

Unified Scheduling Framework for Dual Connectivity in Multi-RAT NSA 5G Networks

Roberto P. Antonioli, Diego A. Sousa, Emanuel B. Rodrigues,
Igor M. Guerreiro, Carlos F. M. e Silva, Tarcisio F. Maciel and Fco. Rodrigo P. Cavalcanti

Abstract—Supporting the dual connectivity technology in non-stand-alone 5G scenarios involves three main functionalities: a user device (UD)-base station (BS) association (UBA) mechanism, a flow control algorithm (FCA) to determine the data split transmitted by each BS and radio resource management (RRM) algorithms. However, previous works only focused on the design of at most two of these functionalities. In this letter, we go beyond and design a unified scheduling framework including the three aforementioned functionalities. The proposed framework targets the user satisfaction maximization and runs in a completely decentralized fashion, where the UBA algorithm runs independently at each UD, the FCA is executed at the master LTE BS and the RRM is performed individually by each BS. Simulations show the gains of the proposed solution compared to existing solution that addressed at most two functionalities.

Index Terms—Non-stand-alone 5G, dual connectivity, resource management, flow control algorithm, UD-BS association.

I. INTRODUCTION

The dual connectivity (DC) technology has been considered as a promising solution to meet the predicted requirements of fifth generation (5G) applications, such as ultra high throughput, enhanced mobility and high reliability. Powered by this technology, a user device (UD) can be connected to two base stations (BSs) simultaneously. An envisioned solution for DC deployment is the bearer split configuration using non-stand-alone (NSA) 5G [1], which consists of secondary BSs employing the new radio (NR) radio access technology (RAT) anchored at a master long term evolution (LTE) BS. Such configuration relies on the tight interworking between the fourth generation (4G) and 5G standards, allowing a smooth transition from the legacy to the new technology.

Fully harvesting the gains provided by DC in bearer split NSA 5G scenarios involves the design of three main functionalities [2]: i) a mechanism for the UD-BS association (UBA); ii) a flow control algorithm (FCA) for splitting the data traffic between BSs; and iii) radio resource management (RRM) techniques, such as radio resource allocation (RRA) and power control.

Previous works focused only on one or at most two of the aforementioned functionalities simultaneously. For instance, in [3], [4], [5], [6], [7] the authors studied only mechanisms for controlling the UBA, in [8], [9], [10], [11] the authors only proposed FCAs and [12] only proposed a RRM algorithm. In [13], the authors focused only on the RRA and proposed

a modified version of the traditional proportional fair (PF) scheduling algorithm for DC scenarios. A joint FCA and RRA scheme for throughput maximization was proposed in [14], whereas the authors in [15] studied a sub-optimum solution for a utility maximization optimization problem focusing on the FCA and UBA. In [16], a mechanism for UBA and RRA was proposed for maximizing the minimum user rate. However, since it is not possible to avoid the deployment of all three functionalities in NSA 5G networks with DC, those works did not propose suitable frameworks for such scenarios. Thus, complementary functionalities need to be incorporated, which do not guarantee that the whole framework would effectively work targeting a common objective.

Differently from these works, this letter proposes the design of a unified scheduling framework involving the three functionalities. It has the advantage that all functionalities work together towards the same objective to improve the system performance. Our proposal is completely decentralized, where the UBA algorithm runs independently at each UD, the FCA is executed at the master LTE BS and the RRM is performed individually by each BS. Using computational simulations, we demonstrate the gains provided by the proposed complete framework compared to existing solution that addressed at most two functionalities..

II. SYSTEM MODEL

Let us consider the downlink of a system serving a set of UDs denoted by \mathcal{J} and comprised of a set \mathcal{N} of RATs, where $|\mathcal{J}| = J$ and $|\mathcal{N}| = N$. The set of BSs employing RAT n is denoted by \mathcal{B}_n , where $|\mathcal{B}_n| = B_n$, $\mathcal{B} = \bigcup_{n \in \mathcal{N}} \mathcal{B}_n$ and $|\mathcal{B}| = B$.

We consider a resource block (RB) as the minimum allocable unit, which is comprised of adjacent subcarriers in the frequency domain and orthogonal frequency division multiplexing (OFDM) symbols spanning one transmission time interval (TTI) in the time domain. All BSs from RAT n serve their users using the same set of RBs denoted by \mathcal{K}_n , where $|\mathcal{K}_n| = K_n$, $\mathcal{K} = \bigcup_{n \in \mathcal{N}} \mathcal{K}_n$ and $|\mathcal{K}| = K$.

It is assumed that each BS from RAT n disposes of a total power P_n , which is equally distributed among its RBs, such that the power allocated by the BS b of RAT n on RB k to transmit to user j is given by $p_{j,k,b} = \frac{P_n}{K_n}$. The signal to interference-plus-noise ratio (SINR) of user j on RB k of the BS b is given by $\gamma_{j,k,b}$. The BS uses a link adaptation scheme that allows different transmission rates depending on the $\gamma_{j,k,b}$ value, which is used by the BS b to select from a set of 15 modulation and coding schemes, the one that allows the BS to transmit to the UD at the highest data rate given that a predefined target block error rate (BLER) is met. Thus, the rate allocated by the BS b to the user j on RB k is $r_{j,k,b}$, which is a function of $\gamma_{j,k,b}$.

This work was supported by Ericsson Research, Technical Cooperation contract UFC.47. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 as well as by CNPq and FUNCAP.

The authors are with the Wireless Telecommunications Research Group (GTel), Federal University of Ceará (UFC), Fortaleza, Ceará, Brazil. E-mails: {antonioli, diego, emanuel, igor, cfms, maciel, rodrigo}@gtel.ufc.br.

Digital Object Identifier: 10.14209/jcis.2020.22

III. PROBLEM FORMULATION

We formulate a utility-based optimization problem to maximize the total user utility derived from the network, i.e., the total user satisfaction, where the concept of user satisfaction relates to meeting the users' quality of service (QoS) demands [17], [16]. Our optimization problem is:

$$\begin{aligned}
 & \text{maximize} \sum_{\rho_{j,k,b}} \sum_{n \in \mathcal{N}} \sum_{b \in \mathcal{B}_n} \sum_{j \in \mathcal{J}} M \left\{ V \left[U(x_{n,b,j}) \right] \right\} \\
 & \text{subject to} \sum_{j \in \mathcal{J}} \rho_{j,k,b} \leq 1, \forall k \in \mathcal{K}_b, \forall b \in \mathcal{B}, \\
 & \sum_{b \in \mathcal{B}} \mathbb{H} \left(\sum_{k \in \mathcal{K}} \rho_{j,k,b} \right) \lambda_{b,n} \leq l_{j,n}, \forall j \in \mathcal{J}, \forall n \in \mathcal{N}, \\
 & \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} \sum_{j \in \mathcal{J}} \rho_{j,k,b} \left(1 - \sum_{n \in \mathcal{N}} \lambda_{b,n} \omega_{k,n} \right) = 0, \\
 & \rho_{j,k,b} \in \{0, 1\}, \forall j \in \mathcal{J}, \forall k \in \mathcal{K}_n, \forall n \in \mathcal{N}, \forall b \in \mathcal{B},
 \end{aligned} \tag{1}$$

where: $\mathbb{H}(a, b)$ is the Heaviside function, which returns either 1, if $a \geq b$, or 0, otherwise; $\rho_{j,k,b}$ is an assignment variable that assumes the value 1 if the RB k of BS b from RAT n is assigned to user j and 0, otherwise; $\omega_{k,n}$ is a binary parameter equal to 1 if RB k can be used by the BSs of RAT n , i.e., $k \in \mathcal{K}_n$, and 0, otherwise; $\lambda_{b,n}$ is a binary parameter equal to 1 if the BS b belongs to the RAT n , i.e., $b \in \mathcal{B}_n$, and 0, otherwise; $l_{j,n}$ is a parameter that indicates the number of BSs of RAT n that UD j can be simultaneously connected to; $U(x_{n,b,j})$ is a user utility function based on a generic variable $x_{n,b,j}$ that represents a QoS metric of user j on BS b of RAT n ; $V(\cdot)$ is a service utility function that differentiates the services in the system; and $M(\cdot)$ is a novel RAT utility function that differentiates the RATs in the system.

The first and last set of constraints state that the RBs of each BS are discrete and that the same RB of a given BS cannot be shared by two or more users in the same TTI. The second set of constraints controls the number of connections of each UD, while the third set of constraints controls which RBs can be used by each BS from each RAT [16].

Due to the non-convexity of problem (1) [16], [18], we adopt a problem-splitting approach in which we design a distributed solution with UBA, FCA and RRA mechanisms.

IV. UNIFIED FRAMEWORK FOR NSA 5G WITH DC

Our framework enables network operators to answer the three following questions, which are related to each of the aforementioned functionalities:

- 1) What is the best connectivity solution for each UD in the system, i.e., should the UD be receiving data simultaneously from two BSs or only from one BS?
- 2) How should the traffic of each UD in DC be split between BSs so that the correct amount of data is forwarded to each BS without overloading or underutilizing their resources?
- 3) How many and which radio resources should be assigned for each UD such that the UDs in DC experience enhanced throughputs without leaving UDs with one active connection in starvation?

Let us now describe in more detail the proposed scheduling framework. We assume that there are two RATs in the system, i.e., $N = 2$ and that user u can be connected to one BS from

each RAT n , i.e., $l_{j,n} = 1, \forall j, n$. $x_{n,b,j}$ can be specialized to be several types of QoS metrics, such as throughput, packet delay or buffer size [19]. For instance, if $x_{n,b,j}$ is the throughput $T_{n,b,j}$ of user j on BS b of RAT n , it is given by:

$$T_{n,b,j}[t] = (1 - f_{\text{thru}}) \cdot T_{n,b,j}[t-1] + f_{\text{thru}} \cdot R_{j,b}[t], \tag{2}$$

where f_{thru} is a filtering constant, t is the time instant and $R_{j,b}[t] = \sum_{k \in \mathcal{K}_n} r_{j,k,b}[t] \cdot \rho_{j,k,b}$. In this case, the solution found for problem (1) using the utility theory is a closed form solution that selects the user j^* connected to BS b of RAT n to transmit on RB k according to [19]:

$$j_{b,n,k}^* = \arg \max_j \left\{ w_j^{b,n} \cdot w_j^s \cdot w_j \cdot r_{j,k,b} \right\}, \tag{3}$$

where: $w_j^{b,n}$ is the utility-based RAT weight associated to user j connected to the BS b of RAT n ; w_j^s is the utility-based service weight associated to user j and service s ; and w_j is the utility-based user weight associated to user j . The equations used to compute the utility-based weights are: $w_j^{b,n} = \frac{\partial M(V(U(T_{n,b,j}[t-1])))}{\partial T_{n,b,j}[t-1]}$, $w_j^s = \frac{\partial V(U(T_{n,b,j}[t-1]))}{\partial T_{n,b,j}[t-1]}$ and $w_j = \frac{\partial U(T_{n,b,j}[t-1])}{\partial T_{n,b,j}[t-1]}$. In fact, (3) represents a centralized solution, which is not our focus since we are interested in fully decentralized solutions. Thus, (3) is only used as an insightful starting point for the proposed decentralized framework.

The proposed framework is completely unified in the sense that $U(\cdot)$, $V(\cdot)$ and $M(\cdot)$ are based on a single utility function, i.e., the sigmoid function [11]. The sigmoid function was selected because it can be controlled using only two parameters, namely, a shape parameter and another parameter that controls whether it is an increasing or decreasing function. The easy control of these parameters is essential to adapt the utility weights in our solution. Moreover, this function was chosen because it yields higher performance compared to other utility functions from [20], as shown in [21], [19].

A. UD-BS Association

The proposed UBA algorithm, which runs independently at each UD, works based on the utility-based user weight w_j . Three inputs are used: signal strength, user QoS indicator and user QoS requirement. Such inputs are used by UDs to monitor both QoS and signal levels in order to detect bad QoS even when the signal quality is acceptable. The proposed UBA algorithm is detailed in Alg. 1. The main novelty in Alg. 1 is the new signaling needed in Step 3, which is exchanged between BSs and UDs. Moreover, the weights computation used for selecting the best BS in Steps 4 and 5 is also novel.

Algorithm 1 Proposed UBA algorithm.

- 1: **for** $n = 1$ to N **do**
 - 2: Check BSs with signal strength above a minimum threshold.
 - 3: Receive average UD throughput from those BSs.
 - 4: Compute w_j using the received average UD throughput values and its QoS requirement.
 - 5: Compute w_j times the signal strength for each BS.
 - 6: Connect to the BS with highest computed product in step 5.
 - 7: **end for**
-

The signaling in step 3 of Algorithm 1 is a new signaling required for our framework, which consists of BSs computing the average UD throughput and broadcasting this value via radio link control (RLC). In step 4, the UD uses this QoS

indicator (average UD throughput) and its QoS requirement to compute w_j , which is the utility of connecting to that BS. In step 6, the UD connects to the BS of each RAT providing the maximum value of the product between the signal strength and w_j compute in step 5. This procedure is constantly executed by the UDs, allowing UDs to be in DC or single connectivity (SC) depending on the links and QoS conditions.

After connecting to the BSs, the UD monitors its perceived QoS on each BS (at most one LTE and one NR) and constantly computes w_j for each BS. For UDs in DC, in case the utility drops below a predefined utility threshold, it means that the connection to that BS is no longer beneficial. Then, the UD temporarily disables the user plane to that BS, i.e., the UD goes into a fast switching (FS) mode [3]. The UD keeps monitoring its perceived QoS and if it increases to some percentage of the QoS requirement, the connection is reactivated. Otherwise, it triggers a handover or disconnection procedure from that BS because the connection is not providing enough QoS benefits.

Using such an approach, even when the signal strength is acceptable, monitoring the QoS enables the possibility of perceiving that one of the connections of a UD in DC is not beneficial. This is essential in overloaded scenarios, where maintaining the majority of the UDs in DC is not the best option. Consequently, this scheme reduces the number of UDs in DC when the system load increases, which in turn reduces the total interference in the system [3].

B. Radio Resource Management

The proposed RRA algorithm focuses on the downlink transmission and is executed independently at each BS, regardless of the RAT associated to that BS. The RRA procedure is describe in detail in Alg. 2, which is an extension for DC scenarios of [19]. The novelty is the exchange of data among the BSs regarding the UDs in DC, done in Step 1.

Algorithm 2 Proposed RRA algorithm.

- 1: Exchange throughput of UDs with more than one connection.
 - 2: Compute the QoS metric $x_{n,b,j}$ for each served UD j .
 - 3: Calculate w_j^s and w_j for each served UD j .
 - 4: **for** $k = 1$ to K_n **do**
 - 5: Compute the value of $r_{j,k,b}$ for each served UD j .
 - 6: Set $\rho_{j,k,b}$ using $j_{b,n,k}^* = \arg \max_j \{w_j^s \cdot w_j \cdot r_{j,k,b}\}$.
 - 7: **end for**
-

The utility-based user weight w_j has the important role of diminishing the priority for receiving resources for UDs experiencing QoS above their QoS requirement. This behavior is useful when the QoS of UDs in DC increases above their QoS requirement (which happens very quickly due to the resource aggregation from multiple connections), then their priority for receiving resources is reduced and UDs with one connection get some resources. The BS also computes the utility-based service weight w_j^s for each UD, which differentiates the priority of the different services in the system. The service utility is also modeled as a sigmoid function, which has its shape parameter dynamically adapted to protect the user satisfaction of a higher priority service [19].

Using the proposed RRA mechanism, network operators are able to maximize the user satisfaction by providing enhanced

throughput experiences for UDs in DC, without letting the UDs with one active connection starve. Therefore, the satisfaction of UDs in DC and UDs with one active connection (i.e., UDs in SC or FS) are simultaneously maximized.

C. Flow Control Algorithm

Besides using $r_{j,k,b}$, the utility-based user weight w_j and the utility-based service weight w_j^s , the proposed FCA also uses the utility-based RAT weight $w_j^{b,n}$ for determining the best split ratios for each UD in DC since it needs to distinguish the priority of each RAT. The FCA runs independently at each packet data convergence protocol (PDCP) layer of each LTE BSs. The proposed FCA algorithm is detailed in Alg. 3. This algorithm was initially proposed in [11] without any integration and unification with the other functionalities for DC scenarios, which are the novel aspects analyzed herein.

Algorithm 3 Proposed FCA.

- 1: Compute the QoS metric $x_{n,b,j}$ for each served UD j .
 - 2: Compute total achievable rate $\hat{R}_{j,b} = \sum_{k \in \mathcal{K}_n} r_{j,k,b}$ on LTE connection of each served UD j .
 - 3: Receive user satisfaction from its NR BS and total achievable rate on NR connection of users in DC.
 - 4: Compute $w_j^{b,n}$, w_j^s and w_j for each served UD j .
 - 5: Compute w_j^{LTE} and w_j^{NR} by multiplying $w_j^{b,n}$, w_j^s , w_j and $\hat{R}_{j,b}$ from the respective RATs.
 - 6: Compute $\text{ratio}_j^{\text{NR}} = \frac{w_j^{\text{NR}}}{w_j^{\text{NR}} + w_j^{\text{LTE}}}$ and $\text{ratio}_j^{\text{LTE}} = \frac{w_j^{\text{LTE}}}{w_j^{\text{NR}} + w_j^{\text{LTE}}}$.
 - 7: Send $\text{ratio}_j^{\text{NR}}$ of data from UD j via backhaul to the NR BS.
 - 8: Send $\text{ratio}_j^{\text{LTE}}$ of data from UD j to LTE BS lower layers.
-

The utility-based RAT weight $w_j^{b,n}$ is also modeled using the same sigmoid function used for w_j^s and w_j , showing the unification of our framework. The RAT utility differentiates the priority of the RATs in the system and is adapted in a similar manner as the service utility to meet a RAT-related predefined goal. Since the NR BS has higher transmission capacity than the LTE BS (due to the shorter TTI and higher bandwidth), the RAT utility is used to send as much data as possible to be transmitted by the NR BS, meaning that the data split ratio of the NR BS is higher than the data split ratio of the LTE BS. However, if the user satisfaction at the NR BS drops below a certain threshold, then the RAT utility is adapted to offload more data to the LTE BS. Using this approach, we exploit the higher transmission capacity of the NR BS without overloading it to a point where the QoS provision is damaged.

In step 6 of Algorithm 3, for each connection of UDs in DC, the FCA computes the weights w_j^{LTE} and w_j^{NR} by multiplying $w_j^{b,n}$, w_j^s , w_j and $\hat{R}_{j,b}$ from the respective RATs. Notice that there is one weight for each connection due to two reasons: 1) UDs experience a different channel on each link, which affects the achievable data rate on each link; 2) each RAT has a different weight defined by the RAT utility. In step 7 of Algorithm 3, the FCA computes the split ratios for each RAT, namely $\text{ratio}_j^{\text{NR}}$ and $\text{ratio}_j^{\text{LTE}}$, and then sends the corresponding amount of data to each BS. Using this approach, we track the channel quality and QoS experiences of the UDs in DC to compute the best data split ratios to be sent to each BS.

V. PERFORMANCE EVALUATION IN 4G-5G NETWORKS

The simulation environment is aligned with the bearer split configuration using NSA 5G presented in the 3rd Generation Partnership Project (3GPP) specifications [1]. The network topology is comprised of an hexagonal grid of three-sector macro LTE BSs and one NR BS randomly deployed on each sector of the LTE BSs [22]. Since there are two RATs in the system, $N = 2$, where $n = 1$ for the LTE and $n = 2$ for the NR. All BSs employ a link adaptation scheme based on a lookup table created from the link level curves in [23] for a target BLER of 10^{-4} . The UD's dropping criterion follows a hotspot UD distribution per LTE sector, where 75% are deployed within the NR coverage area and the other 25% are uniformly deployed within the LTE sector. Each UD j is configured to have one constant bit rate (CBR)-like bearer and can be connected to one BS from each RAT n , thus $l_{j,n} = 1, \forall j, n$. A UD is considered satisfied if its total throughput is higher than or equal to a throughput requirement of 20 Mbps [17], [16]. The f_{thru} is set to 0.001 [19]. Table I summarizes the adopted parameters used for setting the simulation environment. For more details about the simulation environment, see [2], [11].

TABLE I
SIMULATION PARAMETERS FOR LTE AND NR [1], [22].

Parameter	LTE ($n = 1$)	NR ($n = 2$)
Scenario	3GPP Urban Macro	3GPP Urban Micro
Inter-site distance	1000 m	-
BS height	25 m	10 m
Carrier frequency	3.5 GHz	28 GHz
System bandwidth	20 MHz	100 MHz
Subcarrier spacing	15 kHz	60 kHz
Num. of RBs (K_n)	100	125
TTI duration	1 ms	0.25 ms
Noise figure	9 dB	9 dB
BS Tx power (P_n)	49 dBm	35 dBm
Tx array antenna	8 x 1	4 x 4

We compare our solution with: 1) a baseline framework with the UBA algorithm from [3], the FCA from [8] and the PF RRA algorithm for DC [13]; 2) the joint UBA and RRA solution from [16] with the FCA from [8]; and 3) the solution from [16] with the FCA from [11]. Several key performance indicators (KPIs) are evaluated: 1) overall percentage of satisfied UDs, which is the number of UDs with throughput higher than the minimum throughput requirement over the total number of UDs [11], [16], [17]; 2) percentage of satisfied UDs considering the UDs in DC [11]; 3) percentage of satisfied UDs considering the UDs in SC [11]; 4) total cell throughput [11], [16], [17]; 5) percentage of UDs in SC [11], [16]; and 6) the Jain's fairness index.

Fig. 1 presents the overall percentage of satisfied UDs. One can see that the three benchmark solutions perform equally worse than the proposed algorithm. This occurs because those solutions did not develop a complete framework for DC scenarios. The solution from [16] jointly performs the UBA and RRA targeting a common objective of maximizing the minimum user rate, but without the FCA. We added two different FCAs from [8] and from [11], but the performances of the combined solutions did not achieve a performance close to our solution, showing the importance of having the three functionalities working with the same objective. Meanwhile,

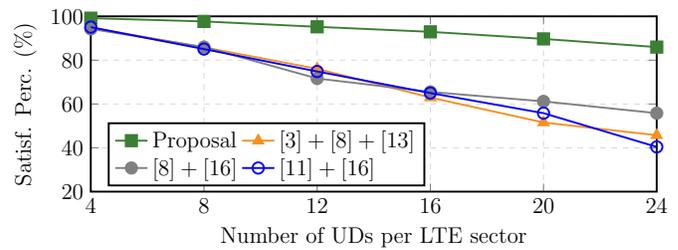


Fig. 1. Percentage of satisfied UDs when increasing the number of UDs.

our complete framework develops the three functionalities with the same objective. Our FCA balances the data splits such that some data is offloaded to the NR BSs when needed, which reduces the burden on the LTE BSs and, consequently, increases the QoS provision and user satisfaction. The RRA algorithm considers the UDs QoS to balance the allocation among UDs in SC and DC, thus simultaneously increasing the satisfaction of all UDs. Finally, the UBA controls the number of active connections by also considering the level of QoS perceived by the UDs. Altogether, this results in gains of up to 18% in terms of user satisfaction.

Finally, Fig. 2 depicts a radar chart from which we simultaneously consider all six aforementioned KPIs for a given load of UDs, which was chosen to be of 16 UDs per LTE sector. Our framework achieves a good balance for all performance metrics, which can be seen from the fact that all metrics are close to 100%. On the other hand, the comparison algorithms highly degrade the satisfaction of UDs in SC attempting to satisfy the UDs in DC, for instance. The balanced performance accomplished by the proposed unified framework occurs due to the design of a framework with a common objective and that considers the important tasks that each functionality should perform.

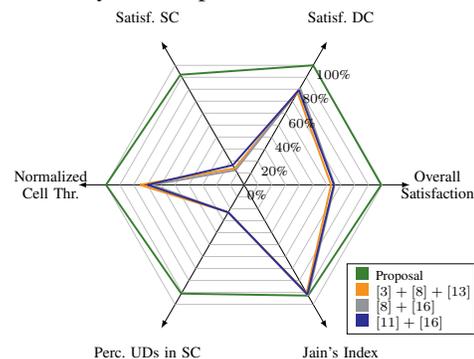


Fig. 2. Radar chart illustrating a comparison between 5 different performance metrics for a system load of 16 UDs per LTE sector.

VI. CONCLUSIONS

The proposed framework maximizes the user satisfaction in the system without penalizing the total throughput. This is accomplished by: i) diminishing the percentage of UDs in DC when the system load increases to reduce the interference; ii) controlling the data split ratios such the transmit buffers are neither overload nor underutilized; and iii) assigning RBs in a balanced manner to maximize the satisfaction of both UDs in DC and SC. Thus, our unified framework exploits the essential behaviors needed in NSA 5G scenarios with DC and maximizes the system performance. As a future work, one could investigate how to solve problem (1) using other techniques, such as game theory or machine learning.

REFERENCES

- [1] 3GPP, "3GPP specification series: Radio technology beyond LTE," 3rd Generation Partnership Project (3GPP), Tech. Rep., 2019.
- [2] R. P. Antonioli, G. C. Parente, C. F. M. e Silva, D. A. Sousa, E. B. Rodrigues, T. F. Maciel, and F. R. P. Cavalcanti, "Dual connectivity for LTE-NR cellular networks: Challenges and open issues," *Journal of Commun. and Information Systems*, vol. 33, no. 1, Aug. 2018, doi: 10.14209/jcis.2018.28.
- [3] V. F. Monteiro, M. Ericson, and F. R. P. Cavalcanti, "Fast-RAT scheduling in a 5G multi-RAT scenario," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 79–85, Jun. 2017, doi: 10.1109/MCOM.2017.1601094.
- [4] A. Awad *et al.*, "Dynamic network selection in heterogeneous wireless networks: A user-centric scheme for improved delivery," *IEEE Consum. Electron. Mag.*, vol. 6, no. 1, pp. 53–60, Jan. 2017, doi: 10.1109/MCE.2016.2614419.
- [5] H. Wang, G. Pocovi, C. Rosa, and K. I. Pedersen, "Configuration of dual connectivity with flow control in a realistic urban scenario," in *Proc. of the IEEE Vehic. Tech. Conf. (VTC)*, Sep. 2015, pp. 1–5, doi: 10.1109/VTCFall.2015.7390877.
- [6] C. Wang, Z. Zhao, Q. Sun, and H. Zhang, "Deep learning-based intelligent dual connectivity for mobility management in dense network," *CoRR*, vol. abs/1806.04584, Aug. 2018, <http://arxiv.org/abs/1806.04584>.
- [7] H. Lee, S. Vahid, and K. Moessner, "Machine learning based RATs selection supporting multi-connectivity for reliability," in *International Conference on Cognitive Radio Oriented Wireless Networks*. Springer, 2019, pp. 31–41, doi: 10.1007/978-3-030-25748-4_3.
- [8] 3GPP, "3GPP contribution R2-132859: Throughput evaluation and comparison of with and without UP bearer split," 3rd Generation Partnership Project (3GPP), Tech. Rep., Aug. 2013.
- [9] H. Wang, C. Rosa, and K. I. Pedersen, "Dual connectivity for LTE-advanced heterogeneous networks," *Wireless Networks*, vol. 22, no. 4, pp. 1315–1328, May 2016, doi: 10.1007/s11276-015-1037-6.
- [10] D. Pérez *et al.*, "Long term evolution-wireless local area network aggregation flow control," *IEEE Access*, vol. 4, pp. 9860–9869, Jan. 2017, doi: 10.1109/ACCESS.2016.2643690.
- [11] R. P. Antonioli *et al.*, "Adaptive bearer split control for 5G multi-RAT scenarios with dual connectivity," *Computer Networks*, vol. 161, pp. 183–196, Oct. 2019, doi: 10.1016/j.comnet.2019.07.005.
- [12] Y. Yang, X. Deng, D. He, Y. You, and R. Song, "Machine learning inspired codeword selection for dual connectivity in 5G user-centric ultra-dense networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8284–8288, Aug. 2019, doi: 10.1109/TVT.2019.2923314.
- [13] Y. Wang, K. I. Pedersen, T. B. Sorensen, and P. E. Mogensen, "Utility maximization in LTE-advanced systems with carrier aggregation," in *Proc. of the IEEE Vehic. Tech. Conf. (VTC)*, May 2011, pp. 1–5, doi: 10.1109/VETECS.2011.5956494.
- [14] M. Pan, T. Lin, C. Chiu, and C. Wang, "Downlink traffic scheduling for LTE-A small cell networks with dual connectivity enhancement," *IEEE Commun. Lett.*, vol. 20, no. 4, pp. 796–799, Apr. 2016, doi: 10.1109/LCOMM.2016.2522404.
- [15] S. Singh, S. p. Yeh, N. Himayat, and S. Talwar, "Optimal traffic aggregation in multi-RAT heterogeneous wireless networks," in *Proc. of the IEEE Internat. Conf. on Commun. (ICC)*, May 2016, pp. 626–631, doi: 10.1109/ICCW.2016.7503857.
- [16] V. F. Monteiro *et al.*, "Distributed RRM for 5G multi-RAT multiconnectivity networks," *IEEE Syst. J.*, vol. 13, no. 1, pp. 192–203, Mar. 2019, doi: 10.1109/JSYST.2018.2838335.
- [17] R. P. Antonioli, G. Fodor, P. Soldati, and T. F. Maciel, "User scheduling for sum-rate maximization under minimum rate constraints for the MIMO IBC," *IEEE Wireless Commun. Lett.*, vol. 8, no. 6, pp. 1591–1595, Dec. 2019, doi: 10.1109/LWC.2019.2930255.
- [18] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks-part II: algorithm development," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 625–634, Mar. 2005, doi: 10.1109/TWC.2004.843067.
- [19] R. P. Antonioli, E. B. Rodrigues, T. F. Maciel, D. A. Sousa, and F. R. P. Cavalcanti, "Adaptive resource allocation framework for user satisfaction maximization in multi-service wireless networks," *Telecommun. Syst.*, vol. 68, no. 2, pp. 259–275, Jun. 2018, doi: 10.1007/s11235-017-0391-3.
- [20] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 678–700, May 2013, doi: 10.1109/SURV.2012.060912.00100.
- [21] E. B. Rodrigues, F. R. M. Lima, T. F. Maciel, and F. R. P. Cavalcanti, "Maximization of user satisfaction in OFDMA systems using utility-based resource allocation," *Wireless Commun. and Mob. Computing*, vol. 16, no. 4, pp. 376–392, Mar. 2016, doi: 10.1002/wcm.2526.
- [22] 3GPP, "Study on scenarios and requirements for next generation access technologies," 3rd Generation Partnership Project (3GPP), TR 38.913, Aug. 2017, v.14.3.0.
- [23] C. Mehlführer *et al.*, "Simulating the long term evolution physical layer," in *Proc. of the European Signal Process. Conf. (EUSIPCO)*, Glasgow, Scotland, Aug. 2009, pp. 1471–1478.