Optimal Split Bearer Control and Resource Allocation for Multi-Connectivity in 5G New Radio

Published Version:
Optimal Split Bearer Control and Resource Allocation for Multi-Connectivity in 5G New Radio / Elias, Jocelyne; Martignon, Fabio; Paris, Stefano. - ELETTRONICO. - (2021), pp. 9482505.187-9482505.192. (Intervento presentato al convegno Joint 30th European Conference on Networks and Communications and 3rd 6G Summit, EuCNC/6G Summit 2021 tenutosi a Porto, Portugal nel 8-11 June 2021) [10.1109/EuCNC/6GSummit51104.2021.9482505].

Availability:
This version is available at: https://hdl.handle.net/11585/832133 since: 2021-09-13

Published:
DOI: http://doi.org/10.1109/EuCNC/6GSummit51104.2021.9482505

(Article begins on next page)

J. Elias, F. Martignon and S. Paris, "Optimal Split Bearer Control and Resource Allocation for Multi-Connectivity in 5G New Radio," 2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit), Porto, Portugal, 2021, pp. 187-192

The final published version is available online at ***https://dx.doi.org/10.1109/EuCNC/6GSummit51104.2021.9482505***

Optimal Split Bearer Control and Resource Allocation for Multi-Connectivity in 5G New Radio

Jocelyne Elias[1], Fabio Martignon[2], and Stefano Paris[3]

[1]University of Bologna, jocelyne.elias@unibo.it
[2]University of Bergamo, fabio.martignon@unibg.it
[3]Nokia Bell Labs, stefano.paris@nokia-bell-labs.com

***Abstract*—Wireless cellular systems are evolving towards the integration of services and devices with a diverse set of throughput, latency, and reliability requirements. To this end, 3GPP has introduced Multi Connectivity (MC) as a more flexible architecture for 5G NR (New Radio), where multiple wireless connections can be used simultaneously to split or duplicate data traffic. While multi connectivity improves single user performance, the inherent increase of data transmissions on the wireless channel may result in higher interference, thus reducing the overall system performance.**

This paper analyzes the problem of admission control and resource allocation in multi connectivity scenarios, considering different requirements and 5G NR features. We formulate the problem as an optimization program and provide an heuristic approach to solve it. Numerical results in a realistic network deployment confirm that our solution can effectively allocate radio resources in order to increase admission rate and system throughput.

***Index Terms*—5G, Multi-Connectivity, Split-Bearer, Optimization.**

I. Introduction

The proliferation of new services and devices as well as the paradigm shift of sectors like industry, healthcare, and transportation towards interconnected systems are calling for the design of more complex architectures and resource management schemes for next generation mobile networks. Indeed, wireless cellular systems are foreseen as the essential technology to connect all these new devices and enable the integration of these new services.

To satisfy the diversity of throughput, latency, and reliability requirements, wireless protocols and standards have been rapidly evolving to make use of higher frequencies and larger spectrum, flexible waveforms and access schemes, more complex schemes for sharing and scheduling radio resources, as well as more flexible architectures to connect and host the elements of the mobile network. For example, 5G New Radio (NR) has introduced Multi-Connectivity (MC) [1] as a simple and effective way for improving latency, reliability, and throughput of cellular communications. In MC a User Equipment (UE) is connected to two next-generation NodeBs (gNBs), each handling up to two cells configured on different carrier frequencies. Therefore, up to four cells can be used for data transmission. The hosting node, typically called Master gNB (MgNB), hosts the SDAP layer (Service Data Adaptation Protocol) that receives packets of the data radio bearer from the Core Network (CN) and passes them to the hosting PDCP layer (Packet Data Convergence Protocol). The hosting node maintains the main control and signalling with the UE, and it can activate an assisting node, also called Secondary gNB (SgNB), to setup auxiliary connections with the UE. Depending on the QoS requirements of the UE, the assisting node can be used to duplicate data packets or transit only part of the data traffic in order to increase reliability or throughput, respectively. Specifically, the PDCP layer controls the flow of data packets of the data radio bearer by deciding which packets must be transmitted through MgNB and SgNB.

Deciding which radio connections to activate and how to split the flow of packets to serve a data radio bearer is a key problem in the design of next-generation mobile networks. Indeed, the use of multiple connections and the ability to split the radio bearer across multiple cells enable the use of spare resources to reduce system blockage. In this paper we formulate and solve the PDCP Split-Bearer Decision problem. The problem consists in (1) deciding which users to serve and (2) deciding whether and how to split the traffic of admitted users across multiple cells (also referred to as *legs*) to meet the bandwidth requirements of the services. All key features of 5G NR are captured and modeled. We propose effective, exact approaches that achieve the optimal solution as well as a decomposition (two-step) approach that, coupled with a carefully chosen bound to the set of legs that each user can exploit, permits to consistently reduce the computing time while still achieving close-to-optimum solutions, even in real-size network scenarios.

Numerical results, obtained in realistic cellular deployments and traffic scenarios, as defined by standards [2], show the effectiveness of the proposed models and approaches, and permit to capture and quantify the trade-off that mobile operators must face between admitting more users and providing them the necessary resources to fulfill their stringent requirements, specifically in terms of allocated rate on the radio interface and for all different transmission legs.

This paper is structured as follows: Section II discusses related work, while Section III presents the system model we consider for the formulation of the problem. Section IV illustrates the proposed mathematical formulation of the admission control and resource allocation problem for multi-connectivity scenarios. Section V presents the decomposition and heuristic approaches we propose to diminish consistently the time necessary to compute a close-to-the-optimum solution. Numerical results obtained in realistic cellular network scenarios, with a large number of gNBs and mobile devices, are presented and discussed in Section VI. Finally, conclusions and directions for future research are illustrated in Section VII.

II. Related Work

Multi-Connectivity has been proposed as a simple and effective way to satisfy data rates, latency and reliability requirements of next generation mobile networks. Different options for connecting to multiple radio access points, including PDCP solutions, are discussed in [3]. Since multi-connectivity in 5G ultra-dense scenarios makes increasingly difficult to perform efficient resource allocation, machine learning techniques have been proposed to improve network performance [4], [5].

PDCP data duplication has been introduced into 5G NR on top of multi connectivity as a simple and effective method to improve the latency and reliability by exploiting the spatial, temporal and frequency diversity offered by multiple cells configured on different carrier frequencies [6], [7]. An analytical evaluation of the outage probability and resource utilization of multi-connectivity with PDCP data duplication is presented in [8]. Authors in [9] formulate PDCP data duplication as a mathematical optimization problem with latency and reliability constraints, and propose a heuristic approach to solve the problem. Another heuristic scheme is used in [10] to dynamically select data duplication only for users whose latency requirements are critical.

In contrast to data duplication, split-bearer has been proposed as a method to increase the throughput of data communications in 5G NR. The work in [11] presents a control mechanism to dynamically select the subset of connections for data transmission based on channel state information and cell load, while [12] proposes a utility based method to allocate resources in multi-connectivity among multiple user devices. Both schemes assume that all UEs can be fully satisfied and admitted in the system. Therefore, they do not consider the admission problem.

Differently from previous work, we formulate the joint admission control and resource allocation problem for multi-connectivity scenarios and propose a heuristic approach to speed up the computation of the optimal solution.

III. System model

In this section, we describe the system model that we consider for the formulation of the PDCP Split-Bearer Decision Problem. We model the Radio Access Network (RAN) as a bipartite graph $G = (\mathcal{U}, \mathcal{L}, \mathcal{L}_u)$, where $\mathcal{U}$ and $\mathcal{L}$ represent the set of users (UEs) and the set of all cells, respectively. $\mathcal{L}_u \subseteq \mathcal{L}$ models the set of cells that can be used by UE $u \in \mathcal{U}$ to transmit a data packet. Since the connection between a UE and a cell uniquely identifies a transmission leg, each element $l \in \mathcal{L}_u$ is called indifferently *leg* or *cell* in the rest of the paper. Among all legs in $\mathcal{L}_u$ that can be used to transmit a data packet, the wireless link between the MgNB and the UE u is denoted as primary link. We identify this primary leg using the function $l_0(u) : \mathcal{U} \to \mathcal{L}$.

G_{ul} models the channel gain between a cell $l \in \mathcal{L}$ and user $u \in \mathcal{U}$, while P_{ul} represents the power used by a cell to serve the UE. For a downlink transmission the cell selects the highest Modulation and Coding Scheme (MCS) according to the Signal to Interference plus Noise Ratio (SINR) in order to match the target error probability. We model the list of available MCSs as the set $\mathcal{M}$. For each MCS $m \in \mathcal{M}$, the threshold $\gamma_m \in \mathbb{R}$ represents the minimum SINR value to reliably transmit a data packet, while R_m identifies the transmission rate corresponding to MCS m.

A simplified RAN example with the main parameters of our system model is illustrated in Figure 1. The RAN is composed of two UEs and three cells (one per gNB). The cell l_1 provided by gNB$_1$ is the primary leg for UE u, whereas cell l_3 of gNB$_3$ is the primary leg for UE i. Both UEs share the cell provided by gNB$_2$ as secondary leg. Since cells of gNB$_1$ and gNB$_3$ are configured on the same carrier frequency, they may interfere with each other. Therefore, transmissions on legs $(u; l_1)$ and $(i; l_3)$ must be coordinated to avoid cross interference. All parameters and decision variables used in our mathematical formulation are summarized in Table I. The parameters represent the state of the network at decision time.

IV. PDCP Split Bearer Decision Problem

In this section we formulate the *PDCP Split-Bearer Decision (PSD)* problem, which consists in deciding whether and how to split the traffic across multiple legs to meet the bandwidth requirements of user services. To this end, we define the following decision variables. Binary variables x_{ulm} indicate whether the MCS m of the leg l is used to serve user u (i.e., $x_{ulm} = 1$). In our formulation the first leg represents the primary leg, namely the best leg connecting the UE u to its MgNB. This leg is identified by the funciton $l_0(u)$. Binary variables y_{ul} are auxiliary variables representing whether user

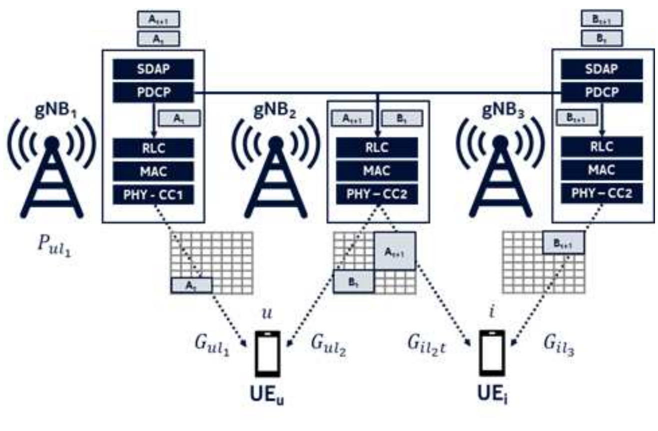

Fig. 1: Example of RAN with two UEs and three cells. Each UE enables two legs (i.e., it is connected to two cells). The node gNB$_2$ is the assisting node of both gNB$_1$ and gNB$_3$.

Parameter	Description
$\mathcal{U}$	Set of users (UEs)
$\mathcal{L}$	Set of legs
$\mathcal{M}$	Set of Modulation and Coding Schemes
$\mathcal{U}_l$	UEs that can be served by leg $l \in \mathcal{L}$
$\mathcal{L}_u$	Set of legs that can be used to serve traffic of UE $u \in \mathcal{U}$
$\mathcal{L}_l$	Legs that interfere with leg $l \in \mathcal{L}$
$l_0(u) : \mathcal{U} \to \mathcal{L}$	Function to identify UE's primary leg
P_{ul}	Power used by leg $l \in \mathcal{L}$ to serve UE $u \in \mathcal{U}$ [W]
G_{ul}	Channel gain for UE $u \in \mathcal{U}$ on leg $l \in \mathcal{L}$
$\mathcal{N}_0$	Noise
γ_m	Minimum SINR threshold for using MCS $m \in \mathcal{M}$
R_m	Rate for MCS $m \in \mathcal{M}$ [bit/s]
R_{max}	Maximum transmission rate [bit/s]
D_u	Traffic data rate of UE u [bit/s]
Variable	**Description**
x_{ulm}	Indicates whether the MCS m of the leg l is used to serve user u
y_{ul}	Auxiliary variable indicating whether leg l is used to serve user u
z_u	Indicates whether the traffic of UE u is fully served

TABLE I: Parameters and variables used in our system model and mathematical formulation.

leg l is used to serve user u. They are used to simplify the formulation of SINR constraints and the description of the optimization problem, but they can be omitted when solving the problem to reduce the memory needed to store decision variables. Finally, binary variables z_u indicate whether the traffic of UE u is fully served by the network (i.e., $z_u = 1$). The PSD problem can be formulated as the optimization problem (1)-(8):

$$\max \sum_{u \in \mathcal{U}} z_u - \alpha \sum_{u \in \mathcal{U}, l \in \mathcal{L}_u} y_{ul} \qquad \text{s. t.:} \tag{1}$$

$$\sum_{l \in \mathcal{L}_u} y_{ul} \geq z_u \qquad \forall u \in \mathcal{U} \tag{2}$$

$$\sum_{m \in \mathcal{M}} x_{ul_0(u)m} \leq y_{ul_0(u)} \qquad \forall u \in \mathcal{U} \tag{3}$$

$$\sum_{m \in \mathcal{M}} x_{ulm} \leq y_{ul} \qquad \forall u \in \mathcal{U}, l \in \mathcal{L}_u \setminus \{l_0(u)\} \tag{4}$$

$$y_{ul_0(u)} \geq y_{ul} \qquad \forall u \in \mathcal{U}, l \in \mathcal{L}_u \setminus \{l_0(u)\} \tag{5}$$

$$\frac{P_{ul} G_{ul}}{\mathcal{N}_0 + \sum_{j \in \mathcal{L}_l, k \in \mathcal{U}_j : k \neq u} P_{kj} G_{uj} y_{kj}} \geq \gamma_m x_{ulm}$$
$$\forall u \in \mathcal{U}, l \in \mathcal{L}_u, m \in \mathcal{M} \tag{6}$$

$$\sum_{m \in \mathcal{M}} R_m x_{ulm} \geq D_u z_u \qquad \forall u \in \mathcal{U} \tag{7}$$

$$x_{ulm}, y_{ul}, z_u \in \{0, 1\} \qquad \forall u \in \mathcal{U}, l \in \mathcal{L}_u, m \in \mathcal{M}. \tag{8}$$

The optimization function (1) maximizes the number of served UEs, and pursues also the minimization of the used resources, α being the weight that permits to trade-off between the two objectives. Constraints (2) state that at least one leg must be activated to user u when such user is served. Constraints (3) and (4) force the use of at most one MCS for the primary leg and the secondary leg, respectively. Note that if UE u is not served, no MCS is selected and all corresponding variables are set to zero. Constraints (5) force the activation of the primary leg if any secondary leg is used to transmit part of the traffic of UE u. Indeed, if the primary leg is not used, the traffic cannot be split on any other secondary leg. Constraints (6) model the SINR perceived by any UE u for leg l. A constraint is defined for each MCS according to its SINR threshold. We observe that all MCSs that satisfy their SINR threshold can be selected to serve a user u. However, constraints (3) and (4) limit the choice to a single value. Constraints (7) guarantee that the aggregated rate of all activated legs is enough to satisfy the traffic demand D_u of the user u. Finally, the set of constraints (8) defines the range of the decision variables. We underline that the objective function and the constraints can be modified to maximize other performance indicators, like reliability or end-to-end delay.

We observe that constraints (6) are not linear. However, they can be linearized by appropriately defining a large constant M as follows:

$$P_{ul} G_{ul} x_{ulm} + (1 - x_{ulm}) M \geq \tag{9}$$
$$\geq \left(\mathcal{N}_0 + \sum_{j \in \mathcal{L}_l, k \in \mathcal{U}_j : k \neq u} P_{kj} G_{uj} y_{kj} \right) \gamma_m$$
$$\forall u \in \mathcal{U}, l \in \mathcal{L}_u, m \in \mathcal{M}. \tag{10}$$

The constraints simply state that when a leg has been activated, the received power at the UE must be larger than the sum of interference and noise. In contrast, when the leg is not activated for any modulation and coding scheme ($x_{ulmt} = 0$), the big M makes the constraint valid for any value of interference generated by other legs.

Comment: Please note that if the weighting parameter α in objective function (1) is set as $\alpha = \frac{1}{|\mathcal{L}||\mathcal{U}|+1}$, we first optimize the number of accepted users and then we carefully allocate radio resources to them to satisfy their required needs. This is a natural procedure that is in line with the operating mode and interests of mobile operators.

V. Problem decomposition and Heuristic

The PSD problem described in the previous section is NP-hard. Therefore, the computing time necessary to solve the problem increases steeply with the instance size. Even medium-size instances as those illustrated in Section VI require hours to be solved using state of the art solvers. This is essentially due to the constants α and M that make relaxation techniques inefficient.

In order to reduce the overall complexity, we propose a *decomposition approach* that consists in solving sequentially the *users' admission problem* and the *radio resource allocation problem*. The approach firstly fixes the z_u variables. Then, it determines the values of all other decision variables using fixed

z_u variables as input parameters. We observe that when the parameter α in objective function (1) is set as described at the end of Section IV (that is, $\alpha = \frac{1}{|\mathcal{L}||\mathcal{U}|+1}$), this procedure provides the optimal solution as the original PSD problem. Specifically, our decomposition procedure for solving the PSD problem proceeds as follows:

1) Solve problem (1)-(8) with the following objective function to be *maximized* (the number of accepted users): $\sum_{u\in\mathcal{U}} z_u$.
2) Fix z_u variables to the optimal values found in Step 1, which now become parameters, and then solve again the (1)-(8) problem with the following objective function to be *minimized* (the overall number of legs, hence of resources, allocated to the users accepted in step 1): $\sum_{u\in\mathcal{U},l\in\mathcal{L}_u} y_{ul}$.

Starting from the observation that in all considered scenarios and network instances a very limited number of UEs activates more than two legs, we further propose a heuristic to speed up the computation and find a good solution, which is close to the optimum. In particular, for each user u we limit the set of legs that can be used to serve the traffic only to the two with the strongest signal $P_{ul}G_{ul}$: $\mathcal{L'}_u = \{l_1, l_2\} \subseteq \mathcal{L}_u$, where $\forall i \in \mathcal{L}_u, i \neq l_1, l_2$ we have $P_{ul_1}G_{ul_1} > P_{ui}G_{ui}$, $P_{ul_2}G_{ul_2} > P_{ui}G_{ui}$, and $l_1 \neq l_2$.

VI. Numerical results

In this section, we illustrate the numerical results obtained solving the PSD problem to the optimum as well as with the decomposition procedure and heuristic detailed in the previous section. We first describe the experimental methodology employed in our simulations. Then we illustrate the performance of our proposed solutions.

In our simulations, we consider the 3GPP Urban Macro deployment scenario [2], where the network topology is composed of 7 macro Base-Stations, with 3 sectors each, inter site distance equal to 150 m, and a number of small cells per sector that we vary from 1 to 7. The position of each small cell is drawn at random from a uniform distribution. The transmission power is set to to 43 dBm and 26 dBm for macro and small cells, respectively. A number of UEs comprised in the $[2, 10]$ range are randomly scattered in each sector. Each UE is characterized by a traffic demand uniformly extracted at random in different ranges (from $[1, 5]$ to $[5, 9]$ Mbit/s) to take into account different bandwidth requirements.

For each scenario, we consider 15 random extractions, and we measure the average number of UEs accepted in the system, the number of activated legs and the time to compute the solution. All numerical results have been obtained on a server equipped with an Intel(R) Xeon(R) E5-2640 v4 CPU@2.40GHz and 126 GB of RAM. The optimal solution has been computed using CPLEX solver, using default values for its parameters.

We first provide a comparison of our proposed decomposition approach (*DA*) and heuristic with the Single Connectivity case, where only one leg (with the Master gNB) is activated for each UE, thus clearly quantifying the benefits that can be obtained implementing Multi Connectivity.

Then, we illustrate the results obtained with DA and heuristic in larger topologies, analyzing numerically the impact on the system performance of the number of legs that can be activated. Finally, we study the impact of the traffic offered to the network by each UE as well as of the number of installed small cells.

A. Comparison with Single Connectivity

In order to quantify the benefit that can be achieved by activating more than one leg for each user, both for the mobile operator that can serve a larger number of users and for the users themselves that can see their bandwidth demands satisfied, we implemented a variation of model (1)-(8) that forces UEs to be connected with a single leg to the best available gNB.

Tables II and III show the results measured for the single and dual connectivity scenarios, respectively, in a network with 5 to 7 small cells per sector, as well as 5 to 7 users per sector (hence with a total of 105 to 147 UEs, and 105 to 154 gNBs, respectively), each with a demand uniformly extracted in the interval $[3, 8]$ Mbit/s.

Single Connectivity	5 SCs	7 SCs
5 UEs	54.3	58.8
7 UEs	67.6	73.9

TABLE II: Average number of accepted users for the Single Connectivity case (with variable numbers of UEs and Small Cells).

Dual Connectivity	5 SCs	7 SCs
5 UEs	65.5	73.3
7 UEs	87.0	98.3

TABLE III: Average number of accepted users for the Dual Connectivity case (2 legs available per UEs, with variable numbers of UEs and Small Cells).

It can be observed that the number of served users when each UE can activate at most 2 legs (Dual Connectivity) is, in these scenarios, 20.6% (for the case with UEs=Small Cells= 5 per sector) up to 33% (for UEs=Small Cells= 7 per sector) higher than in the Single Connectivity scenario. The gain is, hence, more evident when a large number of (interfering) UEs is present in the network, as well as when a large number of available connections to gNBs (Master BSs and small cells) exist.

B. Analysis of multiple cell choices

Figures 2 and 3 show, respectively, the average number of accepted users and the average number of selected legs per UE as a function of the number of UEs per sector, when a different number of legs can be activated (either 2 or 5). Two scenarios are considered, with 2 and 5 small cells deployed in each sector. Each user offers to the network a demand randomly extracted in the 1 to 5 Mbit/s range. Curve label $LX.SCY$ indicates that each UE can use at most X *Legs* out of Y available *Small Cells*.

If we compare the two curves in the middle (L2.SC5 and L5.SC2) in Figure 2, we observe that the mobile operator can obtain similar results, in terms of number of served users, by either increasing the network coverage (here, from 2 to 5 small cells per sector) or activating more legs for each user (from 2 to 5). In dense scenarios, deploying more small cells becomes crucial to increase the network capacity. This is confirmed by the results of the two lower curves (L2.SC2 vs L2.SC5) in Figure 3. On average, having more installed Small Cells results in the activation of more legs per user and consequently serving more users. For example, when we have 10 UEs per cell, 2 Small Cells per sector and each UE can select up to 5 legs, the number of served users increases by 33% (134.1 versus 100.9) while the number of active legs grows by 13% (1.31 with respect to 1.16).

C. Analysis of demand and number of small cells

In this set of tests, we increased the traffic offered to the network by each UE, varying it in the interval 1 to 9 Mbit/s. Specifically, five traffic scenarios are defined: D1, where each UE has a traffic demand randomly extracted from a uniform distribution between 1 and 5 Mbit/s, D2 (between 2 and 6), D3 (between 3 and 7), D4 (between 4 and 8), and D5 (between 5 and 9 Mbit/s). The results are shown in Figures 4,5,6, for different network scenarios, where we increase the number of small cells in each sector from 1 to 7 (recall that there are 21 sectors in total in our scenario), while the number of UEs per sector is set to 5 (for a total of 105 UEs).

The *number of accepted users* depends, of course, on the bandwidth demand of each user, as well as on the number

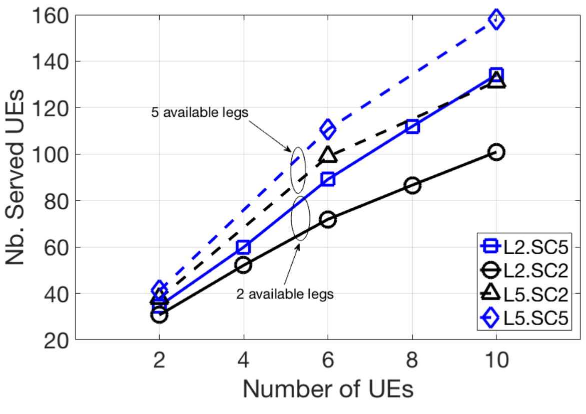

Fig. 2: Average number of served UEs as a function of (i) the number of UEs per sector (x axis), (ii) the number of small cells per sector (2,5) and (iii) the number of legs that can be activated (2,5).

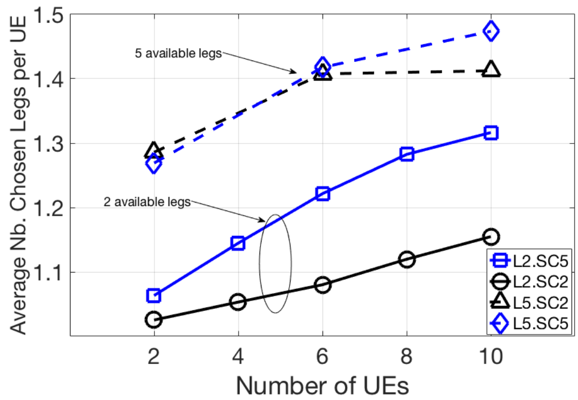

Fig. 3: Average number of selected legs as a function of (i) the number of UEs per sector (x axis), (ii) the number of small cells per sector (2,5) and (iii) the number of legs that can be activated (2,5).

of small cells deployed in each sector, and it is shown in Figure 4. For instance, when 3 small cells are deployed in each sector, our model accepts 61.5 users, in average, for the lowest demand ([1..5] Mbit/s), 64.7 for the medium one ([3..8] Mbit/s) and 71 for the highest demand ([5..9] Mbit/s). If we focus on a given demand scenario, the increase in the number of accepted users is from 62.6 to 83.27 (33%, for the lowest demand, scenario D1) and from 48.9 to 76.4 (i.e. 56%, for the highest demand, scenario D5).

The average *number of legs* assigned to each user increases with both the demand and the number of available small cells per sector, as Figure 5 illustrates. Such increase is more evident when both these parameters increase; specifically, the increase is of the order of 12.2% when passing from 1 to 7 small cells with the lowest demand (scenario D1), and up to 18.7% with the highest one (scenario D5).

Finally, the average *computing time* necessary to obtain a solution, expressed in seconds, is shown in Figure 6. Overall, in all scenarios where at least 3 small cells are deployed, the computing time is indeed limited (at most up to 740 seconds, in average). In general, when the number of available small cells per sector increases, more potential connections (and, consequently, an increased network capacity) are available for each user with respect to the case in which only few of them are deployed (the most demanding case is that of a single small cell); in the first case our heuristics can find a solution more easily. In general, in real 5G network scenarios, it is expected that a significant number of small cells will be deployed (e.g., at least 4 per sector [13]), hence our proposed solution approaches represent a viable solution for planning next-generation mobile networks.

VII. Conclusions

In this paper we developed an optimization framework for solving the PDCP Split-Bearer Decision problem in 5G

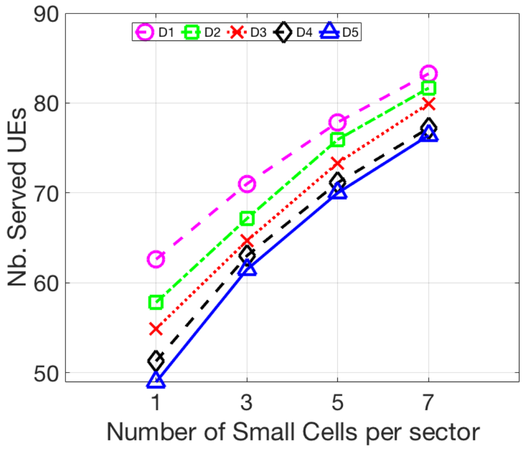

Fig. 4: Number of served users as a function of the number of small cells per sector and the UEs' demand (five scenarios with increasing demand, D1, D2, D3, D4 and D5, are considered).

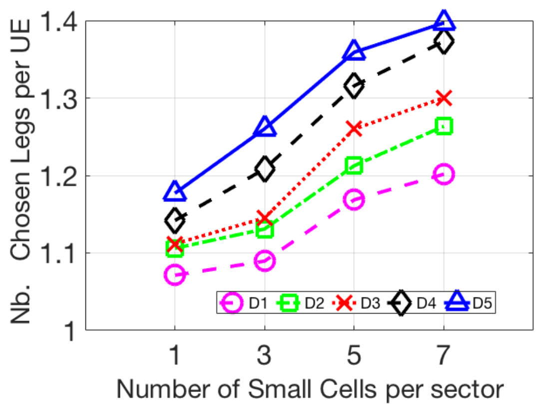

Fig. 5: Number of chosen legs per UE as a function of the number of small cells per sector and the UEs' demand.

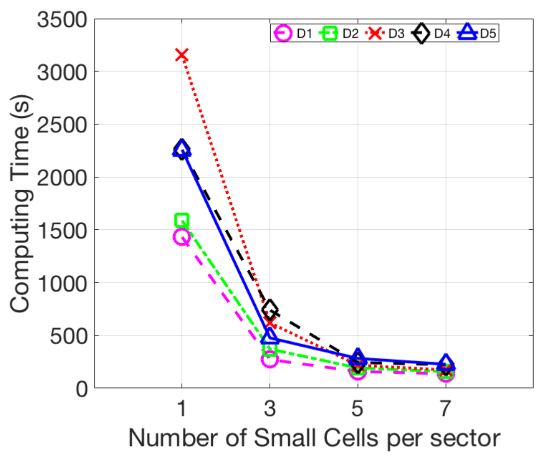

Fig. 6: Computing time in seconds as a function of the number of small cells per sector and the UEs' demand.

networks where Multi-Connectivity is enabled. The problem consists in deciding which users to admit in the network, and whether and how to split the traffic across multiple legs to meet the bandwidth requirements of the services.

We proposed effective, exact approaches that achieve the optimal solution as well as a decomposition (two-step) approach that, coupled with a carefully chosen limitation to the set of legs that each user can exploit, permits to consistently reduce the computing time while still achieving close-to-optimum solutions, specifically for real-size network scenarios.

We performed an extensive numerical analysis, considering realistic cellular deployments and traffic scenarios, which allowed us to capture and quantify the trade-off mobile operators face during network operation between admitting more UEs and providing the necessary resources to respect stringent quality requirements, specifically in terms of allocated rate on the radio interface and for all different transmission legs.

REFERENCES

[1] Joern Krause et al. 3GPP Technical specification 37.340: NG-RAN; Multi-connectivity.

[2] 3GPP TR 36.872. Small cell enhancements for E-UTRA and E-UTRAN - physical layer aspects. Dec. 2013, v:12.1.0.

[3] A. Ravanshid, P. Rost, D.S. Michalopoulos, V.V. Phan, H. Bakker, D. Aziz, S. Tayade, H.D. Schotten, S. Wong, and O. Holland. Multi-connectivity functional architectures in 5G. In *IEEE International Conference on Communications Workshops (ICC)*, pages 1–6, 2016.

[4] Yang Yang, Xinyi Deng, Dazhong He, Yanan You, and Ruoning Song. Machine Learning Inspired Codeword Selection For Dual Connectivity in 5G User-Centric Ultra-Dense Networks. *IEEE Transactions on Vehicular Technology*, 68(8):8284–8288, 2019.

[5] Lingxia Wang, Chungang Yang, and Rose Qingyang Hu. Autonomous traffic offloading in heterogeneous ultra-dense networks using machine learning. *IEEE Wireless Communications*, 26(4):102–109, 2019.

[6] J. Rao and S. Vrzic. Packet Duplication for URLLC in 5G: Architectural Enhancements and Performance Analysis. *IEEE Network*, 32(2):32–40, March 2018.

[7] Adnan Aijaz. Packet Duplication in Dual Connectivity Enabled 5G Wireless Networks: Overview and Challenges. *IEEE Communications Standards Magazine*, 3(3):20–28, 2019.

[8] N.H. Mahmood, A. Karimi, G. Berardinelli, K.I. Pedersen, and D. Laselva. On the resource utilization of multi-connectivity transmission for URLLC services in 5G New Radio. In *IEEE Wireless Communications and Networking Conference Workshop (WCNCW)*, pages 1–6, 2019.

[9] Jaya Rao and Sophie Vrzic. Packet duplication for URLLC in 5G dual connectivity architecture. In *IEEE Wireless Communications and Networking Conference (WCNC)*, 2018.

[10] Nurul Huda Mahmood and Hirley Alves. Dynamic multi-connectivity activation for ultra-reliable and low-latency communication. In *IEEE International Symposium on Wireless Communication Systems (ISWCS)*, pages 112–116, 2019.

[11] Xinran Ba and Yafeng Wang. Load-aware cell select scheme for multi-connectivity in intra-frequency 5G ultra dense network. *IEEE Communications Letters*, 23(2):354–357, 2019.

[12] Konstantinos Alexandris, Chia-Yu Chang, Kostas Katsalis, Navid Nikaein, and Thrasyvoulos Spyropoulos. Utility-based resource allocation under multi-connectivity in evolved LTE. In *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pages 1–6, 2017.

[13] Nurul Huda Mahmood, Melisa Lopez, Daniela Laselva, Klaus Pedersen, and Gilberto Berardinelli. Reliability Oriented Dual Connectivity for URLLC services in 5G New Radio. In *15th IEEE International Symposium on Wireless Communication Systems (ISWCS)*, 2018.