is computed (step 13). In Part B, we determine the order of sub-timeslots for all ONUs basing on the amount of JR data in each ONU. Firstly, all ONUs are sorted into a descending order by the amount of JR data (step 14). The aim is to prioritize the ONU with more JR data. Then, according to the sorting result, an appropriate order number of timeslot is assigned to each ONU (step 16). The mapping method is to start in the middle number and alternately select the front or rear one, which is intended to satisfy the Rule (1). After that, following Rules (3) and (4), we determine which sub-timeslot is allocated for transmitting JR data in each timeslot (steps 17-21). For the front timeslots, the first and second

Table	2
-------	---

Descrip	– ption of RS-DBA algorithm.
Rearr (RS-I	ange sub-timeslot dynamic bandwidth allocation
Part	A:
1:	if the sum of requirement is less than limited bandwidth. do
2:	Allocation for each ONU is equal to requirement.
3:	else do
4:	Allocation for each ONU is computed by using max-min fairness algorithm.
5:	end if
6:	for each ONU, do
7:	if allocation is equal to requirement, do
8:	For each type of data, allocation is equal to its request.
9:	else do
10:	For JR data, allocation is equal to its request; the rest
	is allocated to non-JR data.
11:	end if
12:	end for
13:	According to this bandwidth allocation, computing the size
Dant	of sub-timesiot allocated to each type of data in each ONU.
14.	D . M ONLIG are corted into descending order by the amount of
14.	JR data
15	for <i>i</i> -th ONU do
16:	Computing $a = M/2 + 1$, $b = i/2 $, $c = mod(i, 2)$.
	Computing $d = a - b \cdot (-1)^c$. The <i>d</i> -th timeslot is alloca-
	ted to <i>i</i> -th ONU.
17:	if d is less than a, do
18:	The first sub-timeslot is allocated to non-JR data, and
	the second sub-timeslot is allocated to JR data.
19:	else do
20:	The first sub-timeslot is allocated to JR data, and the
	second sub-timeslot is allocated to non-JR data.
21:	end if
22:	end for
23:	Obtaining the size and order of sub-timeslots for all ONUs.

sub-timeslots are assigned to non-JR and JR data respectively. For the rear timeslots, the reverse applies. According to the results of *Part A* and *Part B*, we will obtain the size and order of sub-timeslots for all ONUs (step 23).

4.2. Time complexity analysis

The number of ONUs is *M*. For *Part A*, the time complexity of bandwidth allocation for ONUs is constrained by the max–min fairness algorithm. In the worst case, the time complexity of the max–min fairness algorithm is $O(M^2)$. The time complexity of bandwidth allocation for different types of data is O(M). The time complexity of computing the size of the sub-timeslot is O(M). For *Part B*, the time complexity of a sorting algorithm is usually $O(M^2)$. For each loop, the time complexity of determining the order number of the timeslot of an ONU is O(M). The time complexity of determining the data type of each sub-timeslot in a timeslot is O(M). The loop will be executed *M*, so the time complexity of determining the order number of all sub-timeslots is $O(M^2)$. Therefore, the total time complexity of RS-DBA is $O(M^2)$.

5. Simulation and discussion

In this section, we verified the effectiveness of the proposed RS-DBA algorithm via numerical simulation. First, the simulation setup and related parameters are presented. Subsequently, the simulation results are evaluated by comparing the proposal with methods in [27,40]. The scheme in [40] is called mobile-PON, which develops a unified scheduler for matching each RB to a quantized transmission block. The mobile-DBA proposed in [27] is a kind of cooperative DBA (CO-DBA), where the mobile scheduling information is translated to the requests of ONUs. For the convenience of expression, we use M-PON and CO-DBA to represent them, respectively. The purpose of comparing RS-DBA Table 3

Parameter	Value
Simulation cycle	10 s
Number of DUs	[3,16]
Average number of UEs per DU	10
UL bandwidth per DU	20 MHz
TTI length	1 ms
Number of RBs per 1/2 TTI	100
Number of subcarriers per RB	12
Number of RE per subcarrier	7
Modulation index	8 bits
Average data rate of UE	10 Mbps
Edge UE ratio	[0.05,0.45]
Quantization bits	4 bits
PON bandwidth	10 Gbps
Transmission cycle	0.5 ms
Grant cycle	50 µs
Number of RBs per TB	10
Guard time	500 ns

and CO-DBA is to prove that our proposal can improve the latency performance of JR. In addition, because the M-PON can also enhance the JR, we compare RS-DBA with it to investigate the advantages of our proposal.

5.1. Simulation setup

The simulation setup is configurated by referring to many works [28,29,40]. We assume the number of DUs ranges from 3 to 16. There is an average of ten UEs in each DU. We consider the uplink bandwidth of each DU is 20 MHz, where 100 RBs are scheduled during one half of transport time interval (TTI). The number of subcarriers per RB is 12. Each subcarrier consists of 7 resource elements (REs). We assume 64 OAM is used in all REs and thus set the modulation index to 8 bits. The data that each UE generates is based on Poisson distribution, and the average data rate of each UE is 10 Mbps. The edge UE ratio, which is the ratio of the number of cell edge UEs to that of all UEs, ranges from 0.05 to 0.45. The quantization bits in FH are assumed to 4 bits. We consider 10G-EPON in the simulations, where the transmission cycle is 0.5 ms and the grant cycle is set to 50 μ s. In M-PON, the number of RBs accommodated on a transport block (TB) is 10. The guard time is set to 500 ns. The settings of parameters used in simulations are summarized in Table 3.

5.2. Results and discussion

In this work, we implemented the numerical simulations using MATLAB 2018b. The simulations were conducted in two cases, where the number of DUs and the edge UE ratio were considered the variables. The average packet latency, efficiency, and storage consumption were evaluated to compare the performance of the three methods.

Figs. 5, 6, and 7 show simulation results of the first case, in which the number of DUs is at the range from 3 to 16 and the edge UE ratio is set to 0.3. As shown in Fig. 5, with the number of DUs increasing, the average packet latencies for different schemes and different types of data are growing. It is easy to understand that the more DUs in the network, the more data are forwarded, leading to higher latency in FH. For the CO-DBA, the slope of average packet latencies of JR data is similar to that of non-JR data. It is due to the fact that different types of data are regarded as homogeneous data in the CO-DBA scheme. Besides, it is shown that the latency of JR data is large than that of non-JR data, which is consistent with the aforementioned analysis in Section 1. We also find that the proposed RS-DBA and the M-PON can reduce the average latency of JR data at the expense of increased non-JR data latency. The difference between RS-DBA and M-PON is that the latter achieves lower JR data latency and higher non-JR data

Fig. 5. Average packet latency against the number of DUs.

Fig. 6. Efficiency against the number of DUs.

latency. This is because M-PON prioritizes the forwarding of all JR data and postpones that of all non-JR data in a transmission cycle, but RS-DBA just performs such forwarding control in each ONU transmission window. In addition, it is observed that the slopes of average packet latencies of JR data for RS-DBA and M-PON are smaller than that for CO-DBA. Correspondingly, the slopes of average latencies of non-JR data for RS-DBA and M-PON are greater than that for CO-DBA. The reason is RS-DBA and M-PON schemes prioritize the transmission of JR data. As a result of that, their performance depends on the amount of JR data. The more DUs in the network, the more JR data are forwarded, leading to better performance for RS-DBA and M-PON. We also find that the JR data average latency of RS-DBA is close to that of CO-DBA when there are a few DUs in the network. This is because the random ONU forwarding order used in CO-DBA approximates the optimal ONU forwarding order obtained by RS-DBA. Therefore, the RSS-DBA does not achieve a distinctive effect in the reduction of JR data latency. Moreover, as shown in Fig. 5, the average latencies of non-JR data for RS-DBA and M-PON will reach the FH latency limitation (150 µs [40]) when the number of DUs is 15. Therefore, the RS-DBA and M-PON are not suitable for the massive-DU scenario.

In Fig. 6, we observe efficiencies of RS-DBA and CO-DBA decrease with the number of DUs increases, but that of M-PON remains unchanging. It is due to the fact that more DUs share the FH, more guard time overheads are required in CO-DBA and RS-DBA during a grant cycle.

Fig. 7. Storage consumption against the number of DUs.

Fig. 8. Average packet latency against the edge UE ratio.

However, the efficiency of M-PON is affected by the size of TB and guard time, which are constants. Namely,

$$Efficiency = \frac{TB/PON \ bandwidth}{Guard \ time + TB/PON \ bandwidth}.$$
 (26)

With the values set in Table 3, the efficiency of the M-PON scheme is calculated to be about 84%, which is in accord with the simulation results. We also find CO-DBA and RS-DBA have identical efficiency curves. The reason is that these two schemes have the same bandwidth allocation except for the forwarding order of ONUs. In other word, the size of timeslots used for transmitting data in CO-DBA and RS-DBA are exactly equal, resulting in the same efficiency. In addition, we can observe the efficiency of RS-DBA is greater than that of M-PON when the number of DUs is under 15. This is because the TB used in M-PON is smaller than the ONU transmission window used in RS-DBA, leads to more guard time overheads in M-PON which consume plenty of PON bandwidth. When the number of DUs exceeds 15, too many DUs share the FH, resulting in the ONU transmission window is shorter than TB.

In Fig. 7, it is shown that the average storage consumption for JR data grows with an increasing number of DUs. It is due to the fact that the more DUs in the network, the more JR data are generated, which leads to more storage consumption for JR. We observe the average storage consumptions of RS-DBA and M-PON are lower than that of CO-DBA. The reason is that RS-DBA and M-PON prioritize the transmission

Fig. 9. Efficiency against the edge UE ratio.

Fig. 10. Storage consumption against the edge UE ratio.

of JR data, thus making JR data reach CU in a relatively concentrated period. As a result of that, the occupied storage is released early. In addition, the average storage consumption of RS-DBA is higher than that of M-PON. This is because RS-DBA prioritizes the forwarding of JR data in each ONU transmission window rather than a transmission cycle. We also find the standard deviation of storage consumption is variable. It indicates the steadiness of JR data forwarding control for different schemes. The more JR data in the network, the more difficult for forwarding control, leading to a bigger standard deviation. The CO-DBA scheme never controls the forwarding of JR data, and thus has the largest standard deviation. The RS-DBA prioritizes the forwarding of JR data in each local transmission period. However, the M-PON implements forwarding control in a global transmission cycle. Therefore, as shown in Fig. 7, the standard deviation of M-PON is smaller than that of RS-DBA.

Figs. 8, 9, and 10 show simulation results of the second case, where the edge UE ratio ranges from 0.05 to 0.45 and the number of DUs is set to 10. In Fig. 8, we can find the average latency of different types of data is growing with the edge UE ratio increases. The reason is that the edge UE pours two copies of JR data into FH. Therefore, the more edge UEs in the network, the more data are generated. We also observe the general trend of curves in Fig. 8 is similar to that in Fig. 5. The main difference is the slopes of average packet latencies of JR data for RS-DBA and M-PON are greater than that for CO-DBA. This is because

Fig. 11. Storage occupancy in the duration from 5400 to 6400 µs, the number of DUs is 10, and the edge UE ratio is 0.30.

the growth of JR data is higher than that of overall data as the edge UE ratio increases. Similarly, Figs. 9 and 6 are alike in the features of curves. The main difference is that the efficiencies of RS-DBA and CO-DBA are almost changeless. It is due to the fact that the efficiency of DBA scheme depends on the guard time overhead which is related to the number of DUs rather than the edge UE ratio. The contrast between Figs. 10 and 7 shows that the standard deviations of different schemes tend to be equal as the edge UE ratio rises. The reason is JR data will be the dominant data when the value of the edge UE ratio is large, which means the effectiveness of RS-DBA and M-PON will decline.

Fig. 11 illustrates the situation of storage occupancy in a simulation duration when the time is from 5400 to 6400 μ s. The number of DUs and the edge UE ratio are set to 10 and 0.30, respectively. As shown in Fig. 11, the storage occupancy peaks of RS-DBA and M-PON are higher than that of CO-DBA. The phenomenon verifies the aforementioned analysis that JR data of RS-DBA and M-PON reach CU in a relatively concentrated period. We observe the curves of M-PON return to zeros faster than that of RS-DBA and CO-DBA. It indicates that M-PON accomplishes the forwarding of JR data faster and thus achieves lower average latency of JR data, which is in accord with the previous simulation results. We also find that the area covered by the curve of RS-DBA is smaller than that of CO-DBA, although the storage occupancy peak of RS-DBA is higher. It means that RS-DBA can achieve low storage consumption as compared with CO-DBA by utilizing virtualization technology.

Figs. 12 and 13 show the comparison between the results of RS-DBA and the optimal solution from the full search method. Considering the complex calculation of the full search method, we conduct this comparison in a moderate network scenario where the number of DUs ranges from 3 to 8. The edge UE ratio is set to 0.3. We can observe the results of RS-DBA gradually approach the optimal values with the increased number of DUs, for both the average packet latency and the storage consumption. It indicates that we can adopt a mixed strategy to improve the JR performance. For the small-scale network, the full search method is used to obtain optimal gains. When the network scale gets larger, RS-DBA should be employed to achieve sub-optimal gains with low computational complexity.

6. Conclusion

In this paper, we considered CoMP JR in TDM-PON-based C-RAN and investigated the impact of FH on JR in terms of latency and storage consumption. To improve the performance of JR, we presented a novel DBA scheme that classifies uplink data and controls the forwarding of different types of data. A practical heuristic algorithm called RS-DBA was proposed to realize the DBA scheme, where timeslots are subdivided and appropriate forwarding order is arranged. We evaluated the

Fig. 13. Storage consumption against the number of DUs.

effectiveness of the proposed RS-DBA algorithm by comparing it with CO-DBA and M-PON algorithms. Simulation results showed that RS-DBA can reduce the average latency of JR data by up to 50% in the best situation as compared with CO-DBA. The decreased degree of average storage consumption in CU ranges from 4% to 40%. In addition, the RS-DBA algorithm can achieve more efficiency than M-PON. The maximum efficiency improvement can reach 10%.

CRediT authorship contribution statement

Guojun Zhu: Conceptualization, Investigation, Formal analysis, Methodology, Software, Writing – original draft. **Yunfeng Peng:** Formal analysis, Resources, Data curation, Supervision, Writing – review & editing. **Tonghui Ji:** Investigation, Formal analysis, Software, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- X. Ge, S. Tu, G. Mao, C.-X. Wang, T. Han, 5G ultra-dense cellular networks, IEEE Wirel. Commun. 23 (1) (2016) 72–79, http://dx.doi.org/10.1109/MWC. 2016.7422408.
- [2] O.T.H. Alzubaidi, M.N. Hindia, K. Dimyati, K.A. Noordin, A.N.A. Wahab, F. Qamar, R. Hassan, Interference challenges and management in B5G network design: A comprehensive review, Electronics 11 (18) (2022) 2842, http://dx. doi.org/10.3390/electronics11182842.
- [3] S. Andreev, V. Petrov, M. Dohler, H. Yanikomeroglu, Future of ultra-dense networks beyond 5G: Harnessing heterogeneous moving cells, IEEE Commun. Mag. 57 (6) (2019) 86–92, http://dx.doi.org/10.1109/MCOM.2019.1800056.
- [4] B.U. Kazi, G.A. Wainer, Next generation wireless cellular networks: Ultradense multi-tier and multi-cell cooperation perspective, Wirel. Netw. 25 (2019) 2041–2064.
- [5] S. Mukherjee, D. Kim, J. Lee, Base station coordination scheme for multi-tier ultra-dense networks, IEEE Trans. Wireless Commun. 20 (11) (2021) 7317–7332, http://dx.doi.org/10.1109/TWC.2021.3082625.
- [6] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, V. Jungnickel, Coordinated multipoint: Concepts, performance, and field trial results, IEEE Commun. Mag. 49 (2) (2011) 102–111, http://dx.doi.org/10. 1109/MCOM.2011.5706317.
- G. Song, W. Wang, D. Chen, T. Jiang, KPI/KQI-Driven coordinated multipoint in 5G: Measurements, field trials, and technical solutions, IEEE Wirel. Commun. 25 (5) (2018) 23–29, http://dx.doi.org/10.1109/MWC.2018.1800041.
- [8] H. Touati, H. Castel-Taleb, B. Jouaber, S. Akbarzadeh, Model-based optimization for JT CoMP in C-RAN, in: IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPS, 2019, pp. 403–409, http://dx. doi.org/10.1109/INFCOMW.2019.8845102.
- [9] F. Zanferrari Morais, C. André da Costa, A.M. Alberti, C. Bonato Both, R. da Rosa Righi, When SDN meets C-RAN: A survey exploring multi-point coordination, interference, and performance, J. Netw. Comput. Appl. 162 (2020) 102655, http://dx.doi.org/10.1016/j.jnca.2020.102655.
- [10] M. Elhattab, M. Kamel, W. Hamouda, Edge-aware remote radio heads cooperation for interference mitigation in heterogeneous C-RAN, IEEE Trans. Veh. Technol. 70 (11) (2021) 12142–12157, http://dx.doi.org/10.1109/TVT.2021. 3111148.
- [11] L.M.P. Larsen, A. Checko, H.L. Christiansen, A survey of the functional splits proposed for 5G mobile crosshaul networks, IEEE Commun. Surv. Tutor. 21 (1) (2019) 146–172, http://dx.doi.org/10.1109/COMST.2018.2868805.
- [12] A. Pizzinat, P. Chanclou, F. Saliou, T. Diallo, Things you should know about fronthaul, J. Lightwave Technol. 33 (5) (2015) 1077–1083, http://dx.doi.org/ 10.1109/JLT.2014.2382872.
- [13] I.A. Alimi, A.L. Teixeira, P.P. Monteiro, Toward an efficient C-RAN optical fronthaul for the future networks: A tutorial on technologies, requirements, challenges, and solutions, IEEE Commun. Surv. Tutor. 20 (1) (2018) 708–769, http://dx.doi.org/10.1109/COMST.2017.2773462.
- [14] J. Bai, Q. Liang, C. Huang, S. Shao, Y. Tang, A simple transmission scheme for coordinated multipoint uplink transmission with limited fronthaul, in: 2016 IEEE 84th Vehicular Technology Conference, VTC-Fall, 2016, pp. 1–5, http: //dx.doi.org/10.1109/VTCFall.2016.7880876.

- [15] Y. Qi, M.Z. Shakir, M.A. Imran, K.A. Qaraqe, A. Quddus, R. Tafazolli, Fronthaul data compression for uplink CoMP in cloud radio access network (C-RAN), Trans. Emerg. Telecommun. Technol. 27 (10) (2016) 1409–1425, http://dx.doi.org/10. 1002/ett.3088.
- [16] K. Miyamoto, S. Kuwano, J. Terada, A. Otaka, Split-PHY processing architecture to realize base station coordination and transmission bandwidth reduction in mobile fronthaul, in: 2015 Optical Fiber Communications Conference and Exhibition, OFC, 2015, pp. 1–3, http://dx.doi.org/10.1364/OFC.2015.M2J.4.
- [17] K. Miyamoto, S. Kuwano, T. Shimizu, J. Terada, A. Otaka, Performance evaluation of ethernet-based mobile fronthaul and wireless comp in split-PHY processing, J. Optical Commun. Netw. 9 (1) (2017) A46–A54, http://dx.doi.org/ 10.1364/JOCN.9.000A46.
- [18] K. Miyamoto, S. Ibi, T. Shimizu, J. Terada, A. Otaka, S. Sampei, Unified design of LLR quantization and joint reception for mobile fronthaul bandwidth reduction, in: 2017 IEEE 85th Vehicular Technology Conference, VTC Spring, 2017, pp. 1–5, http://dx.doi.org/10.1109/VTCSpring.2017.8108426.
- [19] D. Boviz, Y. El Mghazli, Fronthaul for 5G: Low bit-rate design enabling joint transmission and reception, in: IEEE Globecom Workshops, GC Wkshps, 2016, pp. 1–6, http://dx.doi.org/10.1109/GLOCOMW.2016.7848911.
- [20] F. Effenberger, D. Cleary, O. Haran, G. Kramer, R.D. Li, M. Oron, T. Pfeiffer, An introduction to PON technologies [topics in optical communications], IEEE Commun. Mag. 45 (3) (2007) S17–S25, http://dx.doi.org/10.1109/MCOM.2007. 344582.
- [21] G. Kramer, B. Mukherjee, G. Pesavento, IPACT a dynamic protocol for an ethernet PON (EPON), IEEE Commun. Mag. 40 (2) (2002) 74–80, http://dx. doi.org/10.1109/35.983911.
- [22] M. Ma, Y. Zhu, T. Cheng, A bandwidth guaranteed polling MAC protocol for ethernet passive optical networks, in: IEEE INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428), Vol. 1, 2003, pp. 22–31, http://dx.doi.org/10.1109/ INFCOM.2003.1208655.
- [23] Y. Luo, N. Ansari, Bandwidth allocation for multiservice access on EPONs, IEEE Commun. Mag. 43 (2) (2005) S16–S21, http://dx.doi.org/10.1109/MCOM.2005. 1391498.
- [24] Y. Luo, S. Yin, N. Ansari, T. Wang, Resource management for broadband access over time-division multiplexed passive optical networks, IEEE Network 21 (5) (2007) 20–27, http://dx.doi.org/10.1109/MNET.2007.4305168.
- [25] M. Mirahmadi, A. Shami, Traffic-prediction-assisted dynamic bandwidth assignment for hybrid optical wireless networks, Comput. Netw. 56 (1) (2012) 244–259, http://dx.doi.org/10.1016/j.comnet.2011.08.018.
- [26] T. Tashiro, S. Kuwano, J. Terada, T. Kawamura, N. Tanaka, S. Shigematsu, N. Yoshimoto, A novel DBA scheme for TDM-PON based mobile fronthaul, in: Optical Fiber Communication Conference, 2014, pp. 1–3, http://dx.doi.org/10. 1364/OFC.2014.Tu3F.3.
- [27] H. Uzawa, H. Nomura, T. Shimada, D. Hisano, K. Miyamoto, Y. Nakayama, K. Takahashi, J. Terada, A. Otaka, Practical mobile-DBA scheme considering data arrival period for 5G mobile fronthaul with TDM-PON, in: European Conference on Optical Communication, 2017, pp. 1–3, http://dx.doi.org/10.1109/ECOC. 2017.8345831.
- [28] H. Uzawa, K. Honda, H. Nakamura, Y. Hirano, K.-i. Nakura, S. Kozaki, J. Terada, Dynamic bandwidth allocation scheme for network-slicing-based TDM-PON toward the beyond-5G era, J. Opt. Commun. Netw. 12 (2) (2020) A135–A143, http://dx.doi.org/10.1364/JOCN.12.00A135.
- [29] D. Hisano, Y. Nakayama, Two-stage optimization of uplink forwarding order with cooperative DBA to accommodate a TDM-PON-based fronthaul link, J. Opt. Commun. Netw. 12 (5) (2020) 109–119, http://dx.doi.org/10.1364/JOCN. 384367.
- [30] H. Nomura, H. Ujikawa, H. Uzawa, H. Nakamura, J. Terada, Novel DBA scheme integrated with SR- and CO-DBA for multi-service accommodation toward 5G beyond, in: European Conference on Optical Communication, 2019, pp. 1–4, http://dx.doi.org/10.1049/cp.2019.0873.
- [31] IEEE Standard for Packet-based Fronthaul Transport Networks, IEEE Std 1914.1-2019, 2020, pp. 1–94, http://dx.doi.org/10.1109/IEEESTD.2020.9079731.
- [32] 3GPP, R3-173402-overall of proposed L1 processing diagram, Rev3 (2017).
- [33] K. Miyamoto, N. Shibata, S. Kuwano, J. Terada, A. Otaka, Wireless performance and mobile fronthaul bandwidth of uplink joint reception with LLR combining in split-PHY processing, J. Commun. Netw. 20 (6) (2018) 536–545, http://dx. doi.org/10.1109/JCN.2018.000086.
- [34] I. Chih-Lin, J. Huang, R. Duan, C. Cui, J. Jiang, L. Li, Recent progress on C-RAN centralization and cloudification, IEEE Access 2 (2014) 1030–1039, http: //dx.doi.org/10.1109/ACCESS.2014.2351411.
- [35] P. Shang, L. Zhang, M. You, Y. Yang, Q. Zhang, Performance of uplink joint reception CoMP with antenna selection for reducing complexity in LTE-A systems, in: IEEE Wireless Communications and Networking Conference, WCNC, 2015, pp. 977–982, http://dx.doi.org/10.1109/WCNC.2015.7127602.
- [36] T.H. Jacobsen, R. Abreu, G. Berardinelli, K.I. Pedersen, I.Z. Kovács, P. Mogensen, Multi-cell reception for uplink grant-free ultra-reliable low-latency communications, IEEE Access 7 (2019) 80208–80218, http://dx.doi.org/10.1109/ACCESS. 2019.2923324.

- [37] N.P. Anthapadmanabhan, A. Walid, T. Pfeiffer, Mobile fronthaul over latencyoptimized time division multiplexed passive optical networks, in: IEEE International Conference on Communication Workshop, ICCW, 2015, pp. 62–67, http://dx.doi.org/10.1109/ICCW.2015.7247156.
- [38] M. Kamel, W. Hamouda, A. Youssef, Ultra-dense networks: A survey, IEEE Commun. Surv. Tutor. 18 (4) (2016) 2522–2545, http://dx.doi.org/10.1109/ COMST.2016.2571730.
- [39] M. Dorigo, L. Gambardella, Ant colony system: A cooperative learning approach to the traveling salesman problem, IEEE Trans. Evol. Comput. 1 (1) (1997) 53–66, http://dx.doi.org/10.1109/4235.585892.
- [40] S. Zhou, X. Liu, F. Effenberger, J. Chao, Low-latency high-efficiency mobile fronthaul with TDM-PON (mobile-PON), J. Opt. Commun. Netw. 10 (1) (2018) A20–A26, http://dx.doi.org/10.1364/JOCN.10.000A20.