Hierarchical Scheduling for Suppression of Fronthaul Delay in C-RAN with Dynamic Functional Split

Zdenek Becvar, Pavel Mach, Mohammed Elfiky, and Mitsuo Sakamoto

Abstract

The cloud radio access network (C-RAN) can potentially reduce a network's deployment cost and energy consumption. However, a connection between a centralized baseband unit and distributed remote radio heads, known as a fronthaul, introduces an additional delay to both control and user planes. This delay is a serious limiting factor for radio resource management functionalities, such as scheduling, because the radio resources are assigned to users according to outdated channel quality information if the fronthaul delay is non-zero. This article provides an overview of existing scheduling approaches suitable for C-RAN and identifies their potential limitations. Based on these limitations, we outline a framework for hierarchical scheduling. The hierarchical scheduling mitigates a negative impact of the fronthaul delay on the throughput of non-cell-edge users and enables efficient retransmission of erroneous data. Besides, cell-edge users can still benefit from interference mitigation techniques requiring centralized control. We compare individual scheduling approaches and show that hierarchical scheduling increases the network throughput (by up to 26 percent) and reduces the number of retransmissions with respect to the existing solutions.

Introduction

Mobile networks face challenges related to a reduction of a deployment cost and an energy consumption. To meet these challenges, a cloud radio access network (C-RAN) concept has been introduced [1]. The architecture of C-RAN incorporates a centralized baseband unit (BBU) and distributed remote radio heads (RRHs). The RRHs usually perform digital signal processing, digital-to-analog conversion, power amplification, filtering, and so on. The BBU then provides computing power virtualized into a pool of resources shared by multiple RRHs for network control and management [1]. The BBU is interconnected with the RRHs via a fronthaul, which is typically represented by transport links with high capacity and low delay.

Besides the low energy consumption and cost reduction benefits, C-RAN also enables flexibility and large-scale coordination of the network control and management functionalities [2]. For example, the centralized control in the BBU facilitates interference mitigation techniques, such as inter-cell interference coordination (ICIC) and coordinated multipoint (CoMP), as the BBU has knowledge about the whole network or about a larger area under the BBU's coverage. However, shifting a part of the control functionalities from the RRHs to the BBU inevitably introduces new challenges due to high requirements on the fronthaul throughput and delay.

To relax the requirements on the fronthaul, the control functionalities can be split among the BBU and the RRHs as defined by the 3rd Generation Partnership Project (3GPP) [3, 4]. The 3GPP outlines eight options of functional split, each determining a subset of the control functions carried out locally in the distributed RRHs and the functions performed centrally in the BBU. While split option 1 is similar to a conventional base station with almost all communication control functionalities (except radio resource control) in the distributed RRHs, split option 8 keeps only a radio frequency unit in the RRHs, whereas all control functions are centralized [1]. The local distributed processing in the RRHs relaxes the requirements on the fronthaul and reduces the delay. On the contrary, the centralized processing facilitates the cost and energy consumption benefits of C-RAN.

Each option of functional split is preferable for different network conditions and requirements of services [5]. Thus, in [3], the RAN-as-a-service concept is proposed to allow adaptable centralized management and processing according to the actual service demands. Still, a flexible and fully dynamic change of the deployed functional split according to the user's requirements, network status, and current load of the fronthaul is a challenge [6]. With the control functions performed in a centralized way, any realistic fronthaul with non-zero delay negatively affects the overall network performance and the quality of service for users. The reason is that the channel quality information required for radio resource management is outdated. In [7], the authors propose a cloud-based radio-over-fiber network concept optimizing an allocation of communication and processing resources between the RRHs and the BBU to reduce congestion of the fronthaul and to balance load among them.

The authors provide an overview of existing scheduling approaches suitable for C-RAN and identify their potential limitations. Based on these limitations, they outline a framework for hierarchical scheduling. The hierarchical scheduling mitigates a negative impact of the fronthaul delay on the throughput of non-cell-edge users and enables efficient retransmission of erroneous data.

Zdenek Becvar, Pavel Mach, and Mohammed Elfiky are with Czech Technical University in Prague; Mitsuo Sakamoto is with Hon Hai Precision Ind. CO., LTD.

Digital Object Identifier: 10.1109/MCOM.001.2000697

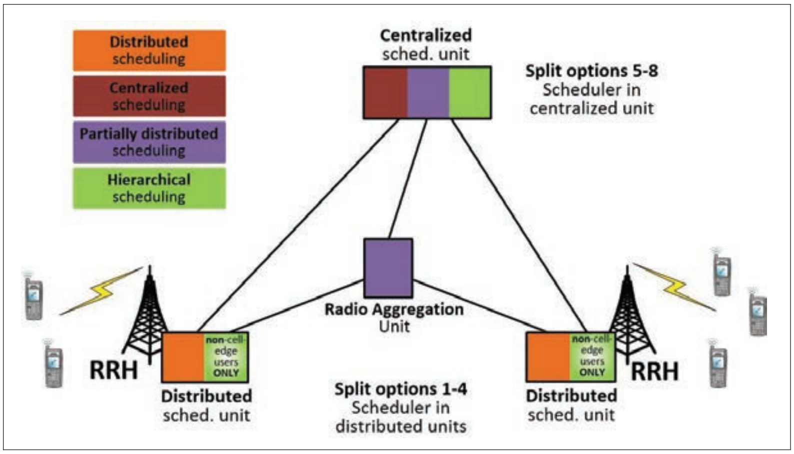

FIGURE 1. Overview of various scheduling options for C-RAN. Individual colors indicate nodes involved in the scheduling for each scheduling type. Hierarchical scheduling is active in the distributed units only for non-cell-edge users.

In this article, we focus on radio resource scheduling in mobile networks with C-RAN architecture. The objective of scheduling in the mobile networks is to assign particular radio resources to individual users [8]. Depending on the split of control functionalities, in C-RAN, the scheduling is performed either in a distributed way (split options 1–4 [4]) or in a centralized way (split options 5–8). Furthermore, a combination of both options toward a partially distributed solution, introduced in [9], enables emphasizing the benefits of both solutions.

We first overview key aspects of the conventional scheduling approaches in C-RAN and identify their limitations and challenges to increase the efficiency of the scheduling in C-RAN. Motivated by the challenges, we outline and discuss a hierarchical scheduling scheme. The hierarchical scheduling splits the scheduling-related functionalities between the centralized and distributed units to mitigate a negative impact of the fronthaul delay on the network throughput. The centralized unit handles especially the scheduling for cell-edge users, as these suffer from inter-cell interference. Then the distributed units manage and adjust the allocation of resources for non-cell-edge users. The simulation results indicate that the hierarchical scheduling is superior to other approaches and mitigates impairment of the throughput in the case of a realistic fronthaul with non-zero delay.

Scheduling in C-RAN-Based Mobile Networks

This section overviews key aspects of distributed, centralized, and partially distributed scheduling. Also, pros and cons of each are highlighted to identify the challenges for scheduling in C-RAN.

Distributed Scheduling

Distributed scheduling is analogous to traditional scheduling in 4G networks, where each base station schedules the radio resources for the users individually without any interaction with other base stations. In C-RAN, the distributed scheduling corresponds to the case when the scheduling is carried out in the RRHs (i.e., function split options 1–4), as indicated in Fig. 1 (orange). This solution benefits from fast response to any changes in the user's channel quality and fast retransmission of erroneous packets. However, scheduling performed individually by each RRH without any coordination with neighboring RRHs results in strong mutual interference, and the performance of cell-edge users is degraded. To mitigate the interference, the RRHs can coordinate their transmission to the cell-edge users, for example, by means of ICIC or CoMP. Then several neighboring RRHs interact with each other and perform scheduling considering the interference to the users served by other RRHs. However, this solution implies mutual signaling exchange among the coordinated neighboring RRHs and, consequently, a significant load of the direct connections among the RRHs and high complexity of network management. Thus, the coordination is not easy and significantly increases the cost of the network deployment.

Summary of *Distributed Scheduling:*
- *Pros:* low packet delay and easy management of the erroneous packets retransmission as the fronthaul does not impair scheduling
- *Cons:* complex implementation resulting in high cost (more features included in the RRHs); high amount of signaling for coordinated interference mitigation

Centralized Scheduling

Centralized scheduling is performed solely in the centralized unit, as shown in Fig. 1 (red). This case is represented by function split options 5–8, where the scheduling-related functionalities are located in the centralized BBU. On one hand, this solution can efficiently facilitate interference mitigation techniques, as the centralized unit contains information about the whole network or at least about a larger area covered by this unit. Centralized scheduling also preserves C-RAN benefits of lower cost and energy consumption. On the other hand, centralized scheduling requires delivery of the scheduling-related information (channel quality, buffer status, etc.) from the RRHs to the centralized unit. In this case, the fronthaul quality plays a substantial role. A fronthaul with high delay can heavily degrade the overall network performance and outweighs the gains introduced by the interference mitigation techniques, because the scheduling is performed with outdated channel state information. The fronthaul delay also impairs and prolongs the retransmissions of erroneous packets.

Summary of *Centralized Scheduling:*
- *Pros:* low cost; easy coordination of scheduling among RRHs to suppress interference
- *Cons:* performance potentially degraded due to the fronthaul delay; complicated management of erroneous packet retransmission, leading to high delay

Partially Distributed Scheduling

A type of scheduling combining both centralized and distributed approaches is introduced in [9], where the scheduling functions are split between the centralized BBU and partially distributed radio aggregation units (RAUs) (Fig. 1, purple). The RAU is a new semi-distributed entity concentrating the control functions for several underlying RRHs. Hence, multiple RAUs are deployed in the network, and each RAU performs the scheduling for the several underlying RRHs so that each is under the control of just one RAU. In parallel to the scheduling in the RAUs, the BBU performs its own scheduling for the whole network. The decision on whether the scheduling derived in the BBU or in the RAUs is exploited depends on the fronthaul delay. If the fronthaul delay prevents timely delivery of the scheduling from the BBU to the RAUs, the scheduling done by the RAUs is exploited. On the contrary, if the scheduling from the BBU arrives on time, the scheduling by the RAUs is overruled by that from the BBU. From the function split perspective, the scheduling is not done in the RRHs at all; thus, the partially distributed scheduling corresponds rather to the centralized scheduling and split options 5–8.

Summary of *Partially Distributed Scheduling:*
- *Pros:* reduced impact of the fronthaul delay compared to the centralized approach; possible coordination among RRHs to suppress interference
- *Cons:* higher cost than the centralized approach, as every scheduling is performed twice (in the BBU and the RAU) and the new entity is required; performance still limited by the actual status of the fronthaul and its parts (RRH to RAU and RAU to BBU)

Challenges in Scheduling for C-RAN

From the overview of the existing scheduling approaches for C-RAN, we identify the following gaps and challenges related to preservation of the native C-RAN benefits while mitigating its drawbacks:
- *Challenge 1:* Suppress the negative impact of the fronthaul delay on the network throughput via a combination of the centralized and distributed solutions.
- *Challenge 2:* Handle fast retransmission of erroneous packets while still enabling interference mitigation among neighboring RRHs.
- *Challenge 3:* Enable a dynamic split with support for dynamic reallocation of the scheduling-related functionalities among the network nodes, as targeted by the 3GPP in [4]. The dynamic split can bring significant cost savings, as the computing and processing capabilities of individual virtualized entities can be pooled and shared according to the actual load of each node.

The Hierarchical Scheduling Framework

This section provides an overview of hierarchical scheduling, outlined in [10], to address the negative impact of the fronthaul on the network throughput (*Challenge 1*). Then an efficient way of erroneous packet retransmission (*Challenge 2*) by hierarchical scheduling is presented. Furthermore, the flexibility of the radio resource scheduling according to the availability of the processing resources in the centralized and distributed units is discussed (*Challenge 3*).

Hierarchical Scheduling for C-RAN

To suppress the negative impact of the fronthaul (*Challenge 1*), hierarchical scheduling is composed of two tiers: a centralized scheduling unit running in any (semi-) centralized entity, and a distributed scheduling unit located in any distributed entity (Fig. 2). We illustrate the concept for the case with the centralized unit and the distributed units located in the BBU and the RRHs, respectively.

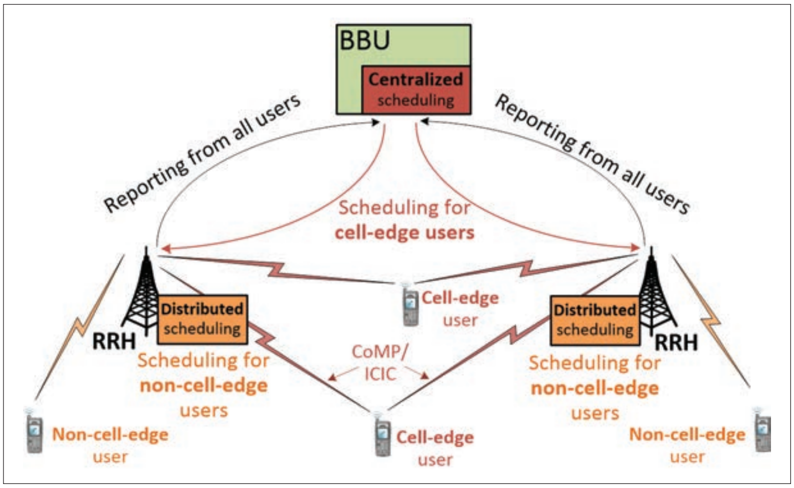

FIGURE 2. Overview of hierarchical scheduling: the centralized scheduling unit performs high-level scheduling for all users exploiting global interference knowledge; distributed scheduling units adjust scheduling for non-cell-edge users and handle retransmission of erroneous packets.

The centralized unit provides the scheduling with an awareness of the mutual interference among the cells. The centralized unit primarily plans the data transmission for the cell-edge users, because these users experience strong interference from the neighboring cells (in downlink) and the users in the neighboring cells (in uplink). Thus, the centralized unit efficiently handles the resource allocation for any inter-cell interference mitigation technique.

The scheduling decision done by the centralized unit is sent via the fronthaul to the distributed unit. The distributed unit then adjusts the scheduling for the non-cell-edge users if the reported channel quality by the users becomes outdated due to the fronthaul delay. The distributed unit is allowed to update the scheduling only for the non-cell-edge users, which are not significantly influenced by the inter-cell interference. In contrast, the distributed unit cannot modify the scheduling for the cell-edge users to ensure efficient inter-cell interference mitigation.

The classification of the users to cell-edge and non-cell-edge can be done based on the channel quality, interference level, total number of occupied resources, and so on. As the hierarchical scheduling supports the cooperative interference mitigation techniques, the classification based on a benefit of the RRHs' cooperation for the interference mitigation is suggested. Hence, the user is classified as cell-edge if the cooperation of at least two RRHs on the transmission to this user reduces the number of resources required to serve this user. Otherwise, the user is labeled as non-cell-edge.

As the centralized scheduling is performed in a (semi-) centralized unit, its scheduling decision is affected by the fronthaul delay. Thus, delivering a new scheduling decision in each transmission time interval (TTI) can be redundant as the scheduling decision from the centralized unit is inaccurate anyway if the actual radio channel quality changes due to the fronthaul delay. Thus, the centralized unit creates long-term scheduling that is understood as the scheduling for *N* consec-

The centralized unit primarily plans the data transmission for the cell-edge users, because these users experience a strong interference from the neighboring cells (in downlink) and the users in the neighboring cells (in uplink). Thus, the centralized unit efficiently handles the resource allocation for any inter-cell interference mitigation technique.

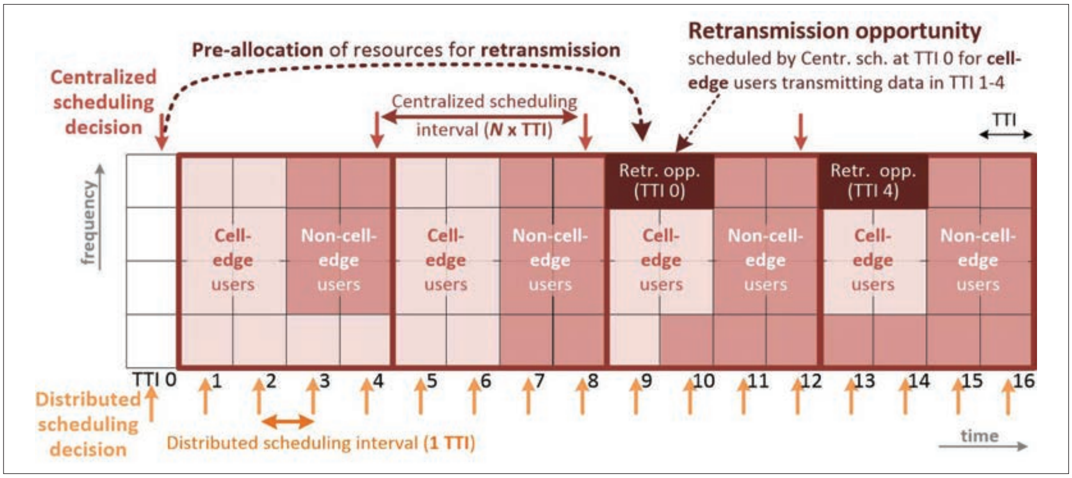

FIGURE 3. Example of scheduling intervals of centralized and distributed units with pre-allocated resources for retransmission.

utive TTIs (i.e., *N*xTTI). The long-term scheduling reduces requirements on the processing power of the centralized unit and lowers the amount of signaling overhead exchanged over the fronthaul. In contrast, the scheduling in the distributed unit is not affected by the fronthaul delay at all. Thus, the distributed unit is allowed to adjust the centralized scheduling for the non-cell-edge users on a short-term basis, for example, every TTI (Fig. 3).

The long-term scheduling over a high number of TTIs (i.e., high *N*) leads potentially to throughput degradation as the scheduling does not reflect the actual radio conditions (i.e., channel information used for the scheduling may not be valid anymore in later TTIs [11]). To cope with the potential throughput degradation for the cell-edge users due to high *N*, we propose to schedule the transmissions to the cell-edge users in the early TTIs located just after the centralized scheduling decision is done, as shown in Fig. 3. In an extreme case, all TTIs at the beginning of *N* can be dedicated solely to the cell-edge users. Then, progressively, fewer resources are allocated to the cell-edge users in the following TTIs. Since the distributed units are able to adapt the long-term scheduling for the non-cell-edge users in each TTI according to the actual channel information, the later scheduling of the non-cell-edge users does not degrade their throughput.

The hierarchical scheduling enables to adapt *N* for individual cell-edge users according to their channel statistics. More specifically, the actual value of *N* is individually and dynamically set for each user according to the magnitude and frequency of the users' channel fluctuation over time. Thus, the negative effect of the outdated channel quality information is suppressed via a low *N* if the channel fluctuates significantly and, in contrast, the amount of signaling overhead is reduced by setting a larger *N* if the channel is stable.

Retransmission of Erroneous Packets

The hierarchical scheduling enables a fast and efficient handling of the erroneous packets retransmission to address *Challenge 2*. The fast error correction is provided by the hybrid automatic repeat request procedure that combines retransmission of the erroneous packets with forward error correction [12]. For the non-cell-edge users, the retransmission process follows a common hybrid automatic repeat request procedure in the mobile networks, that is, the resources for the retransmission are scheduled by the distributed unit if the packet is received with errors by the user (indicated by a request for retransmission). For the retransmissions of erroneous data, the distributed unit selects the most suitable resources that are not dedicated to the cell-edge users at the moment.

The retransmission of erroneous data is more complicated for the cell-edge users. Handling the retransmissions by the centralized unit would lead to a high packet delay as the requests for retransmission should be delivered to the centralized unit, and then a new scheduling decision for the retransmission should be sent back to the distributed unit. To avoid this long-lasting process, the centralized unit pre-allocates specific radio resources for potential retransmissions of the erroneous packets for the cell-edge users during each scheduling period *N* (Fig. 3, with two retransmission opportunities for the scheduling done in TTI 0 and TTI 4). The retransmission for the cell-edge users is then handled solely by the distributed units at the pre-allocated radio resources with no intervention from the centralized unit. Thus, the retransmission process is shortened, and the fronthaul delay does not affect the retransmissions at all.

To avoid wasting the radio resources pre-allocated for retransmission, the distributed unit assigns unused pre-allocated resources to any non-cell-edge user(s) in an arbitrary way, since the interference from other neighboring cells is low.

Dynamic Deployment of Scheduling-Related Functionalities

With the evolution of software defined networking and edge computing, some base stations (including RRHs) are supposed to be accompanied with a certain amount, albeit only small, of computing power. Hierarchical scheduling enables exploiting the virtualized resources already available in these base stations and RRHs in an efficient way. Hierarchical scheduling provides the ability to pool the resources for scheduling among individual units and exploit them opportunistically, where and when these are required. Thus, hierarchical scheduling allows operating even in a fully centralized way when only very simple and low-cost RRHs are deployed.

To address *Challenge 3*, the computing load of both centralized and distributed units related to scheduling is balanced by either an adjustment of the ratio of the cell-edge and non-cell-edge users or an appropriate setting of the centralized scheduling period N. Consequently, we can dynamically control and balance the scheduling-related computing load of the distributed units and improve the quality of service offered to users depending on the current network status including fronthaul quality, radio channel quality, and so on.

The dynamic split that changes the allocation of the scheduling functions over time is implemented by continuous and periodic re-classification of the users to cell-edge and non-cell-edge. This re-classification process considers the network load, which influences the fronthaul delay. With a high fronthaul delay, more users are scheduled directly by the distributed unit to overcome the low-quality fronthaul, and the ratio of non-cell-edge users is increased. In contrast, with a low fronthaul delay, more users are scheduled by the centralized unit to lower the ratio of the non-cell-edge users.

Performance of Scheduling in C-RAN

In this section, we provide an overview of a scenario and models for the performance evaluations. Then we discuss and compare the performance of individual C-RAN scheduling options.

Simulation Scenario and Models

We consider a square area with a size of 1 × 1 km encompassing a single BBU with the centralized scheduling unit located in the middle of the area. Furthermore, 100 RRHs with the distributed scheduling unit and 400 currently active users are deployed randomly with the uniform distribution. Each RRH is connected to the BBU via the fronthaul with a constant delay from 0 to 30 ms. The delay is the same for all RRHs, and we investigate the impact of the delay in the next subsection. Each user is associated with the RRH providing the highest received signal strength. The RRHs transmit with 27 dBm at 2 GHz carrier frequency. The channel between the user and the RRH follows the Urban Microcell model [13] with Rayleigh (standard deviation of 4 dB) and Rician (standard deviation of 7.82 dB) fading to model shadowing and fast fading, respectively, according to [14]. We assume an orthogonal frequency-division multiple access (OFDMA) system with a bandwidth of 20 MHz modeled in line with LTE-A-Pro standardized in 3GPP (TS 36.300).

The modulation and coding scheme for the transmission is determined according to a reported channel quality indicator (CQI) in line with [13]. The users are classified to cell-edge and non-cell-edge according to ICIC gain as indicated above. The retransmission of erroneous packets is implemented in line with the hybrid automatic repeat request procedure defined by 3GPP in [13].

The following C-RAN scheduling options are compared:

- *Distributed scheduling* solely performed in the RRHs (split options 1–4 according to 3GPP [4])
- *Centralized scheduling* deployed in the BBU (split options 5–8)
- *Partially distributed scheduling* proposed in [9]
- *Hierarchical scheduling* introduced in [10] and described above

All compared schedulings exploit the same implementation of ICIC based on [15] for a fair comparison. Moreover, all solutions consider conventional proportional fair scheduling, where the sum of logarithmic average of the users' throughput is maximized. Since [9] does not specify any deployment of the RAUs, we consider a realistic case, where each RAU is collocated with the RRH closest to the center of the cluster of all underlying RRHs.

As hierarchical scheduling supports adaptation of the centralized scheduling period N, we investigate the impact of N on the throughput, and we demonstrate an upper bound performance of hierarchical scheduling (denoted as *Hierarchical*, *optimum*) reached if N is dynamically set to an optimum value N_{opt}. The exact value of N_{opt} is determined individually for each cell-edge user according to its "channel dynamicity." The channel dynamicity is understood as a significance of the changes in CQI per a monitored period of time (in our case 50 ms). The more significantly the CQI changes within the monitored period, the lower N is set for the cell-edge user, because a larger N would lead to outdated information for scheduling. To indicate the theoretical upper bound, N_{opt} is selected so that a perfect prediction of the channel for the monitored period is assumed, and the impact of each possible scheduling period from 1 to N on the throughput for each cell-edge user is evaluated. Then the value of N that yields the maximum throughput is selected as N_{opt}.

Comparison of C-RAN Scheduling Options

The impact of the fronthaul delay on the sum throughput of all users is investigated in Fig. 4. Note that for distributed scheduling, the fronthaul delay represents the delay at the direct connection among the RRHs (e.g., X2 interface). The throughput decreases with increasing fronthaul delay for all investigated schedulings, as the fronthaul delay leads to a delay in the delivery of the channel quality reports to the centralized unit and of the scheduling decision to the distributed unit. The fronthaul delay impairs the throughput of the distributed, centralized, and partially distributed solutions heavily, while hierarchical scheduling notably suppresses the negative impact of the fronthaul delay. The hierarchical scheduling outperforms the distributed, centralized, and partially distributed solutions by 16, 22, and 26 percent, respectively. The gain of hierarchical scheduling is achieved as the scheduling in the centralized unit is adjusted in the distributed units for the non-cell-edge users if the fronthaul delay leads to a notable change in the channel quality. The rapid drop in the throughput of the partially distributed scheduling at certain fronthaul delays (Fig. 4) corresponds to twice (uplink plus downlink) the threshold determining whether the scheduling is distributed or centralized.

The centralized solution reaches the lowest throughput; however, the cost of computing/processing resources required to perform the scheduling is the lowest out of all solutions due to a

With a high fronthaul delay, more users are scheduled directly by the distributed unit to overcome the low quality fronthaul and the ratio of non-cell-edge users is increased. In contrast, with a low fronthaul delay, more users are scheduled by the centralized unit to lower the ratio of the non-cell-edge users.

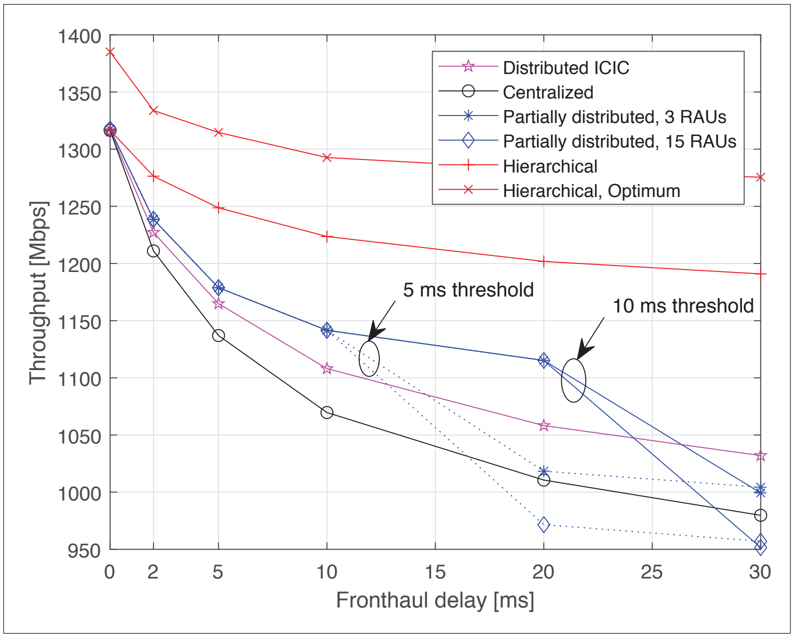

FIGURE 4. Impact of fronthaul delay on network throughput for individual C-RAN scheduling options, $N = 1$.

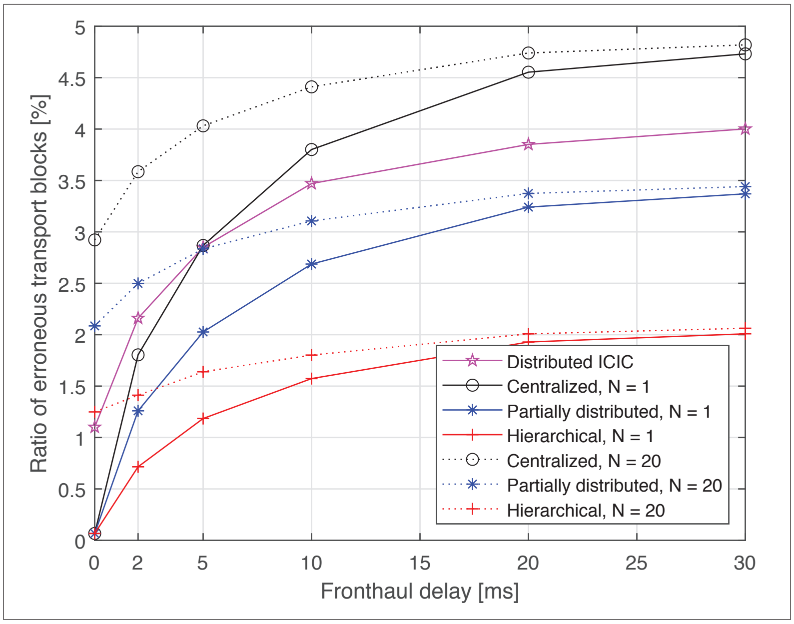

FIGURE 5. Impact of fronthaul delay on ratio of transport blocks received with error and retransmitted.

concentration of the scheduling for all RRHs in the BBU. Distributed scheduling requires full processing power for the scheduling in every RRH; thus, this solution induces the highest cost. A gain in the throughput of the partially distributed solution with respect to the centralized one is at an increased cost introduced by the new entities, the RAUs. However, due to an aggregation of the RAUs for several RRHs, the cost of the partially distributed solution is typically still lower than that of the distributed kind. The cost of hierarchical scheduling is also higher than the cost of the centralized kind, as (at least some) RRHs should be equipped with an additional computing power for the scheduling-related processing. This cost is similar to the partially distributed scheduling, but lower than the cost of distributed scheduling, as only the cell-edge users are scheduled by the RRHs in the hierarchical scheduling.

Figure 5 shows the ratio of the transport blocks requiring retransmission. For all types of scheduling, the transport block error rate increases with the fronthaul delay and saturates at approximately 20 ms delay. The saturation takes place because the channel variation for this fronthaul delay (~20 ms) is so large that the impact of the fast fading is random. Also, the higher N results in a higher error rate as the larger N increases the probability of the channel information used for scheduling being outdated. Hierarchical scheduling reduces the error rate by up to 50, 58, and 41 percent compared to distributed, centralized, and partially distributed scheduling, respectively.

Furthermore, we investigate the impact of the prolonged scheduling period of centralized unit N on the throughput. The longer the scheduling period is, the less computation power and less signaling are needed for scheduling. Figure 6 shows that the throughput first increases with N , and then it starts decreasing at a certain point. The initial increase for small values of N is a result of the reduced scheduling-related overhead, because the scheduling information is sent only every Nth TTI, and more resources remain for users' data. With respect to the conventional setting of $N = 1$, hierarchical scheduling with prolonged N introduces a gain of 5 percent for the ideal fronthaul with no delay and for $N = 2$ (ideal fronthaul illustrated with solid lines in Fig. 6). With further prolonging of N, the throughput starts decreasing, as the scheduling information is not up to date for the later TTIs within N. The hierarchical scheduling benefits from an adjustment of the scheduling by the distributed unit for the non-cell-edge users. Thus, the overhead reduction is preserved, and the throughput degradation is suppressed.

If hierarchical scheduling is accompanied by the optimum adaptation of N according to the dynamicity of the channels (line *Hierarchical, optimum*), the throughput is further improved. The reason is that the value of N_{opt} is selected from the range $1 \leq N_{opt} \leq N$ in such a way that the actual value of N for each user is dynamically set according to the channel fluctuation over time. The maximum throughput achieved for hierarchical scheduling with the optimum scheduling period is roughly 10 percent above the maximum achieved by all other approaches for the fronthaul with zero delay (solid lines, Fig. 6). If the fronthaul delay increases to 10 ms (dotted lines, Fig. 6), the benefit resulting from the prolonged scheduling period of the centralized unit is less significant. This is because N becomes overwhelmed with the fronthaul delay, and anyway, the scheduling is done too far in advance with respect to the time of data transmission (fronthaul delay plus N). However, Fig. 6 confirms that hierarchical scheduling suppresses the negative impact of the fronthaul delay and outperforms distributed, centralized, and partially distributed scheduling by 11, 17, and 8 percent, respectively, for 10 ms fronthaul delay and $N = 20$.

Conclusion

In this article, we have provided an overview of the scheduling approaches for C-RAN. Motivated by limitations of the existing solutions, we have proposed a new framework for hierarchical scheduling in the networks with C-RAN architecture. The hierarchical scheduling encompasses the distributed units and the centralized unit. While the centralized unit performs the long-term scheduling, especially for the cell-edge users that can benefit from interference mitigation techniques, the distributed units eliminate a negative impact of fronthaul delay on non-cell-edge users and enables efficient handling of error correction. We show that hierarchical scheduling notably improves the network throughput (by up to 26 percent) via suppression of the negative impact of the fronthaul delay. This gain can be further increased by several percentages via dynamic setting of the scheduling period of the centralized unit.

Acknowledgments

The authors would like to thank Dr. Morten Høgdal, Dr. Andrijana Popovska Avramova, and Dr. Aleksandra Checko from MTI Radiocomp ApS, Denmark, for their valuable discussions and suggestions. The work was supported by Hon Hai Precision Industry Co., Ltd., trading as Foxconn Technology Group, Taiwan, and by projects of the CTU in Prague nos. SGS17/184/OHK3/3T/13 and SGS20/169/OHK3/3T/13.

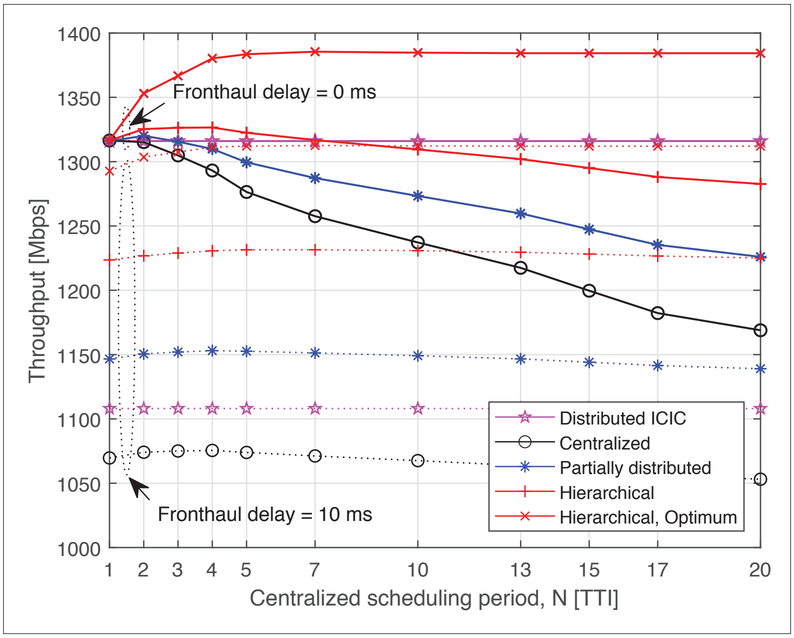

FIGURE 6. Impact of scheduling period of centralized unit, *N*, on throughput for fronthaul delay of 0 ms (solid lines) and 10 ms (dotted lines).

References

[1] A. Checko *et al.*,"Cloud RAN for Mobile Networks — A Technology Overview," *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 1, 1st qtr. 2015.

[2] C. Lin I *et al.*, "Recent Progress on C-RAN Centralization and Cloudification," *IEEE Access*, vol. 2, 2014.

[3] P. Rost *et al.*, "Cloud Technologies for Flexible 5G Radio Access Networks," *IEEE Commun. Mag.*, vol. 52, no. 5, May 2014.

[4] 3GPP TR38.801, "Study on New Radio Access Technology: Radio Access Architecture and Interfaces," v.14.0.0, Mar. 2017.

[5] L. M. P. Larsen, A. Checko, and H. L. Christiansen, "A Survey of the Functional Splits Proposed for 5G Mobile Crosshaul Networks," *IEEE Commun. Surveys & Tutorials*, vol. 21, no. 1, 1st qtr. 2019.

[6] M. Peng *et al.*, "Fronthaul-Constrained Cloud Radio Access Networks: Insights and Challenges," *IEEE Wireless Commun.*, vol. 22, no. 2, Apr. 2015.

[7] H. Yang *et al.*, "C-RoFN: Multi-Stratum Resources Optimization for Cloud-Based Radio over Optical Fiber Networks," *IEEE Commun. Mag.*, vol. 54, no. 8, Aug, 2016.

[8] F. Capozzi *et al.*, "Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and a Survey," *IEEE Commun. Surveys & Tutorials*, vol. 15, no. 2, 2bd qtr. 2013.

[9] M. Huang and X. Zhang, "Distributed MAC Scheduling Scheme for C-RAN with Non-Ideal Fronthaul in 5G Networks," *Proc. IEEE WCNC 2017.*

[10] Z. Becvar *et al.*, "Hierarchical Resource Scheduling Method of Wireless Communication System," U.S. patent no. US10624105B2, 2020.

[11] N. Pontois *et al.*, "User Pre-Scheduling and Beamforming With Outdated CSI in 5G Fog Radio Access Networks," *Proc. IEEE GLOBECOM*, 2018.

[12] 3GPP TS36.213, "Evolved Universal Terrestrial Radio Access (EUTRA); Physical Layer Procedures," v.15.6.0, 2019.

[13] 3GPP TR38.901, "Study on Channel Model for Frequencies from 0.5 to 100 GHz," v.15.0.0, June 2018.

[14] W. Sun *et al.*, "Transmission Mechanism and Performance Analysis of Multiuser Opportunistic Beamforming in Rayleigh and Rician Fading Channels," *IEEE Trans. Vehic. Tech.*, vol. 67, no. 10, Oct. 2018.

[15] G. Nardini *et al.*, "Practical Large-Scale Coordinated Scheduling in LTE-Advanced Networks," *Wireless Network*, vol. 22, 2016.

Biographies

Zdenek Becvar received his M.Sc. and Ph.D. in telecommunication engineering from Czech Technical University in Prague (CTU), Czech Republic, in 2005 and 2010, respectively. He is an associate professor in the Department of Telecommunication Engineering at CTU. From 2006 to 2007, he joined Sitronics R&D Centre, and in 2009, he was involved in research activities of Vodafone R& D at CTU in 2009. From 2013 to 2017, he was a representative of CTU in 3GPP/ETSI.

Pavel Mach received his M.Sc. and Ph.D. degrees from CTU in 2006 and 2010, respectively. Now, he is a senior researcher at the same university in the 5Gmobile research lab focusing on key challenges related to future mobile networks.

Mohammed Elfiky received his M.Sc. degree in telecommunication engineering from the University of Malaya, Malaysia, in 2014. He is working toward a Ph.D. in the Department of Telecommunication Engineering, CTU.

Mitsuo Sakamoto received his B.E. in electrical communication engineering from Toyota Technological Institute in 1989 and his Ph.D. from Kyusyu University in 2006. He has over 30 years of experience overseeing wireless telecommunications systems for Nokia, Orange, eMobile, Fujitsu, and so on. He has experience with standardization in 3GPP/3GPP2, ITU-R/WRC/APT, ARIB/5GMF, and MIC. Since 2016, he has been a director of Foxconn Advanced Communication Academy.