IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 3, 2021, accepted October 19, 2021, date of publication October 22, 2021, date of current version October 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3122331

Dynamic Spectrum Allocation Following Machine Learning-Based Traffic Predictions in 5G

RAKIBUL ISLAM RONY, (Student Member, IEEE), ELENA LOPEZ-AGUILERA, AND EDUARD GARCIA-VILLEGAS

Department of Network Engineering, Universitat Politécnica de Catalunya, BarcelonaTech, 08034 Barcelona, Spain

Corresponding author: Rakibul Islam Rony (rakibul.islam.rony@upc.edu)

This work was supported in part by the EU Horizon 2020 Research and Innovation Program (5GAuRA) under Grant 675806, and in part by the Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement from the Generalitat de Catalunya under Grant 2017 SGR 376.

ABSTRACT The popularity of mobile broadband connectivity continues to grow and thus, the future wireless networks are expected to serve a very large number of users demanding a huge capacity. Employing larger spectral bandwidth and installing more access points to enhance the capacity is not enough to tackle the stated challenge due to related costs and the interference issues involved. In this way, frequency resources are becoming one of the most valuable assets, which require proper utilization and fair distribution. Traditional frequency resource management strategies are often based on static approaches, and are agnostic to the instantaneous demand of the network. These static approaches tend to cause congestion in a few cells, whereas at the same time, might waste those precious resources on others. Therefore, such static approaches are not efficient enough to deal with the capacity challenge of the future network. Thus, in this paper we present a dynamic access-aware bandwidth allocation approach, which follows the dynamic traffic requirements of each cell and allocates the required bandwidth accordingly from a common spectrum pool, which gathers the entire system bandwidth. We perform the evaluation of our proposal by means of real network traffic traces. Evaluation results presented in this paper depict the performance gain of the proposed dynamic access-aware approach compared to two different traditional approaches in terms of utilization and served traffic. Moreover, to acquire knowledge about access network requirement, we present a machine learning-based approach, which predicts the state of the network, and is utilized to manage the available spectrum accordingly. Our comparative results show that, in terms of spectrum allocation accuracy and utilization efficiency, a well designed machine learning-based bandwidth allocation mechanism not only outperforms common static approaches, but even achieves the performance (with a relative error close to 0.04) of an ideal dynamic system with perfect knowledge of future traffic requirements.

INDEX TERMS 5G, automation, machine learning, BW allocation, spectrum sharing.

I. INTRODUCTION

Ever since the inception of cellular networks, the popularity of wireless connectivity is increasing everyday. Every generation of cellular networks, i.e., 1st, 2nd, 3rd and 4th Generation (1G, 2G, 3G and 4G) improved the performance of the previous generation in terms of capacity, coverage, data rate, etc, to meet the expectations of an evolving and connected society. Moreover, with the increasing popularity of mobile devices, tactile internet applications, video streaming and multi-fold varieties of use cases (e.g., broadband access everywhere, higher user mobility, extreme real-time communications, ultra-reliable communications), wireless networks are becoming more popular as a cost-effective solution for ubiquitous connectivity. In that sense, the latest generation, that is, the 5th Generation (5G), is considered revolutionary, since it is promising a level of services and facing new challenges like never before. For instance, 5G is expected to provide around one hundred times more capacity (i.e., 10 Mbps per m^2) to a hundred times more connected devices (i.e., 10^6 connected devices per Km^2) compared to 4G. Additionally, the expected end to end latency for delay-sensitive applications is expected to be less than 1ms [1]. According to [2], in 2020, the number of mobile

The associate editor coordinating the review of this manuscript and approving it for publication was Nafees Mansoor.

subscriptions has totaled 7.9 billion globally, and mobile data traffic grew by 49% within the previous year. By the end of 2026, the amount of mobile subscriptions is expected to be 8.8 billion, among which, 88% will consist of mobile broadband connections. Furthermore, with this mobile broadband connectivity, each user will be expecting a huge capacity, i.e., 25 Gigabyte (GB) on average per month [2].

To meet the aforementioned demands (high capacity, low latency, etc.), provide ubiquitous coverage, and be able to serve different scenarios (Internet of Things (IoT), Device to Device (D2D), Vehicle to everything (V2X), enhanced mobile broadband (eMBB) communications, etc.), wireless networks are becoming more complex everyday. Disrupting features of future wireless networks, such as Massive Multiple Input Multiple Output (MIMO), multi layer heterogeneous Radio Access Technologies (Multi-RATs), Centralized Radio Access Network (CRAN), Network slicing, Edge computing, Network Function Virtualization (NFV), etc. are the key enablers of 5G to meet the promised level of Quality of Service (QoS). Thus, implementing highly dense, complex and multi layered 5G networks will require a higher degree of automation [3]. Although, according to [4], the existing 5G networks do not provide such level of flexibility or automation yet, Artificial Intelligence (AI) offers solutions to tackle the complexity of 5G and beyond [3], [4]. Among many other AI techniques, (e.g., autonomous vehicles, robotics, computer vision, etc.), Machine Learning (ML) is arguably the most convenient mechanism since it depends on the availability of large amounts of data, something that abounds in a modern mobile network.

On the other hand, cell densification has always been the most effective and fastest way to increase the area capacity of the network, however, this process is costly and introduces additional challenges, i.e., interference management. Deployment of small cells (SC) and techniques like eICIC are popular solutions for the aforementioned cost and interference challenges, respectively. Largely deployed SCs along with network-wide deployed IoT devices will result into a Ultra Dense Network (UDN) [5]. The aggregated capacity requirement of such UDN will be enormous and potentially unbearable for current capacity enhancement techniques, such as carrier aggregation or Frequency Reuse (FR), because of poor frequency resource management: while some coverage areas are overprovisioned, others appear overloaded. Thus, along with these techniques, it is important to ensure the best usage of the scarce frequency resources by allocating them according to the network's varying demand.

As our contribution, we provide an intelligent access-aware dynamic spectrum allocation technique whereby resources are shared from a common spectrum pool. First, to ensure the best utilization of the available spectrum, we propose an allocation mechanism that follows the instantaneous capacity requirements in the access network. Then, we analyse the adaptation of a ML technique to predict the capacity requirements throughout different times of the day. We compare the performance of our methods with common static approaches and with an ideal (though unfeasible) dynamic mechanism.

The remaining of this paper is structured as follows. In Section II, we review the state of the art by presenting a detailed discussion on different spectrum allocation techniques, and the growing popularity of different ML techniques in wireless networks. In Section III, we present the system model and describe the proposed dynamic access-aware spectrum allocation technique. In Section IV, the description of the evaluation scenario, the related simulation assumptions and comparative traditional approaches are provided. Finally, in Section V, detailed analysis and comparative results for the proposed intelligent access-aware spectrum allocation technique are presented. Additionally, Section V analyses the performance of the ML-approach, which provides bandwidth assignments according to traffic predictions. Section VI concludes the paper.

II. STATE-OF-THE-ART

5G is expected to use CRAN for its deployment, where most of the RAN functionalities will be centralized into a Central Unit (CU), and the Access Points (APs) will be left only with basic radio frequency functionalities. This centralized approach allows controlling the network from a central controller, having the network-wide knowledge. Additionally, it is expected that in future networks, the available spectrum will be pooled together into a spectrum pool, which is shared by all the APs [5]. This spectrum pool will be controlled by a Software Defined Network (SDN)-based central controller, which allows the functionality to control and distribute Bandwidth (BW) among different APs according to their current requirements. Furthermore, to be resource-efficient in 5G complex UDN, it will be necessary to learn the network behaviour and predict the future requirements such that the BW distribution approach can be made ahead of the dynamic requirements and be ready with the correct resource distribution. Therefore, in this era of automation, future cellular networks are expected to adopt ML techniques to perform dynamic resource allocations. In this section, we discuss different spectrum allocation and ML approaches.

A. SPECTRUM ALLOCATION TECHNIQUES

Licensed spectrum has always been the most expensive and scarce resource from wireless networks service provider's point of view. With the aggressive growing of capacity demand on wireless networks, the spectrum became more precious, given that the most straightforward way to meet the higher capacity requirements is to increase the assigned BW [6]. However, this approach also brings additional challenges. BW distribution of the spectrum among different cells for ensuring the best utilization of the scarce resources has always been a very challenging task.

Throughout the previous generations of cellular networks, Fixed Spectrum Allocation (FSA) has been a popular approach for BW distribution among different APs. In this solution, an initial capacity plan is performed, and BW is allocated to the APs according to the maximum requirements it is expected to serve during peak hours, and to the available resources. Later on, BW allocation remains static and agnostic to the dynamic capacity requirements of the different APs in the network [7].

Compared to FSA, in Dynamic Spectrum Allocation (DSA), allocation of spectrum follows the instantaneous requirements of the APs, according to the available resources. A simple approach is presented in [7], where a DSA is triggered after a particular time interval, which estimates the load for the next time interval, computes the spectrum requirements and, if needed, allocates additional spectrum for the subsequent time interval. In [7], authors argued that DSA can be potentially used to share the idle spectrum in a multi-operator environment sharing a spectrum pool.

DSA technique has been very popular since its inception as it increases the spectral efficiency of the network by allowing the under-utilized frequency bands to be employed in an efficient way [8]. With the evolution of wireless technologies, DSA techniques also evolved. For example, different approaches emerged in the Cognitive Radio (CR) environment, where DSA is used to allocate the idle channels to Secondary Users (SU), which usually are unlicensed and have lower priority in the network [9]–[13], based on Fuzzy logic, Q-learning, randomized rounding algorithm, etc., to learn, estimate and allocate the required spectrum.

In [14], authors proposed a Reinforcement Learning (RL)-based DSA technique for spectrum allocation to the IoT users in a cellular network. In this proposal authors successfully showed that DSA technique can be used to identify the under-utilized spectrum in the network to be reused for a sensor-aided IoT network, subsequently, enhancing the spectrum re-usability.

In [15], RL-based DSA technique is used, where each cell can take its own decision of spectrum allocation to its users with the objective of maximization the overall SINR. In this work, authors tested the DSA technique in a decentralized approach, where each cell acts as an individual DSA agent, although collecting the spectrum allocation information from the neighbouring cell or DSA agent. In another work [16] from the same authors, RL-based DSA technique was tested in a centralized approach, where a central DSA agent controls and takes the spectrum allocation decision for all the cells in a considered area. The two studies showed that DSA-based spectrum allocation technique can outperform traditional FR techniques, both in an homogeneous (i.e., macrocell only), or heterogeneous (coexistence of macrocell and femtocell/SC together) scenario. However, both the presented works utilize DSA technique to minimize the intercell interference in OFDMA networks. In this paper, however, we propose a technique similar to DSA, where spectrum allocation to the cells follows the access network's dynamic requirements of each cell in a real network scenario.

ML-based Dynamic Frequency and Bandwidth Assignment (DFBA) focusing on spectrum allocation to the small cells in a cellular network was studied in [17]. The authors presented a technique to learn and predict Long Term Evolution (LTE) KPIs (e.g., SINR per resource block, Medium Access Control (MAC) level throughput, delay, etc.) and assign/rearrange spectrum allocation to the LTE-based SCs in the network. However, this work also does not consider the opportunity to enhance the spectral efficiency and system fairness by allocating the scarce spectrum according to the current load of the cells.

During the previous generations of cellular networks (i.e., 1G and 2G), compilation of a set of neighbouring cells were known as clusters, and the entire available BW was distributed among the cells within a cluster, avoiding overlapping portions of the spectrum. Thus, the whole available spectrum could be reused again in each cluster [18]. On the other hand, in newer generations of wireless networks, i.e., 4G and 5G, as part of the strategy to reach the required capacity, the technology allows the use a FR factor equal to one [18], which means that the same frequency band is reused in each cell in the network. The interference introduced due to such aggressive reuse is expected to be handled (i.e., controlled down to a minimum level) by new generation of advanced technologies, such as, eICIC, beamforming, etc. However, reusing the entire BW is neither efficient nor a fair approach since, in dynamic 5G networks, different APs will have distinct levels of load to serve.

Moreover, in future UDN, due to the wide variety of use cases, different sizes of APs (e.g., Macro Base station (MBS), SCs) will serve different numbers and types of users. Additionally, due to the mobility and higher user density, the capacity requirements of each AP can differ largely. Moreover, the number of served users or the amount of traffic carried by each AP will vary dynamically with the time of the day [19]. Thus, in 5G networks, the static allocation of BW is not an efficient approach, rather a dynamic solution is required, where, unlike the discussed related work, BW distribution must follow the current requirements of individual APs. In this paper, utilizing a spectrum pool, we adopt the DSA technique to allocate required BW to MBSs according to their current requirements, which, in our study, are based on real traffic traces. A similar approach was presented in [20], where authors address the problem of dynamic changes in the required capacity in a multi-service (i.e., cellular network, vehicular network and IoT) 5G network. In [20], authors presented how underused frequency bands from one service, e.g., IoT networks, can be requested and used by another service, e.g., vehicular networks, to meet individual dynamic capacity requirements during a congestion situation. Another approach of spectrum sharing can be found among coexisting different Mobile Network Operators (MNOs) to enhance the capacity of their network [21], [22].

In this paper, however, we consider each cell belonging to a single MNO, as an independent entity having its own dynamic capacity requirements. A spectrum pool, which we consider is managed by an intelligent SDN-based controller (similar to the centralized DSA approach in [16]), allocates and rearranges small portions of licensed spectrum bands (e.g., 5 MHz chunks) extracted from the entire available BW (e.g., 100 MHz). Even though 100% frequency reuse is allowed (i.e., FR = 1), our approach seeks to meet each cell's requirement while minimizing the bandwidth assigned. In this way, we also mitigate the effect of that aggressive frequency reuse, minimizing interference without sacrificing capacity. Additionally, utilizing a ML technique, we predict the throughput requirements of each cell and allocate BW according to the predicted behaviour. To validate the approach, we also compare real and predicted throughput requirements, and compare the BW allocation for real and predicted approaches. Such approach can be further studied in a multi-MNOs scenario, which we leave for future work.

B. MACHINE LEARNING IN WIRELESS NETWORKS

Benefits of using ML techniques in wireless communications can be multi-fold. Utilizing ML techniques, a network is able to analyze the behaviour, learn from it, predict future status, and prepare itself for it. In wireless networks, by relating the system parameters to the desired objective [23], ML techniques can address the challenge of traditional optimization approaches, which tend to leave a large gap between theoretical and real-time design of the network [24]. Therefore, for the development and automation of 5G networks, ML-based approaches are getting enormous attention and can be adopted in different aspects of cellular networks, e.g., interference management [24], beamforming [25]–[27], link quality estimation [28], 5G-based IoT [29], energy efficiency [3], [30], resource management [31], etc. There are different ML techniques: i) supervised learning, where the model learns by studying a set of labeled data, ii) unsupervised learning, where the model learns from a set of unlabeled data and, iii) reinforcement learning, where the model learns by assigning positive and negative rewards for its actions. Their corresponding different learning models (e.g., Support Vector Machines (SVM), K-means clustering, Gradient follower (GF), etc.), and their usage guideline for future wireless networks are well summarized in [32]–[34].

In this study, unlike the DSA techniques discussed earlier, we decouple the ML and DSA agent. We use a ML technique to predict the traffic, i.e., access network requirements, since reliable traffic prediction is already being considered as a key enabler for future wireless networks to improve QoS by reducing uncertainty [35]. The central controller collects the predicted traffic requirements of access network for each cell from ML agent and, after necessary calculations, it (the central controller) assigns/rearranges the spectrum allocation to each cell accordingly. In this way, the central controller offloads the ML complexity into a dedicated separated agent. Additionally, as discussed, most of the DSA techniques use RL-based learning, which require relatively large time, since it is a positive/negative reward-based process. Moreover, unlike RL-based DSA agents, our proposed central controller takes a faster decision following predicted requirements, since the proposed technique restricts the large number of spectrum allocation combinations among the cells by enforcing a few realistic constraints, as discussed in the next section. Therefore, our proposed spectrum allocation technique requires a simple time-series predictive ML tool.

In [36], utilizing collected trace data, authors have developed a model to predict the aggregated traffic and, subsequently, reduce the monitoring effort. Classification and clustering of cells were used to improve the traffic predictions in [3], [37]. However, these approaches require additional cell-level data. On the other hand, advanced ML-techniques such as Neural Networks (NN) reduce the dependency on additional features, and can be efficiently used to predict the traffic by analyzing a time series traffic data. A NN (aka. Artificial Neural Networks (ANN)), is a supervised ML technique, which emulates the way human brain works, by simulating artificial neurons with basic functionalities created from the complex computation [33]. A typical NN is composed of three types of functional layers, i.e., input layer, one or more hidden layers, and an output layer. All the layers consist of a set of nodes, which are connected with adjustable weight coefficients to each of the nodes in the next layer. The weights connect the input data (via input layer), to the activation and transfer function (inside hidden layer) and generates an output (in the output layer) if the weighted sum activates the neuron [38]. Since NN can solve non-linear complex problems by finding true relations between the input and output parameters, and confirms maximum level of generalization, NNs are very commonly used as a ML technique in recent studies [39]–[47].

In [48] authors have summarized a few studies on the benefit of NN network-based spectrum prediction techniques, and concluded that without needing much prior knowledge of the system, ANN show the best performance.

Long-Short Term Memory (LSTM), a kind of recurrent ANN, avoids long-term dependency on input [49] by using additional information about whether to remember or forget it. In [49], authors have presented LSTM-based multi-step traffic prediction of a LTE network. However, LSTM is more complex to implement compared to the classical NN. On the other hand, LSTM is more suitable to go beyond the available time series data and predict traffic requirements for one or more time-steps in the future. In this study, however, we focus on traffic prediction during the available time series, which can be validated on the existing data. According to [50], Autoregressive Neural Network (NARNET), a NN-based ML technique, is arguably one of the best ML technique to predict non-linear time series data. Additionally, in [51], authors showed that NARNET outperforms other studied techniques in terms of Coefficient of determination or R-squared (R2), which implies higher reliability for forecasts. More detail about NARNET is discussed in Section IV.B.

III. ACCESS-AWARE DYNAMIC SPECTRUM ALLOCATION AND SYSTEM MODEL

As discussed in Section II.A, following a static approach, spectrum is allocated to the APs without timely knowledge of the varying access network requirements, and it is usually based on the expected traffic during the peak hour. Subsequently, in most of the cases, either the spectrum is over-provisioned or under-provisioned [19]. In a dense 5G network, due to the aggressive capacity demand by a large number of connected devices, such static approach for spectrum allocation becomes unfair. Therefore, access-aware spectrum allocation, which allocates the spectrum to each of the APs following current requirements, is unavoidable. Such an approach is presented in [5], where access-aware spectrum allocation is applied to a two-tier heterogeneous network (i.e., coexistence of MBS and SCs), and SCs are anchored to the MBSs via wireless links. In such a complex network, three different types of links compete for the same pool of spectrum resources: i) direct link, the link between MBS and its users; ii) backhaul link, connecting the SCs to the anchored MBS and, iii) access link, the link between SC and its users. In [5], authors state that, in comparison with the traditional approaches (i.e., unaware of access demands), aforementioned access-aware spectrum allocation solution ensures a better network performance in terms of user throughput, spectrum allocation fairness and spectral efficiency. However, the access-aware approach presented in [5] can still be considered static, since the dynamic changes in the access network requirements are not taken into consideration. In this work, we propose a dynamic access-aware spectrum allocation, where BW allocation is performed following the dynamic changes in the access networks' requirements varying with time. Unlike [5], this work is based on a real deployment (cf. Section IV), where all APs can be considered, in fact, as MBS. Each MBS in this realistic model constitutes (one or) multiple cells. For this reason, the term MBS is henceforth used instead of the more generic AP.

A. SYSTEM MODEL

We assume, BS is the set of MBSs, each having nMC_i cells, so that $\sum_{i=1}^{numMBS} nMC_{\mathrm{i}} = numMC$, where $numMBS$ is the number of MBSs, and $numMC$ is the total number of cells. MC_i is then the set of nMC_i cells of the i^{th} MBS. B_i is the total BW assigned to that MBS, so that $B_i = \sum_{j=1}^{nMC_{\mathrm{i}}} b_{\mathrm{ij}}$, where b_{ij} are the elements of matrix B, representing the BW assigned to j^{th} cell in i^{th} MBS. The minimum frequency resource (i.e., 5 MHz in our scenario) is defined as b_m, and b_M is the maximum BW (i.e., 100 MHz in our scenario). Finally, R is an array, the elements r_{ij} therein represent the capacity requirements for j^{th} cell in i^{th} MBS, as read from the traces, or predicted by the ML model (c.f. Section V.C.). Similarly, c_{ij} are the elements of the matrix C, representing the capacity available in j^{th} cell of i^{th} MBS, according to the currently assigned bandwidth, and following Eq 1.

$$C = B * \log_2(1 + \frac{S}{N+I}) bits/s \tag{1}$$

For a given amount of assigned BW, to compute the achievable throughput (C) in bits per second (bps) for a cell, we utilize the Shannon - Hartley theorem (Eq. 1). B is the total allocated BW for a cell (b_{ij}) in Hz, S is the received signal power from the serving cell, which is calculated subtracting the pathloss from the transmitted power and antenna gains, I is the received power from the interfering links, and N is the sum of thermal noise and noise figure. Note that the received signal and the interference are obtained after 1,000 random droppings of one UE within the cell's coverage area. Thus, both S and I refer to average values throughout the cell's coverage area, ensuring the generality of the results.

B. PROPOSED ALGORITHM

Our proposed access-aware BW allocation algorithm is triggered at time intervals. It starts by sorting, in descending order, all the cells according to their requirements, so that the cell serving higher traffic has more chances to be assigned more resources in a BW-limited scenario. Afterwards, for each MBS, and for each cell in that MBS in the BS set, two conditions are checked; i) if the achievable capacity c_{ij} is less than the required capacity r_{ij} and, ii) if the anchoring MBS of that cell still has frequency resources available. Here, to limit the co-channel interference with the adjacent cell, which degrades the network performance due to strong coverage overlapping for the same frequency band, we avoid frequency re-use within a MBS but it allows the assignment of frequency resources up to the system bandwidth. Thus, the cumulative allocated BW of all cells belonging to one MBS (B_i) cannot exceed the system BW (i.e., b_M). It can be reused in the other MBS. If both the aforementioned conditions for a cell are met, the cell is assigned with an additional minimum size frequency resource ($b_m = 5$ MHz, for example) taken from the entire available system BW (e.g., $b_M = 100$ MHz), as explained in Section II.A. Hereafter, allocated BW to that cell and to the anchoring MBS is updated and the achievable capacity c_{ij} is re-calculated. If the new achievable capacity reaches the required capacity r_{ij}, the cell is considered as satisfied and counted as allocated for this time interval. On the other hand, if the new achievable capacity still does not meet the required capacity (let us call it unsatisfied cell), the cell waits for the next iteration, where according to the aforementioned two conditions, only the unsatisfied cells of the MBSs with available BW are assigned additional frequency resources. On the contrary, if the cumulative assigned BW for the MBS, which is anchoring the unsatisfied cell, reaches the maximum allowed BW, the cell is counted as allocated, whether it is satisfied or not. The iteration stops for this time interval when all the cells are considered as allocated.

The discussed BW allocation approach is depicted in Algorithm 1.

IV. EVALUATION SCENARIO

At this point, we want to evaluate the proposed access-aware dynamic BW allocation approach following a real traffic scenario. We use network traces provided by a real network operator, including data collected during seven days with 15 minutes granularity of a LTE based network serving a large number of users in the downtown area of a Greek city.

Algorithm 1 Access-Aware Dynamic BW Allocation

```
Input: numMBS, numMC, nMC, B, R
Initialize B_i ← 0 for all i in (1, numMBS)
Initialize b_ij ← 0 for all i, j
allocated ← 0
while allocated < numMC do
   for all i in (1, numMBS) do
      Sort MC_ij according to their requirements r_ij in decreasing order
      satisfied ← 0
      for all j in (1, nMC_i) do
         if B_i + b_m ≤ b_M and r_ij > c_ij then
            b_ij ← b_ij + b_m; allocate additional b_m to cell MC_ij
            B_i ← B_i + b_m; update total assigned bandwidth to MBS_i
            c_ij ← capacity(MC_ij, b_ij); update MC_ij's capacity according to Eq. 1
            if c_ij ≥ r_ij then
               allocated++
               satisfied++
            end if
         end if
      end for
      if B_i == b_M then
         allocated ← allocated + nMC_i − satisfied
      end if
   end for
end while
```

The network consists of 21 evolved Node B (eNB), comprising 96 cells, with an average of 4.5 cells per eNB (varying from 1 to 9 cells). A simple representation of the considered scenario is depicted in Figure 1.

To observe the behavior of the traffic, we take one cell to start our analysis. Figure 2 depicts the average (i.e., average of seven days converted into a 24 hour scenario) Downlink (DL) data (in Mega Bytes (MB)) for a randomly selected cell varying with the time of the day. It can be observed from Figure 2, that the demand on DL data in the network varies largely with time and thus, a static BW allocation approach designed to meet those varying requirements is not an efficient approach, rather, a dynamic approach is required.

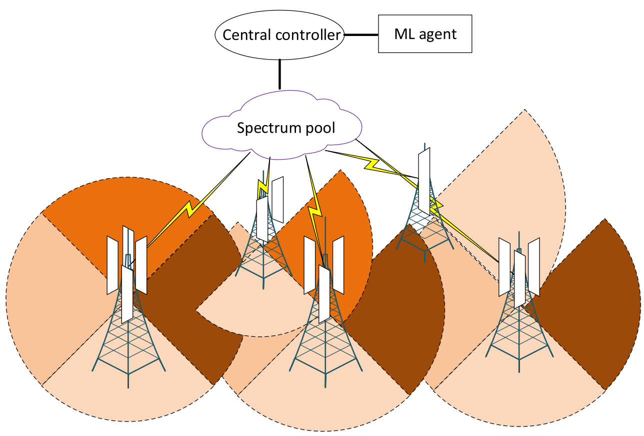

FIGURE 1. Spectrum pool concept to share the available system bandwidth.

A. SIMULATION ASSUMPTIONS

To evaluate our proposed access-aware dynamic BW allocation approach using a real traffic scenario, we reproduce the network scenario, where the traces were collected, into Matlab code. Since our vision is to test our proposal in a 5G-like scenario, we make several assumptions, as recommended by the 5G Infrastructure Public Private Partnership (5GPPP) in [52], yet, assuming the traffic pattern follows the real LTE traces collected.

In [52], 5GPPP claims that in 5G networks, most popular RAT for MBS will be Sub-6 GHz: Carrier frequency (CF) at 3.5 GHz with 100 MHz channel BW. Additionally, 3GPP and ETSI also identify Sub-6 GHz (CF: 3.5 GHz, BW: 100 MHz) as a candidate for 5G New Radio (NR) in [53]. Utilizing Sub-6 GHz operating band, we consider the narrowest transmission BW for a MBS to be 5 MHz, as recommended in [53], which we allocate in every iteration of the proposed BW allocation technique (cf. Section III.B).

As mentioned earlier in Section III, we consider a deployment with no frequency reuse within each MBS; that is, bandwidth assigned to the different cells inside the same MBS does not overlap. Additionally, we consider that directivity of antennas in the different cells and frequency assignment[1] are such that cells from adjacent MBSs do not interfere each

[1]Note that, our mechanism allocates bandwidth, but frequency assignment is out of the scope of this proposal and left for a future work.

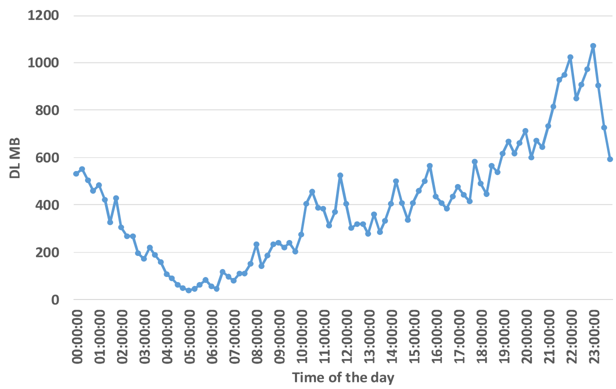

FIGURE 2. Time vs. average DL (in MB) of a randomly selected cell.

TABLE 1. Simulation assumptions following 5GPPP recommendations [52], [54].

Parameters	Value
Number of MBSs	21
Number of Cells	96
RAT	Sub-6 GHz
Carrier Bandwidth (BW)	100 MHz
Carrier frequency	3.5 GHz
BW of a single frequency band	5 MHz
MBS coverage radius	200 - 800m (Depending on the location of the MBS)
MBS transmit power	49dBm in 20 MHz band
Thermal noise level	-174 dBm/Hz
MC and SC noise figure	9 dB in 20 MHz band
Antenna gain	17 dBi
Antenna height	MBS: 25m; User: 1.5m
Average building height	21 m
Channel model	3D model from [54]
Propagation type	MBS to User: LoS; Interferer links: NLoS
Number of interferer links	17
Number of simulations	1000 simulations, each with a random deployment of User for each cell
Training model	Bayesian Regularization (BR)
Number of hidden layers	8
Number of delays (d)	1
Training data set	75% data
Validation data set	15% data
Testing data set	15% data

other significantly, and only one cell per non-serving non-adjacent MBS is considered as a source of interference for the application of Eq. 1. Thus, each cell will be experiencing interference from 17 other cells.

In the following, we describe three different BW allocation approaches to present our evaluation results in the subsequent section.

- *Static Approach 1:* Each cell in a MBS is allocated a non-overlapping 20 MHz channel (the maximum channel BW for LTE [55]), while not exceeding the maximum BW per MBS, which is 100 MHz. If, due to the higher number of cells, 20 MHz per cell is not achievable, the entire available BW is evenly distributed among all the cells in that MBS, which is also a very common approach, where FR = number of cells (usually 3) [16]. For example: If MBS-1 has two cells, each cell will be allocated 20 MHz separated channels, and the remaining 60 MHz will not be assigned in MBS-1. On the other hand, if MBS-2 has 6 cells, each cell will be allocated 16.6 MHz.
- *Static Approach 2:* The available BW is equally split among the cells, not limiting the maximum BW per cell to 20 MHz. In this way, each cell from MBS-1 in the previous example will be allocated 50 MHz. The allocated BW for the cells in MBS-2 remains the same as in static approach 1.
- *Proposed Dynamic Approach:* It follows the operation explained in Section III.B; that is, iteratively, we allocate additional 5 MHz to every cell until its requirements for the next interval are met, or the maximum BW (i.e., 100 MHz per MBS) is reached.

B. PREPARATION OF THE NEURAL NETWORK

According to the discussion presented in Section II.B, given our data set and the objective of the prediction, the most suitable architecture is defined by NARNETs. We focus on a BW distribution problem that must follow the dynamic requirements of a cell, which changes with the time of the day. Hence, we follow a time series data, which are collected from a real network, to understand the dynamic changes in the required throughput.

In both closed loop and open loop scenarios, utilizing past samples (e.g., $y(t-1), y(t-2)\ldots y(t-d)$ of the output, known as delays (d)), and by modeling the underlying characteristics of the time series, NARNET has the ability to self-learn and provide good multi-step predictions ($y'(t)$) of a non-linear time series [50], [56]–[59]. A simple representation of NARNET model is presented in Eq. 2, where function f is the result of the process represented in Figure 3. Each neuron performs a linear combination of all its inputs, applying an adjustable weight to each one, followed by an activation function (i.e., non-linear transformation) f_n.

$$y'(t) = f(y(t-1), y(t-2)\ldots\ldots y(t-d)) \tag{2}$$

To predict the throughput requirements of the access network for each cell, we used NARNET. More precisely, we use DL MB (cf. Figure 2) data extracted from the traces as the input *y(t)* for the NARNET. To predict and compare the results with real traces we used Matlab. While creating a NARNET in Matlab, it allows the selection of the portion of the data set to be utilized during the three phases of learning, i.e., training, validation and testing. During the training phase, a part of the available data is used, and the weight coefficients are adjusted by taking feedback, known as backward propagation, from the comparison between predicted and expected output by means of a given loss function, e.g., Mean Squared

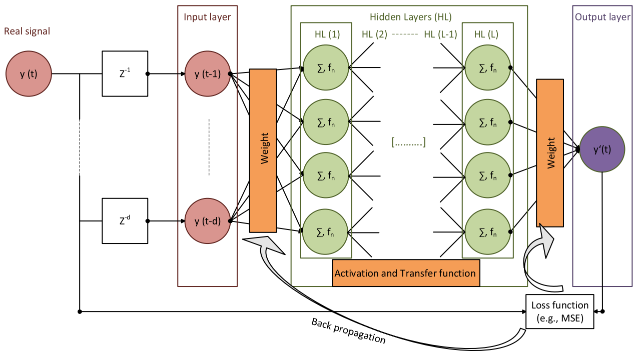

FIGURE 3. Structure of a typical NARNET.

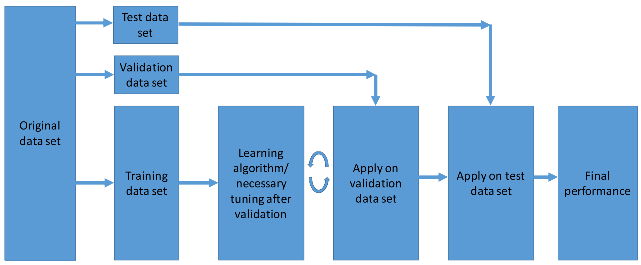

FIGURE 4. Basic representation of training, validation and testing process of NARNET.

Error (MSE) or other estimators. During the validation phase, the training is validated on a different data set, and the results are utilized to generalize (i.e., performance for unseen data) or improve the generalization of the NN. Finally, during the testing phase, the model is tested on the last portion of the data, and the NN performance is computed. A basic representation of the aforementioned process is depicted in Figure 4. Figure 3 presents the structure of a typical NARNET. Here, after trying different combinations, we end up using 70% of the data for training, 15% for validation, and 15% for testing, since that was the partition providing the best performance (i.e., least MSE) for our data set. Three different training algorithms are offered by a Matlab-based NARNET, i) Levenberg-Marquardt (LM), which was originally designed for faster results, however, consumes more memory and is less accurate; ii) Bayesian Regularization (BR), which has the additional objective of minimizing the sum of squared weights and, subsequently, achieves a good generalization of the model [60] (i.e., less prone to overfitting). Compared to LM, BR is a slower model, however, it provides better generalisation for noisy data; and iii) Scaled Conjugate Gradient (SCG), which stops the training when generalization does not improve anymore, consumes less memory, and follows minimization of MSE as the only objective function. A detailed comparative analysis of LM and BR was presented in [60] and the results show that BR outperforms LM for different types of data sets. To find the best configuration for our data set, we performed NARNET training utilizing multiple combinations of following hyper parameters: number of hidden layers (from 1 to 60), delay samples ($d = 1$ to 5) and training models (i.e., LM, BR and SCG). Following different combinations of the aforementioned parameters, we had a large number of predicted data sets, and we measured their corresponding performance by calculating Relative Error (RE) (Eq. 3) for each sample, along with MSE measured by NARNET. Finally, the best combination of the different hyper parameters in NARNET for our data set consists of 8 hidden layers, delay of 1 sample and BR as the training algorithm, which provides an average RE (average of the seven days samples for all the cells in the scenario) of 0.132.

$$RE = \frac{\text{Real data} - \text{Predicted data}}{\text{Real data}} \tag{3}$$

All the assumptions related to the evaluation are summarised in Table 1.

V. RESULTS

In this section we present the obtained results in a comparative manner for different approaches described in Section IV.A. We discuss the results in three different phases; i) peak-hour analysis (i.e., time slot with highest load), ii) 24 hour analysis (average of seven days converted into one day), and iii) machine learning-based traffic prediction and BW allocation for seven days.

Note that, in this work we are interested in testing our proposed dynamic approach in a 5G-like environment, and thus, we scaled up the collected LTE throughput requirements into a 5G scenario, considering that the traffic shows the same behaviour as in LTE. In [61], ITU-T predicted that, compared to recent LTE networks, data rate requirements in 5G will be 10-folded, whereas, according to NTT DOCOMO, the increase ratio will be closer to 100 folds [62]. Therefore, we tried different numbers, i.e., from 10 to 120, as a Scaling Factor (SF), with which we multiply the observed LTE carried traffic to scale it in order to resemble a challenging 5G scenario.

A. PEAK HOUR ANALYSIS

Initially, we select the peak hour (i.e., the time slot when a cell experiences the maximum load during a day) to evaluate our proposed access-aware dynamic BW allocation approach for different values of the SF. In this peak hour scenario, we compare the performance of the proposed approach with two static approaches, as described in Section IV.A.

Analyzing the available data set, we found that most cells experience the highest load of the day around 23:15:00h, as shown in Figure 2. Figure 5 presents the number of unsatisfied cells (i.e., the bandwidth assignment could not meet the actual requirements, as read from the network traces, and scaled by the SF) during the peak hour (23:15:00h) for different levels of SF. Clearly, our proposed dynamic approach outperforms the two traditional static approaches by serving more cells ($\approx$10), while utilizing the same level of system BW (i.e., 100 MHz). At a very high level of SF (i.e., 90 to 120), the number of unsatisfied cells for Static approach 2 and dynamic approach are very close. This is due

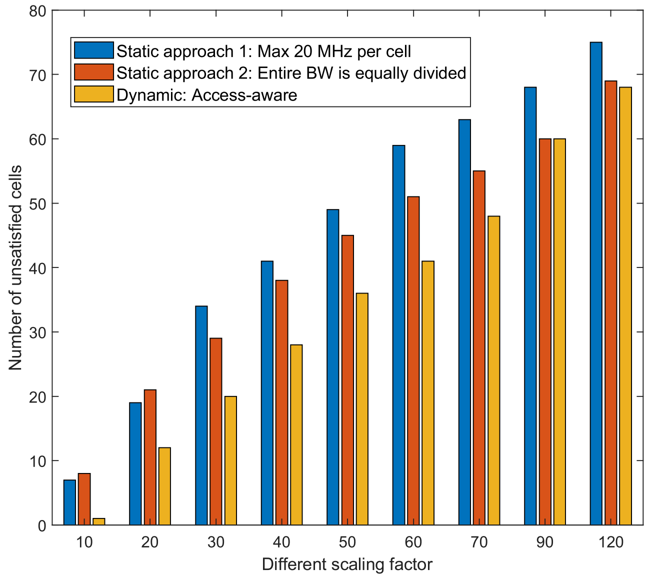

FIGURE 5. Number of unsatisfied cells with different levels of SF.

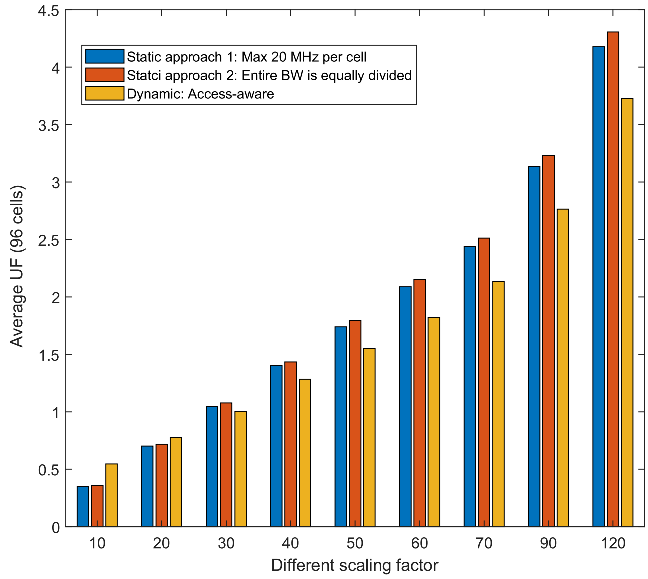

FIGURE 7. Average UF of 96 cells for different levels of SF.

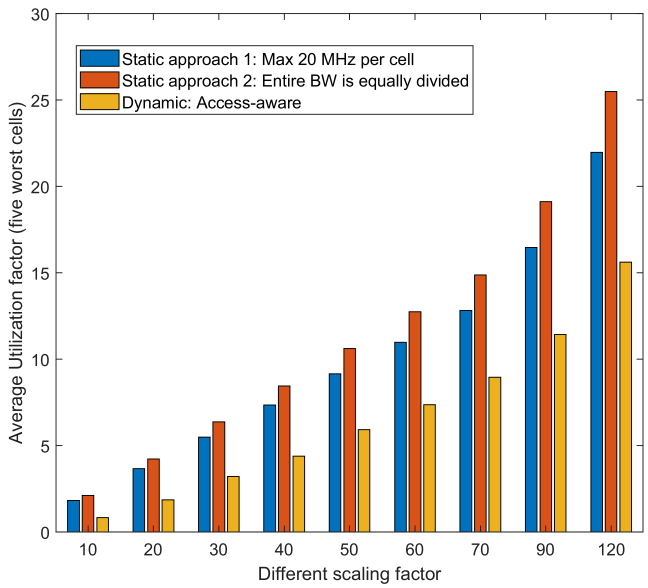

FIGURE 6. Average UF of the worst five cells with maximum load for different levels of SF.

to the very high requirements, which are really difficult to meet by nowadays LTE deployments such as the one used as the basis of our model, designed to support a lighter load. In that extreme, the deployment of a denser network (i.e., more BSs) will be a necessity. On the other hand, in terms of Utilization Factor (UF), which we define as the ratio between required capacity vs. the achieved capacity for a cell, the proposed dynamic access-aware approach performs better for any level of SF. Its performance improves for higher SFs, being it especially efficient for large loads (e.g., SFs 90 and 120). To ensure a fair distribution of BW and to maximize the spectral efficiency, it is better to keep UF value as close to 1 as possible. Values smaller than 1 mean that the network is overprovisioned and frequency resources are wasted, while values larger than 1 imply that capacity requirements are not fully met. Figure 6, depicts the average UF of the five cells with highest load (identified from the previously analyzed unsatisfied cells) for different levels of SFs, which shows a fairer and better distribution in a dynamic access-aware BW allocation approach, compared to the static approaches. The results presented in Figure 6 also show that, utilizing the same level of system BW, the dynamic access-aware approach, is capable of serving more traffic (i.e., UF closer to 1) compared to the static approaches.

Figure 7 shows the UF averaged over the 96 cells. The results present similar trend as in Figure 6, i.e., proposed dynamic access-aware approach performs better (i.e., UF values closer to 1) than the other two static approaches, and for higher values of SF the performance gain becomes more evident as congestion increases (i.e., UF values higher than 1). Additionally, from Figure 7 we can conclude that SF between 20 - 30 consist in the scenarios where we can actually meet the maximum capacity requirements (UF $<=1$) utilizing the available system BW (i.e., 100 MHz per MBS). For larger SF, the studied deployment, designed to support current traffic demands, becomes unable to serve the projected requirements (UF $>=1$). In other words, with a scaling factor lower than 30, the studied network has the potential to carry the scaled traffic and, therefore, remains useful for the study of a future 5G scenario. In contrast, when scaling the traffic x120, a network 4 times denser would be needed (i.e., 4 times the number of deployed MBSs). Therefore, in the subsequent sections we use three levels of SF, i.e., 20, 25 and 30, to present the evaluation results.

B. 24 HOUR ANALYSIS

In this section, we analyze the results in a 24 hour scenario, utilizing the average results of seven days. In Figure 8 we

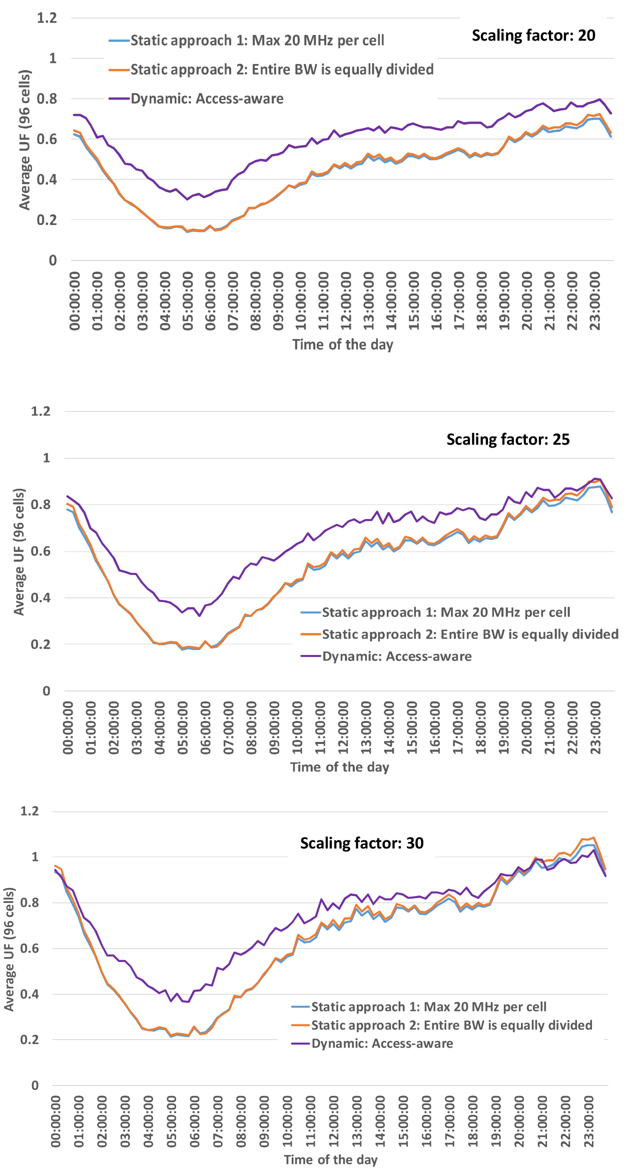

FIGURE 8. Average UF of 96 cells for different approaches.

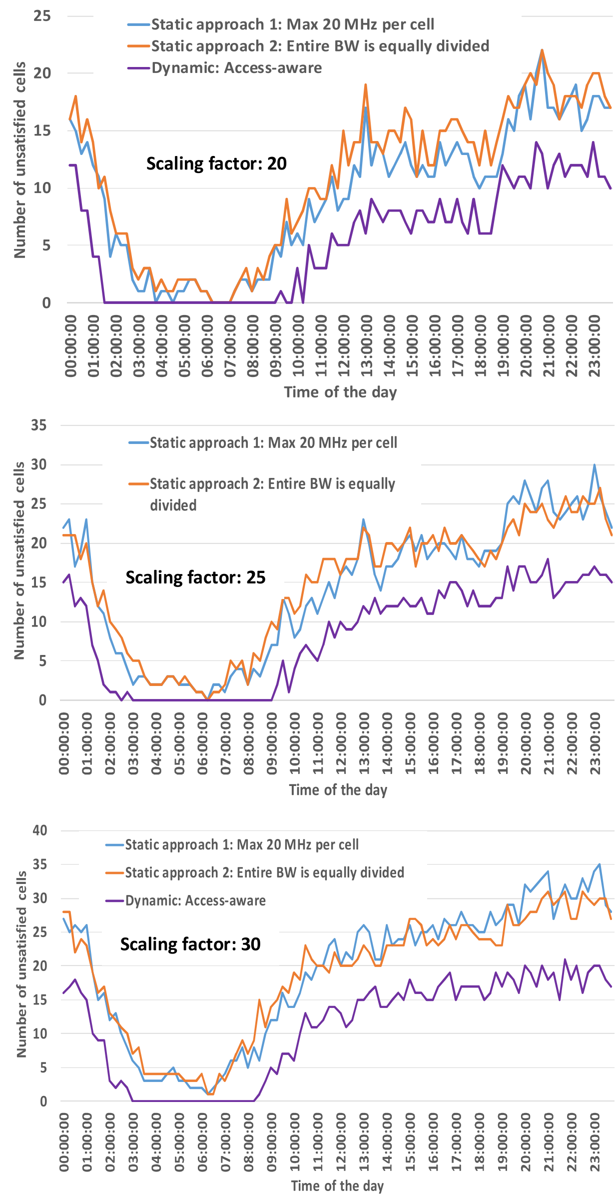

FIGURE 9. Number of unsatisfied cells during the day for different approaches.

present the average UF of 96 cells, varying with the time of the day. As mentioned earlier, UF value closer to value 1 states that available BW is more efficiently used. As depicted in Figure 8 for each SF (i.e., 20, 25 and 30), the proposed dynamic access-aware approach uses less resources than the other two static approaches to serve the same amount of traffic during off-pick hour. On the other hand, during the pick-hour and for higher value of SF (i.e., 30), it is shown that, the dynamic access-aware solution serves more traffic (i.e., UF closer to 1) for a congestion scenario. Therefore, utilizing the same level of system BW, and employing our proposed dynamic distribution, the congestion (i.e., UF > 1) period is 7% shorter (in a 24 hour scenario for SF $= 30$), compared to the Static approach 2.

Figure 9 depicts the number of unsatisfied cells, (i.e., those cells that could not reach the required capacity), during the day for the different approaches. The dynamic access-aware approach tends to satisfy higher number of cells given the BW limitations for a MBS. For instance, for SF 20, during off-peak hour, i.e., low load hour (03:00:00h to 08:15:00h), dynamic access-aware approach is able to serve all the 96 cells without any congestion. However, both the static approaches have a few number of congested cells. On the other hand, during the peak hours (i.e., 20:00:00h to 23:45:00h), dynamic access-aware approach satisfies 9% - 12% more cells compared to the static approaches. With the highest SF of 30, the performance gain of dynamic access-aware approach is even higher (i.e., 10% - 15% more satisfied cells) compared to the two static approaches. Therefore, the proposed dynamic access-aware BW allocation approach

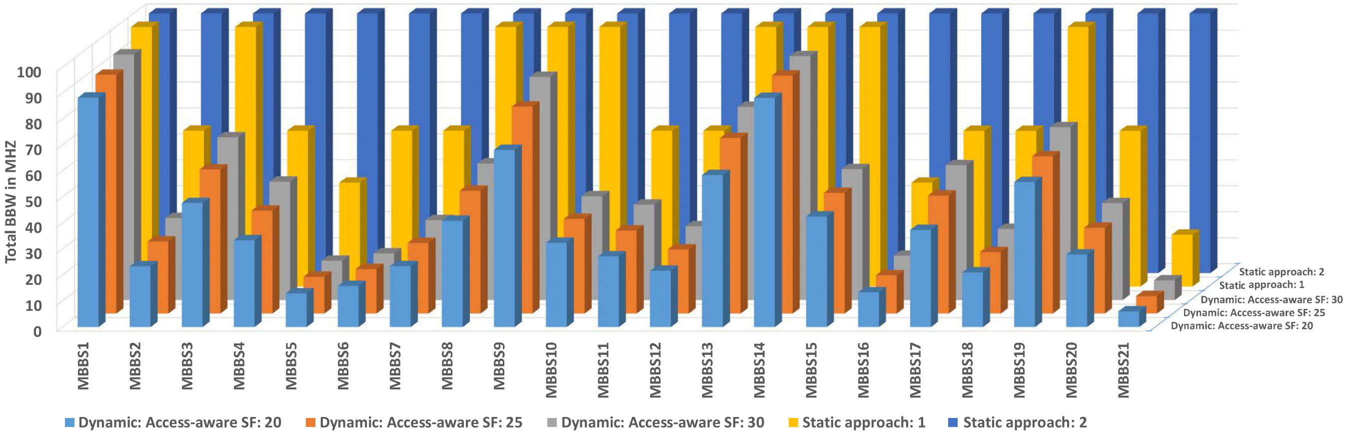

FIGURE 10. Total allocated BW to each MBS for different BW allocation approaches.

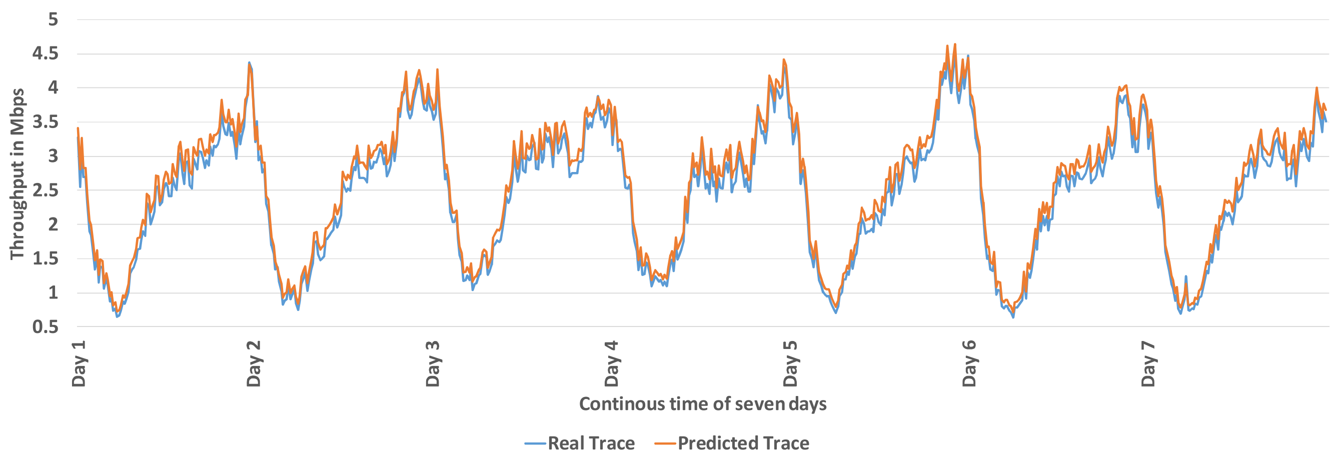

FIGURE 11. Real and predicted average (over 96 cells) throughput requirements for seven days.

appears as a better solution for the congestion problem in the network.

At this point, we also want to observe how much system BW the proposed dynamic access-aware BW allocation approach can save by assigning the necessary amount of BW following the access networks' requirements. In Figure 10, the total BW assigned (average of 24 hours) to each of the studied 21 MBS is depicted for the three different approaches. For Static approach 1, the assigned BW for a MBS depends on the number of cells it has (with a maximum of 20 MHz per cell), remaining agnostic to the access networks' current requirements, and thus, being independent of the SF. On the other hand, for Static approach 2, total assigned BW for each MBS is 100 MHz, independent of the number of cells, or the dynamic requirements of the access network. On the contrary, the total BW assigned for a MBS in the dynamic access-approach is a total reflection of the throughput requirements of the access networks and, to a lesser extent, of the number of cells of a MBS, given that all cells are assigned at least 5 MHz. Moreover, the total assigned BW changes for each cell and MBS, if required, depending on the SF. Thus, in Figure 10, the total BW for a MBS varies for different levels of SF, contrary to the static approaches, where it remains unchanged and agnostic to the current network condition. Additionally, as depicted in Figure 10, for the highest SF studied (i.e., 30), dynamic access-aware approach can save from 5 MHz (i.e., MBS 17) to 60 MHz (i.e., MBS 10) of system BW, compared to Static approach 1, whereas, the number can be between 10 MHz (i.e., MBS 1) to 90 MHz (i.e., MBS 21) compared to Static approach 2.

C. MACHINE LEARNING BASED TRAFFIC PREDICTION AND BW ALLOCATION

In this section we focus on the ML predicted throughput requirements and the variation in the BW allocations based on the results of the predictions.

Thus, this section compares the quality of the BW assignments when the system has perfect prior knowledge of the traffic demands, with the case where the algorithm relies on traffic predictions produced by the neural network described in Section IV.B. In Figure 11, varying with the time of seven days, we present the average (i.e., average of 96 cells) real

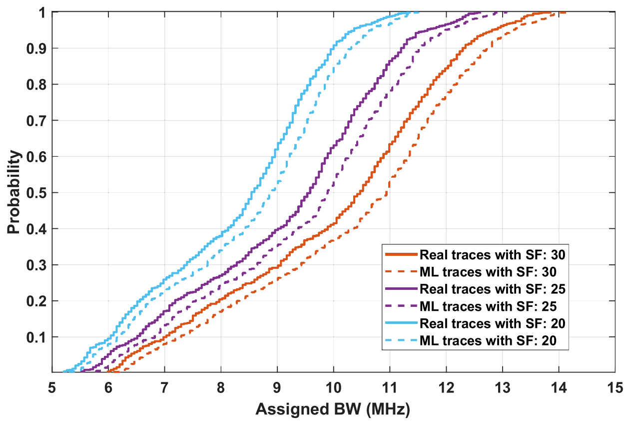

FIGURE 12. CDF of average (over 96 cells) assigned BW for perfect knowledge of real traffic (Real) and ML predicted traffic, using different scaling factors (SF).

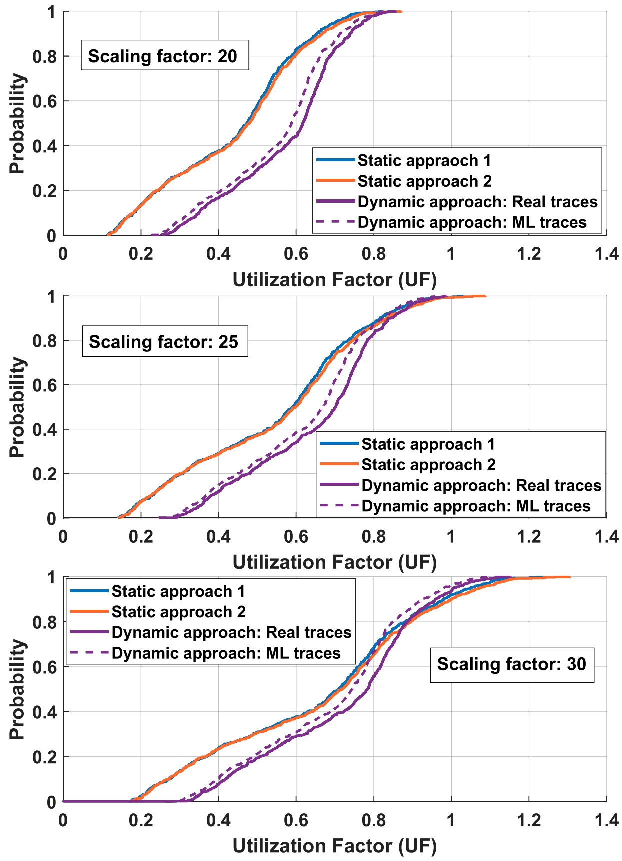

FIGURE 13. CDF of average (over 96 cells) UF for perfect knowledge of real traffic (Real); and ML predicted traffic, using different scaling factors (SF).

traffic requirements as extracted from the traces, and the average requirements predicted by the trained NARNET for the same time instant. As depicted in Figure 11, our selected ML approach provides a prediction showing mean RE $= 0.132$, and follows the behavior of the real trace. However, it is also observed that predicted traces are slightly higher than the real traces, considering the average of 96 cells. This difference has a clear reflection on the BW assignment, because, in our proposed dynamic access-aware BW allocation approach, BW is allocated according to the throughput requirements.

Figure 12 depicts the Cumulative Distribution Function (CDF) plot of the average (over 96 cells) assigned BW (in MHz) to a cell during seven days following real and predicted traces for different levels of SF. As discussed earlier, the predicted traces have usually shown higher values than the real traces, hence, the assigned BW based on those predictions is also slightly higher than strictly needed in practice. For instance, depicted in Figure 12, with SF of 25, a cell is satisfied with 10 MHz or less, on average, 62% of the time, whereas, following ML predictions, this is reduced to 52% of the time. A similar trend follow the results observed for SF values of 20 and 30, where cells are assigned a slightly higher BW when the mechanism is based on predictions instead of having a perfect knowledge of future traffic. In practice, this conservative behavior would be more robust in front of unexpected increments in traffic, at the cost of less adjusted UF.

In Figure 13, we present an analysis of UF for the four BW assignment approaches: i) static approach 1, ii) static approach 2, iii) ideal dynamic approach assuming perfect knowledge of future demands (i.e., following real traces), and iv) dynamic approach following ML predictions. Note that for all the cases, UF is computed with respect to the real requirements, i.e., for iv), UF is the ratio between real throughput requirements vs. the achievable capacity of the network configured based on traffic predictions. Figure 13 presents the CDF plot of average UF (over 96 cells) during seven days (i.e., 672 samples: 7 days with 15 minutes granularity) for different SFs (i.e., 20, 25 and 30). As depicted in Figure 12, due to slightly higher throughput predictions, the predicted BW allocation is also slightly higher than the ideal access-aware dynamic BW allocation approach, and thus, UF (Real traces) in Figure 13 for different levels of SF are slightly better (i.e., closer to value 1) compared to the ML approach. On the other hand, it is also observed that ML-based BW assignment results into better UF (i.e., closer to value 1) compared to the two static approaches during satisfied (UF < 1) and unsatisfied (UF > 1) period.

As shown in Figure 13, for the dynamic approach following the real traces (iii), and the dynamic approach following the predicted traces (iv), UF values remain acceptably close, showing mean RE $= 0.0485$ for SF $= 20$, mean RE $= 0.0463$ for SF $= 25$, and mean RE $= 0.0475$ for high SF $= 30$. Therefore, in terms of utilized resources, the efficiency of a solution based on traffic predictions is similar to the performance obtained if perfect knowledge of future traffic demands were possible. Hence, the predictions from a well trained NARNET, which is performed utilizing a batch of real data set, can be very useful to predict the future requirements and perform an efficient allocation of BW to serve the real traffic in the network. Moreover, as depicted in Figure 5, 6, 7, and 13 for different levels of SF, our proposed dynamic approach always performs better than the static approaches.

As a limitation of this work, we believe that the average UF_Real and UF_ML is not very remarkable (i.e., should be more close to value 1), this is because, during the off-peak hour (03:00:00h to 08:15:00h) the required throughput is usually low compared to the configured capacity utilizing the minimum offered BW (i.e., 5 MHz) considered. This can be an area of improvement for this work to look for more fine-grained channel BW for such future potential mid-band RAT (i.e., Sub-6 GHz). Also note that, limited by the granularity of collected traces, the proposed access-aware dynamic spectrum allocation technique triggers every 15 minutes. We believe, shorter Reporting Output Period (ROP) to collect the traces can make the proposed approach more dynamic and spectral-efficient, without increasing complexity, since the application of the NARNET is computationally cheap, once trained.

VI. CONCLUSION

Expected capacity requirements of future wireless networks continue to challenge the traditional approaches of frequency resource allocation. Moreover, the access network requirements vary dynamically, which requires a dynamic allocation of the frequency resources to ensure the efficient usage of precious resources.

Presented access-aware dynamic BW allocation approach follows the current requirements of each cell and allocates the required BW to serve the carried traffic. The evaluation results show that, with the dynamic access-aware approach, the frequency resources are used more efficiently, i.e., UF closer to 1. Additionally, such intelligent allocation can serve more traffic, specially during the peak hours. Utilizing the same level of system BW, dynamic allocation approach recovers 9% - 12% cells from congestion compared to the traditional static approaches. We have presented the evaluation results by scaling requirements of a real LTE network into a 5G-like scenario, and concluded that, with the higher values of scaling factor, the performance gain of the dynamic access-aware approach are more tangible in terms of number of congested cells.

However, having the perfect network knowledge (i.e., real traces) is a challenge, which cannot be taken for granted. Therefore, in this paper we have presented a NARNET-based ML technique to predict the access network requirements with an acceptable level of error (i.e., RE: 0.132). Presented comparative results show that, in terms of resource utilization and BW allocation, following a NARNET-based predictions, dynamic access-aware BW allocation approach performs very close to the results obtained with perfect network knowledge.

REFERENCES

[1] R. I. Rony, A. Jain, E. Lopez-Aguilera, E. Garcia-Villegas, and I. Demirkol, "Joint access-backhaul perspective on mobility management in 5G networks," in *Proc. IEEE Conf. Standards Commun. Netw. (CSCN)*, Sep. 2017, pp. 115–120.

[2] Ericsson. (Nov. 36, 2020). *Ericsson Mobility Report (Nov. 2020)*. [Online]. Available: https://www.ericsson.com/4adc87/assets/local/mobility-report/documents/2020/november-2020-ericsson-mobility-report.pdf

[3] D. Sesto-Castilla, E. Garcia-Villegas, G. Lyberopoulos, and E. Theodoropoulou, "Use of machine learning for energy efficiency in present and future mobile networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6.

[4] C. V. Nahum, L. De Novoa Martins Pinto, V. B. Tavares, P. Batista, S. Lins, N. Linder, and A. Klautau, "Testbed for 5G connected artificial intelligence on virtualized networks," *IEEE Access*, vol. 8, pp. 223202–223213, 2020.

[5] R. I. Rony, E. Lopez-Aguilera, and E. Garcia-Villegas, "Cooperative spectrum sharing in 5G access and backhaul networks," in *Proc. 14th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2018, pp. 239–246.

[6] N. Palizban, S. Szyszkowicz, and H. Yanikomeroglu, "Automation of millimeter wave network planning for outdoor coverage in dense urban areas using wall-mounted base stations," *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 206–209, Apr. 2017.

[7] G. Salami, O. Durowoju, A. Attar, O. Holland, R. Tafazolli, and H. Aghvami, "A comparison between the centralized and distributed approaches for spectrum management," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 2, pp. 274–290, 2nd Quart., 2010.

[8] F. Shah-Mohammadi and A. Kwasinski, "Fast learning cognitive radios in underlay dynamic spectrum access: Integration of transfer learning into deep reinforcement learning," in *Proc. Wireless Telecommun. Symp. (WTS)*, Apr. 2020, pp. 1–7.

[9] P. Shetkar and S. B. Ronghe, "Spectrum sensing and dynamic spectrum allocation for cognitive radio network," in *Proc. 4th Int. Conf. Converg. Technol. (ICT)*, Oct. 2018, pp. 1–5.

[10] W. Zhang, L. Deng, and Y. C. Kiat, "Dynamic spectrum allocation for heterogeneous cognitive radio network," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2016, pp. 1–6.

[11] H. Zhang, N. Yang, W. Huangfu, K. Long, and V. C. M. Leung, "Power control based on deep reinforcement learning for spectrum sharing," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 4209–4219, Jun. 2020.

[12] P. Lv, M. Fu, Y. Zhuo, H. Zhao, and J. Zhang, "A dynamic spectrum access method based on Q-learning," in *Proc. Int. Workshop Electron. Commun. Artif. Intell. (IWECAI)*, Jun. 2020, pp. 135–141.

[13] K. Sharma, A. Rana, and B. Aneja, "A fuzzy-logic based framework for dynamic channel allocation with improved transmission in cognitive radio," in *Proc. Int. Conf. Signal Process. Commun. (ICSC)*, Dec. 2016, pp. 31–36.

[14] H. Cha and S.-L. Kim, "A reinforcement learning approach to dynamic spectrum access in Internet-of-Things networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[15] F. Bernardo, R. Agusti, J. Pérez-Romero, and O. Sallent, "Intercell interference management in OFDMA networks: A decentralized approach based onreinforcement learning," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 41, no. 6, pp. 968–976, Nov. 2011.

[16] F. Bernardo, R. Agustí, J. Pérez-Romero, and O. Sallent, "An application of reinforcement learning for efficient spectrum usage in next-generation mobile cellular networks," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 40, no. 4, pp. 477–484, Jul. 2010.

[17] B. Bojović, E. Meshkova, N. Baldo, J. Riihijärvi, and M. Petrova, "Machine learning-based dynamic frequency and bandwidth allocation in self-organized LTE dense small cell deployments," *EURASIP J. Wireless Commun. Netw.*, vol. 2016, no. 1, pp. 1–16, Dec. 2016.

[18] Y. A. Abohamra, M. R. Soleymani, and Y. R. Shayan, "Using beamforming for dense frequency reuse in 5G," *IEEE Access*, vol. 7, pp. 9181–9190, 2019.

[19] R. I. Rony, E. Lopez-Aguilera, and E. Garcia-Villegas, "Access-aware backhaul optimization in 5G," in *Proc. 16th ACM Int. Symp. Mobility Manage. Wireless Access (MobiWac)*, 2018, pp. 124–127.

[20] Y. Zhang, D. He, W. He, Y. Xu, Y. Guan, and W. Zhang, "Dynamic spectrum allocation by 5G base station," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun. 2020, pp. 1463–1467.

[21] M. Srinivasan, V. J. Kotagi, and C. S. R. Murthy, "A Q-learning framework for user QoE enhanced self-organizing spectrally efficient network using a novel inter-operator proximal spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2887–2901, Nov. 2016.

[22] G. Zhang, K. Yang, J. Wei, K. Xu, and P. Liu, "Virtual resource allocation for wireless virtualization networks using market equilibrium theory," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2015, pp. 366–371.

[23] T. X. Vu, S. Chatzinotas, V.-D. Nguyen, D. T. Hoang, D. N. Nguyen, M. D. Renzo, and B. Ottersten, "Machine learning-enabled joint antenna selection and precoding design: From offline complexity to online performance," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3710–3722, Jun. 2021.

[24] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for interference management," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5438–5453, Oct. 2018.

[25] H. Huang, W. Xia, J. Xiong, J. Yang, G. Zheng, and X. Zhu, "Unsupervised learning-based fast beamforming design for downlink MIMO," *IEEE Access*, vol. 7, pp. 7599–7605, 2018.

[26] T. E. Bogale, X. Wang, and L. B. Le, "Adaptive channel prediction, beamforming and scheduling design for 5G V2I network: Analytical and machine learning approaches," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5055–5067, May 2020.

[27] F. B. Mismar, B. L. Evans, and A. Alkhateeb, "Deep reinforcement learning for 5G networks: Joint beamforming, power control, and interference coordination," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1581–1592, Mar. 2019.

[28] G. Cerar, H. Yetgin, M. Mohorcic, and C. Fortuna, "Machine learning for wireless link quality estimation: A survey," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 696–728, 2nd Quart., 2021.

[29] J.-R. Jiang, "Short survey on physical layer authentication by machine-learning for 5G-based Internet of Things," in *Proc. 3rd IEEE Int. Conf. Knowl. Innov. Invention (ICKII)*, Aug. 2020, pp. 41–44.

[30] A. Mughees, M. Tahir, M. A. Sheikh, and A. Ahad, "Towards energy efficient 5G networks using machine learning: Taxonomy, research challenges, and future research directions," *IEEE Access*, vol. 8, pp. 187498–187522, 2020.

[31] Y. L. Lee and D. Qin, "A survey on applications of deep reinforcement learning in resource management for 5G heterogeneous networks," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2019, pp. 1856–1862.

[32] M. E. M. Cayamcela and W. Lim, "Artificial intelligence in 5G technology: A survey," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2018, pp. 860–865.

[33] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2392–2431, 4th Quart., 2017.

[34] K. Sharma and X. Wang, "Toward massive machine type communications in ultra-dense cellular IoT networks: Current issues and machine learning-assisted solutions," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 426–471, 1st Quart., 2019.

[35] Y. Xu, W. Xu, F. Yin, J. Lin, and S. Cui, "High-accuracy wireless traffic prediction: A GP-based machine learning approach," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[36] U. Paul, L. Ortiz, S. R. Das, G. Fusco, and M. M. Buddhikot, "Learning probabilistic models of cellular network traffic with applications to resource management," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DYSPAN)*, Apr. 2014, pp. 82–91.

[37] H. Maciejewski, M. Sztukowski, and B. Chowanski, "Traffic profiling in mobile networks using machine learning techniques," in *Proc. Int. Conf. Digit. Inf. Process. Commun.* Berlin, Germany: Springer, 2011, pp. 132–139.

[38] S. Agatonovic-Kustrin and R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *J. Pharmaceutical Biomed. Anal.*, vol. 22, no. 5, pp. 717–727, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0731708599002721

[39] T. Liu, Y. Ye, S. Yin, H. Chen, G. Xu, Y. Lu, and Y. Chen, "Digital predistortion linearization with deep neural networks for 5G power amplifiers," in *Proc. Eur. Microw. Conf. Central Eur. (EuMCE)*, 2019, pp. 216–219.

[40] H. Yang, B. Wang, Q. Yao, A. Yu, and J. Zhang, "Efficient hybrid multi-faults location based on Hopfield neural network in 5G coexisting radio and optical wireless networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 5, no. 4, pp. 1218–1228, Dec. 2019.

[41] R. Khdhir, B. Cousin, K. Mnif, and K. Ben Ali, "Neural network approach for component carrier selection in 4G/5G networks," in *Proc. 5th Int. Conf. Softw. Defined Syst. (SDS)*, Apr. 2018, pp. 112–117.

[42] B. Tian, Q. Zhang, X. Xin, Q. Tian, X. Wu, Y. Tao, Y. Shen, G. Cao, and N. Liu, "Recursive neural network based RRH to BBU resource allocation in 5G fronthaul network," in *Proc. Asia Commun. Photon. Conf. (ACP)*, Oct. 2018, pp. 1–3.

[43] X. Zhao, F. Du, S. Geng, N. Sun, Y. Zhang, Z. Fu, and G. Wang, "Neural network and GBSM based time-varying and stochastic channel modeling for 5G millimeter wave communications," *China Commun.*, vol. 16, no. 6, pp. 80–90, Jun. 2019.

[44] B. Shubyn and T. Maksymyuk, "Intelligent handover management in 5G mobile networks based on recurrent neural networks," in *Proc. 3rd Int. Conf. Adv. Inf. Commun. Technol. (AICT)*, Jul. 2019, pp. 348–351.

[45] A. Mazin, M. Elkourdi, and R. D. Gitlin, "Accelerating beam sweeping in mmWave standalone 5G new radios using recurrent neural networks," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Aug. 2018, pp. 1–4.

[46] A. M. Mahmood, A. Al-Yasiri, and O. Y. Alani, "Cognitive neural network delay predictor for high speed mobility in 5G C-RAN cellular networks," in *Proc. IEEE 5G World Forum (GWF)*, Jul. 2018, pp. 93–98.

[47] Q. Yu, Y. Guan, Y. Yu, and C. Yu, "Genetic algorithm optimized back propagation neural networks in behavioral modeling of power amplifiers excited by 5G NR signal," in *Proc. 9th Asia–Pacific Conf. Antennas Propag. (APCAP)*, Aug. 2020, pp. 1–2.

[48] A. Waqar, S. Khadim, A. Zeb, S. Amir, and I. Khan, "A survey on cognitive radio network using artificial neural network," *Int. J. Future Gener. Commun. Netw.*, vol. 10, no. 11, pp. 11–18, Nov. 2017.

[49] H. D. Trinh, L. Giupponi, and P. Dini, "Mobile traffic prediction from raw data using LSTM networks," in *Proc. IEEE 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2018, pp. 1827–1832.

[50] G. Zhang, H. Zhou, C. Wang, H. Xue, J. Wang, and H. Wan, "Forecasting time series albedo using NARnet based on EEMD decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3544–3557, May 2020.

[51] M. Das and S. K. Ghosh, "Data-driven approaches for meteorological time series prediction: A comparative study of the state-of-the-art computational intelligence techniques," *Pattern Recognit. Lett.*, vol. 105, pp. 155–164, Apr. 2018.

[52] METIS-II. *Performance Evaluation Framework*, document Deliverable D2.1, Jan. 10, 2016. [Online]. Available: https://metis-ii.5g-ppp.eu/wp-content/uploads/deliverables/METIS-II_D2.1_v1.0.pdf

[53] ETSI, *TS: 138 104 V15.2.0-5G; NR; Base Station (BS) Radio Transmission and Reception*, document 3GPP TS 38.104 version 15.2.0 Release 15, Jan. 10, 2018. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/138100_138199/138104/15.02.00_60/ts_138104v150200p.pdf

[54] 5G-PPP. *5G PPP Use Cases and Performance Evaluation Models*. Living document 10, Mar. 2016. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-use-cases-and-performance-evaluation-modeling_v1.0.pdf

[55] O. Liberg, M. Sundberg, Y.-P. E. Wang, J. Bergman, and J. Sachs, "LTE-M," in *Cellular Internet of Things*, O. Liberg, M. Sundberg, Y.-P. E. Wang, J. Bergman, and J. Sachs, Eds. New York, NY, USA: Academic, 2018, ch. 5, pp. 135–197. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128124581000058

[56] P. A. Adedeji, S. Akinlabi, O. Ajayi, and N. Madushele, "Non-linear autoregressive neural network (NARNET) with SSA filtering for a university energy consumption forecast," *Proc. Manuf.*, vol. 33, pp. 176–183, Jan. 2019.

[57] N. Salhab, R. Rahim, R. Langar, and R. Boutaba, "Deep neural networks approach for power head-room predictions in 5G networks and beyond," in *Proc. IFIP Netw. Conf. (Netw.)*, 2020, pp. 579–583.

[58] B. Cha and S.-K. Noh, "Learning using LTE RSRP and NARNET in the same indoor area," in *Proc. 23rd Int. Comput. Sci. Eng. Conf. (ICSEC)*, Oct. 2019, pp. 261–264.

[59] A. He and X. Jin, "NARNET-based prognostics modeling for deteriorating systems under dynamic operating conditions," in *Proc. IEEE 14th Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2018, pp. 1322–1327.

[60] M. Kayri, "Predictive abilities of Bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: A comparative empirical study on social data," *Math. Comput. Appl.*, vol. 21, no. 2, p. 20, 2016.

[61] *IMT Vision—Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond, M Series*, document Rec. ITU-R M.2083-0, 2015, p. 21. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf
[62] N. Docomo, ''Docomo 5G white paper,'' NTT Docomo, Tokyo, Japan, White Paper, Jul. 2014. [Online]. Available: https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/whitepaper_5g/DOCOMO_5G_White_Paper.pdf

RAKIBUL ISLAM RONY (Student Member, IEEE) was born in Bangladesh, in 1986. He received the B.Sc. degree in engineering from the Islamic University of Technology (IUT), Bangladesh, in 2008, and the M.Sc. (Tech.) degree in communications engineering–radio communications from Aalto University, Espoo, Finland, in January, 2014. He is currently pursuing the Ph.D. degree with the Department of Network Engineering, Universitat Politecnica de Catalunya, BarcelonaTech (UPC). At the beginning of his career, he worked as a RF Engineer II at Huawei Technologies Bangladesh Ltd., from 2009 to 2011. During his master's studies, he worked as a Research Assistant at Aalto University to carry out his master's thesis. After completing his M.Sc. degree, he worked as a Senior Lecturer with the Department of Electrical and Electronic Engineering (EEE), Primeasia University, Bangladesh, from 2014 to 2016. From June 2016 to October 2016, he served as an Assistant Professor with the EEE Department, Eastern University. In November 2016, he started working as a Marie Curie Early Stage Researcher (ESR) for H2020 Project 5GAuRA. His research interests include 5G, wireless communications, backhaul, fronthaul, resource management, and cooperative spectrum sharing.

ELENA LOPEZ-AGUILERA received the M.Sc. degree in telecommunications engineering and the Ph.D. degree from the Universitat Politècnica de Catalunya, BarcelonaTech (UPC), in 2001 and 2008, respectively. From 2005 to 2006, she was an Invited Researcher with the Grenoble Computer Science Laboratory. She is currently an Associate Professor and a member with the Wireless Networks Group, Networks Engineering Department, UPC. She has published papers in journals and conferences in the area of wireless communications. She has been involved in projects with public and private funding. Her experience also comprises quality of service, security, radio resource management, location mechanisms, and wake-up radio systems. Her main research interests include the study of IEEE 802.11 WLANs, the Internet of Things enabling technologies in heterogeneous scenarios, and 5G networks.

EDUARD GARCIA-VILLEGAS is currently an Associate Professor with the Universitat Politecnica de Catalunya (UPC) and a member of the Wireless Networks Group (WNG). He participates in the IEEE P802.11 WG and in the research developed within the i2CAT Foundation. After being awarded a Jose Castillejo Grant for young doctors, he collaborates with the MOMENT Laboratory, UCSB. His research interests include IEEE 802.11, radio resource management in wireless networks, aspects of 5G network architectures, and the IoT enabling technologies, such as sensor networks, mesh, and multi-hop *ad-hoc* networks.