Applying Machine Learning and Dynamic Resource Allocation Techniques in Fifth Generation Networks

Christos J. Bouras(✉), Evangelos Michos, and Ioannis Prokopiou

Computer Engineering and Informatics Department, University of Patras, Patras, Greece
bouras@cti.gr, {emichos,st1059554}@ceid.upatras.gr

Abstract. According to Internet of Things (IoT) Analytics, soon, the online devices in IoT networks will range from 25 up to 50 billion. Thus, it is expected that IoT will require more effective and efficient analysis methods than ever before with the use of Machine Learning (ML) powered by Fifth Generation (5G) networks. In this paper, we incorporate the K-means algorithm inside a 5G network infrastructure to better associate devices with Base Stations (BSs). We use multiple datasets consisting of user distribution in our area of focus and propose a Dynamic Resource Allocation (DRA) technique to learn their movement and predict the optimal position, RB usage and optimize their resource allocation. Users can experience significantly higher data rates and extended coverage with minimized interference and in fact, the DRA mechanism can mitigate the need for small cell infrastructure and prove a cost-effective solution, due to the resources transferred within the network.

1 Introduction

Internet of Things (IoT) refers to a system of interconnected devices that possess the ability to communicate (send/receive) data over the same shared network [1]. Based on this architecture, many advanced applications have been created, like smart houses, smart buildings and more. IoT applications are more and more utilized in today's industry and the majority of them focus on long-range communication, while at the same time, they increase the data throughput and minimize power consumption as much as possible. Big Data (BD) derives by IoT sensors and devices and is transferred to servers, which most of the times, is located in Cloud Data Centers worldwide. As a result, the demands for communication and infrastructure keep rising daily. Studies reveal that at least 25 billion devices will be online by 2022 (without including laptops, tablets and smartphones). This alarming increase also comes with an increase on the amount of data that is currently being stored. All this information undoubtedly has to be monitored and analyzed, so that we can keep learning from the available datasets and improve without any manual intervention. Using this technique, IoT devices are becoming smarter and more efficient day by day.

L. Barolli et al. (Eds.): AINA 2022, LNNS 449, pp. 662–673, 2022.
https://doi.org/10.1007/978-3-030-99584-3_57

To face these challenges, many models have been designed that focus on making everything inside an IoT network more Cloud independent. This is where 5th Generation (5G) networks come into play, offering massive connectivity and/or massive machine-type communication (mMTC). Massive access alongside with Machine Learning (ML) aims to achieve effective and secure communications for a large number of distributions to IoT devices via 5G and beyond networks. Massive access key features include low power, massive connectivity, while on the other hand, ML must cope with a massively increased complexity, reducing the number of measurements and facilitating robust decisions, promoting self-organizing networks and future predictions. Those characteristics are constantly proven to be promising for 5G networks [2]. At the same time, Heterogeneous Networks (HetNets) will also come into play by extending the existing macrocell infrastructures. This will be achieved by installing small cells in specific locations inside the macrocell (e.g. in areas near the macro cell borders) so as to provide improved coverage and throughputs for all devices near cell borders, where interference levels significantly experience spikes. At the same time, the use of ML and Artificial Intelligence (AI) is deemed highly necessary for 5G networks, as their application cellular networks is a subject that has recently gained significant interest [3].

Starting off with some of the most popular and efficient ML existing algorithms, the Decision Tree algorithm is a supervised learning algorithm that is mainly utilized in order to efficiently solve the problems of regression and classification, in contrast with other supervised learning algorithms, by classifying the information based on a certain variable [4]. The input variables and output variable correlate with each other through Linear Regression as $y = a + bx$, where y is the output and x is the input. Linear Regression strives towards finding out the coefficients a and b, based on supervised learning [5, 6]. Furthermore, the K-Nearest Neighbors algorithm (KNN) recursively loops through the existing information in order to find the K-nearest instances to the new instance, or on the other hand, the number (denoted as k) of instances that are closer to the new example. The output is either a regression problem or a common class for classification and the aim is to reduce standard deviation at each cluster's points and takes advantage of the Bayes' Theorem in order to calculate how likely is that an event will eventually occur, supposing that another event also occurred [7]. Lastly, the Random Forest algorithm involves numerous decision trees that operate together and simultaneously. Each decision tree reveals a prediction for a class and the most voted class becomes the prediction of the model [8–10]. Last but not least, the K-means algorithm is an unsupervised ML method for the processing of learning data and starts with a first group of randomly selected centroids, which are used for each cluster as starting points, and then performs iterative calculations to optimize the location of the centroid [11–13].

Regarding our motivation, the city of Patras, as well as the majority of the cities of today, has different connectional needs depending on the distribution of the users inside the network. Our goal is to use the knowledge from this user distribution in any given day and suggest the optimal positions for connectivity, as far as the small cells are concerned. When using ML, we observe that K-means often converges to clearly suboptimal local minima depending on the initial conditions possibly not giving the best results. The way we deal with this problem, using the corresponding big dataset, is shown in the following sections of this paper. Furthermore, there exists the UE-BS association issue, which relates to which is the optimal connection between a station and a network device. A data object that deviates greatly from the rest is referred to as an outlier. They signify measurement errors, poor data collection, or simply highlight variables that were not taken into account when collecting the data. They can be the result of a measurement or execution mistake. With the use of DRA also, we seek to minimize the number and effects of the outliers.

Aiming to tackle the aforementioned challenges, in this work we will incorporate the K-means unsupervised learning algorithm in a 5G geographical area, which will help towards optimizing the association problem between a device and a Base Station (BS), assuming their possible positions of the city structure. We will make use of multiple datasets that consist of device spawning (their position and datetime randomly deviate by a small margin) dividing it into 70% of training and 30% of testing dataset and we will then use K-means algorithm, with K equals the number of BS, to learn the user distribution in the network from these datasets and predict future optimal positions of the small cells based on their movement. With the use of ML, we can observe that K-means often converges to clearly suboptimal local minima depending on initial conditions and for that reason, we will be using a large dataset on the users' distribution (position with deviation of some meters) for a relatively representative sample. Additionally, we will also propose a Dynamic Resource Allocation (DRA) technique, where BSs that are low in usage can lend extra resources to neighboring stations to help tackle user congestion. Such an approach can mitigate the need for small cell infrastructure and prove a cost-effective solution, due to the resources transferred within the network. No similar work has been conducted for the specific analysis with the use of K-means.

The rest of this work is organized as follows: In the following section, we showcase our system model on which we are going to examine the association algorithm alongside with the ML mechanisms. In Sect. 3, we demonstrate our proposed mechanisms and Sect. 4 includes the evaluation and comparison and report real experiments with our findings. Section 5 discusses the conclusion and future work.

2 System Model

Starting with the energy consumption model, we consider that in this 5G network model, all BSs do operate at maximum power. This will ultimately result in the highest available throughputs for the network devices. Supposing that each macrocell holds a BS at its center, let $\boldsymbol{P_i^{BS}}$ be the power consumed by the $\boldsymbol{i^{th}}$ BS, which is calculated as [14]:

$$P_i^{BS} = P_i^{cons} \cdot P_i^{rad} + P_i^{BS} \tag{1}$$

where $\boldsymbol{P_i^{rad}}$ corresponds to the outgoing radiated power from the BS, $\boldsymbol{P_i^{cons}}$ is the power consumed because of the feeder/amplifier losses and $\boldsymbol{P_i^{BS}}$ related to the consumed BS-related power.

Regarding the UEs, supposing that $\boldsymbol{P_j^{UE}}$ is the consumed power for the network device, when connected $\boldsymbol{P_j^{UE}}$ is calculated as [15]:

$$P_j^{UE} = P_j^{\text{loss}} \cdot \left(\sum_{\text{a} \in \text{N}_\text{j}^{\text{ant}}} \sum_{\text{i} \in \text{N}_{BS}^{\text{i}}} P_{j,i}^{rad,\text{a}} \right) + P_i^{\text{cons}} \tag{2}$$

where $\boldsymbol{P_j^{\text{loss}}}$ corresponds the power consumed (including system losses) for each of the antennas the device is connected to, $\boldsymbol{N_j^{ant}}$ depicts the different antennas the user is equipped with, $\mathbf{N}_{\boldsymbol{BS}}^{\boldsymbol{i}}$ relates to the set of antennas of a BS, $\boldsymbol{P_{j,i}^{rad,\mathbf{a}}}$ is the radiated power of the $\boldsymbol{a^{th}}$ antenna for the $\boldsymbol{j^{th}}$ user connected to the $\boldsymbol{i^{th}}$ BS and lastly, $\boldsymbol{P_i^{\text{cons}}}$ corresponds to the energy required for the network user to associate with the BS.

Each UE has specific RB demands, depending on the BS it attempts to link/connect to. The device with the lowest number of RBs will attempt to associate to a BS, if the BS has enough RBs to satisfy the device itself. The device's RB demands are proportional to its data rates needs and inversely proportional to the bandwidth of the RB and the Signal-to-Interference-plus-Noise Ratio (SINR) between the UE and the BS. The equation to calculate the required amount of RBs for a device to link to a BS is computed as [16]:

$$r_{j,i} = \left\lceil \frac{th_j}{B_{RB} \cdot \log_2\left(1 + SINR_{j,i}\right)} \right\rceil \tag{3}$$

where $\lceil \bullet \rceil$ corresponds to the operator for the ceiling function, $\boldsymbol{th_j}$ relates to the UE throughput demands, $\boldsymbol{B_{RB}}$ is the RB's bandwidth and $\boldsymbol{SINR_{j,i}}$ denotes the signal quality between the device and the BS.

Regarding the Path Loss (PL) propagation model, in order to measure the signal losses in the simulated network, we construct the distance-dependent path loss model for the macrocell infrastructure (measured in dB) as follows [17]:

$$PL_{macro} = 128.1 + 37.6 \cdot \log_{10}(d) \tag{4}$$

where d corresponds to the distance between the transmitter and the receiver (note that this is measured in kilometers). Consequently, the channel gain can be calculated as:

$$G = 10^{-PL/10} \tag{5}$$

In our simulation, we note that we consider the fact that all BSs have an antenna height equal to 15m, as stated in the 5G NR technical specifications. Any additional wall losses are excluded from our model formulation.

Moving on to the model concerning the user throughputs, let $\boldsymbol{s_{j,i}}$ be subcarrier between the $\boldsymbol{j^{th}}$ UE and the $\boldsymbol{i^{th}}$ BS. Regarding the overall set of subcarriers, we assume that $\boldsymbol{S_{s,j,i}}$ denotes the subcarriers summation between the $\boldsymbol{j^{th}}$ UE and the $\boldsymbol{i^{th}}$ BS. Following the Orthogonal Frequency-Division Multiple Access (OFDMA) standard, the $\boldsymbol{j^{th}}$ UE associated with $\boldsymbol{i^{th}}$ BS has throughput equal to:

$$R_{j,i} = \sum_{s \in S_{s,j,i}^{DL}} B_s \cdot \log_2\left(1 + SINR_{s,j,i}\right) \tag{6}$$

where B_s denotes the subcarrier bandwidth and $\boldsymbol{SINR_{s,j,i}}$ is the SINR between the BS and UE on a subcarrier s. BLER is equal to 10^{-4}. The $\boldsymbol{SINR_{s,j,i}}$ is formulated as follows (all calculations are over a subcarrier s) [14]:

$$SINR_{s,j,i} = \frac{P_{s,i}^{rad} \cdot G_{s,j,i}}{N_0 \cdot \Delta f + \sum_{i'} P_{s,i'}^{rad} \cdot G_{s,j,i'}} \tag{7}$$

where $\boldsymbol{P_{s,i}^{rad}}$ denotes the radiated power from the BS, $\boldsymbol{G_{s,j,i}}$ corresponds to the channel gain between a $\boldsymbol{j^{th}}$ UE and an $\boldsymbol{i^{th}}$ BS, $\boldsymbol{N_0}$ is the white noise power spectral density, $\boldsymbol{\Delta f}$ is the subcarrier spacing and $\sum_{i'} \boldsymbol{P_{s,i'}^{rad}} \cdot \boldsymbol{G_{s,j,i'}}$ relates to the summation of every $\boldsymbol{i'^{th}}$ BS's radiated power (which causes interference in the neighboring cells), multiplied with the channel gain between the interfering BS and the UE. Finally, in order to calculate SINR in (dB), we use the following equation:

$$SINR_{(dB)} = 10 \cdot \log_{10}\left(SINR_{s,j,i}\right) \tag{8}$$

3 Proposed Mechanisms

The K-means algorithm is an unsupervised ML technique for the processing of learning data and begins with a first group of randomly chosen centroids, which are used as the starting points for each cluster and then performs iterative calculations to optimize the centroid positions. K-means is chosen through other clustering variances because of its scalability and adaptability in large datasets as well as its guaranteed coverage. By alternating between assigning data points to clusters based on current centroids, K-means finds the best centroids selecting centroids (the center points of a cluster) based on the current assignment of data points to clusters. This attempts to make the data points of the intra-cluster as close as possible while at the same time, keeping the clusters as distinct as possible. The clustering generated is a form of vector quantization that aims to divide n observations into k clusters in which each observation belongs to non-overlapping subgroups (clusters) in which each data point belongs to only one group with the nearest mean (cluster centers or centroids), serving as the cluster prototype. This results in the data space being partitioned into Voronoi cells. K-means clustering minimizes variances within clusters using squared Euclidean distances, but not normal Euclidean distances, whereas only the geometric median minimizes Euclidean distances (the mean optimizes squared errors).

The proposed UE-BS association algorithm assumes pre-defined context information for users. Aiming at maximizing the efficiency of the proposed model using ML while respecting the pre-defined user data demands, the aforementioned problem transforms into a minimization of required RBs. The proposed low-complexity UE-BS association algorithm requires knowledge of the SINR, the system architecture, the available RBs, the throughput demands for every user and the outcome of the K-means algorithm. To achieve efficiency maximization, we begin iterating from the device with the lowest RB requirements. Repetitively, for each device, we will attempt to associate the device which has the lowest demands towards the BS to which it has maximum signal quality, or in more technical term, maximum SINR. If ML is enabled in the current simulation scenario, then we attempt to take advantage of the distribution prediction K-means produced with K equals the number of BS in the scenario using the resulting centroids as the optimal positions of BS including the corresponding users that we suggest, otherwise we continue without ML. Additionally, the best-case scenario is when both ML and DRA are enabled, offering additional RBs to BSs that are in need because of multiple reasons (device congestion, high interference, low coverage etc.). A DRA connection will always be optimal for the device, because the UE-BS association will be optimal in terms of signal coverage. Each UE-BS association is possible only if there exist remaining RBs, otherwise, we decide to select the next best candidate. As for any remaining BSs, they are discarded in this scenario.

Mechanism 1. K-means Algorithm using input datasets

```
begin
 specify number of clusters K and initialize k means
points randomly
 guess some initial cluster centers
 calculate the distance between each data point and
cluster centers
 for every μ_i = some value, i=1,...,k, do
      categorize each item to its closest mean
      update the mean's coordinates (averages of the
items categorized in that mean so far)
      assign points to nearest cluster center
c_i={j:d(xj,μi)≤d(xj,μl),l≠i,j=1,...,n}
      set the cluster centers to the mean
μ_i=1|c_i|∑jcEixj for every I
  end for
 keep repeating until there is no change to the centroids
 If no data point was reassigned then
      Stop
 end .
```

Mechanism 2. Association Algorithm using ML and DRA support

```
begin
 K = number of BSs
 for each j in N_UE do
      choose device candidate with min (T_device,BS )
      select best BS by finding max(SINR_device,BS )
      if ML is enabled then
           for each BS do
              for each device do
                   clusters[BS_l] = K-means of device on BS_l
              end for
          end for
      end if
      if the available RB_BS are enough then
         if ML is enabled then
            associate device and BS based on the clusters
        else
               associate device and BS
        end if
         if ML DRA is enabled
            for each BS do
                 if any BS needs resources then
                       if neighboring BS has enough RBs
to serve then
                       offer 15% RBs
                       update remaining RBs
                       end if
                  end if
             end for
         else
            update available RBs;
         end if
      else
           select next best BS candidate by max ( SINR_device,BS )
      end if
 end for
 end .
```

4 Performance Evaluation

In this section, we discuss the 5G network simulation scenario, where the Python programming language was used to construct the experiment analysis (datasets, K-means implementation, system model, association algorithms etc.). We consider a two-level ring topology in the geographical area of study, resulting in a total of 19 macrocells. Considering the second level of surrounding macrocells is crucial towards measuring the interference caused by neighboring cells, as it would be a mistake not to consider the negative effects of signal interference from the neighboring cells. All macro BSs are located in the center of the cell and are surrounded by small cell infrastructures that can help towards better user coverage inside the network. We consider that all BSs operate at full power, to provide the highest available throughputs to the devices inside the network.

The geographical area of interest is depicted below in Fig. 1. The map depicted below represents a larger area in the city of Patras and includes all available BSs (macrocell and smallcell infrastructures, as well as the devices distribution alongside the area of interest). To evaluate the ability of our mechanisms to efficiently predict user movement inside the network, a dataset was created based on the devices' location (with deviation of some meters) enough times for the accuracy to be objective and representative, including the (x,y) pair of the devices' positions and their timestamp, for a fixed number of hours per day. The experiment ended after gathering enough information from a whole month, which was entered in the dataset, which was then given as input to the K-means algorithm. The ultimate goal of the approach is to manage to offer better user coverage and data rates, after successfully predicting the changes in user demands each day, based on the user distribution as gathered in the dataset. This means that for the case of predicting resources that e.g. higher probability of devices being in the city center from Monday through Friday (weekdays) and them being out of the urban area on the weekends, the prediction would suggest allocating extra resources to the more crowded areas (due to the extra bandwidth available). The probability distribution is based on real-life scenarios (Fig. 2 and Table 1).

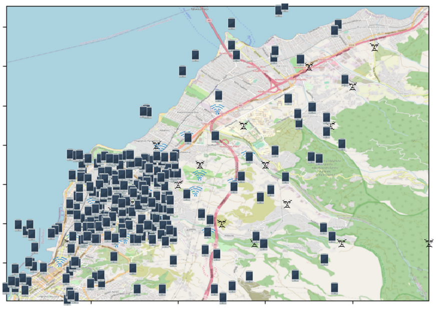

Fig. 1. Snapshot of simulation scenario in patras

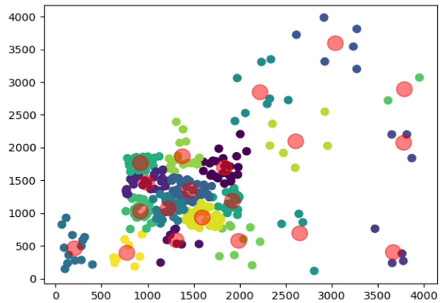

Fig. 2. Optimal smallcell positions using user movement

Figure 3 and Fig. 4 reveal the successful connections percentage, divided into macrocell and smallcell connections, for both the scenarios of weekly and weekend device congestion. The very first thing we can easily notice is that as devices increase inside the same geographical area, it is getting harder and harder for macrocells to serve all devices in the network. This is due to the fact that all macrocell BSs have a pre-defined amount of RBs devoted to them and as devices increase in the network, the more chances there are that the RBs will diminish with a higher rate. This means that such devices can attempt to connect to the additional surrounding layer of small cells, as envisioned officially in 5G networks, where due to multiple factors (e.g. no RBs remaining, outside of area coverage, high interference from neighboring BSs), it is preferable to connect to low-emission and low energy consumption smallcells to be successfully covered in the 5G network.

Additionally, we observe that the applied ML techniques, with or without DRA, can offer an increased amount of macrocells resulting in less smallcell connections because

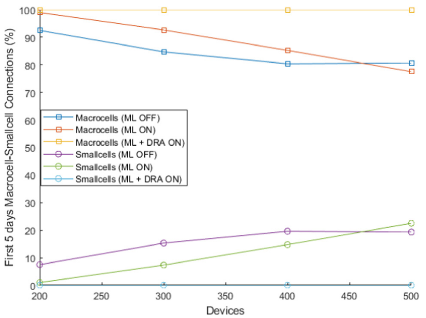

Fig. 3. Weekdays Macrocell/Smallcell Connections (%)

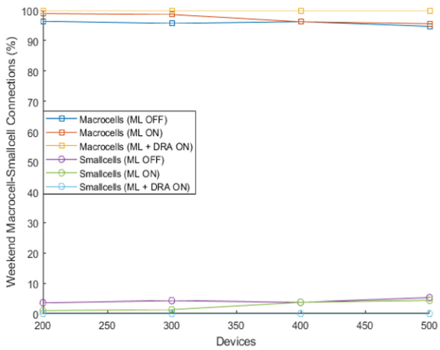

Fig. 4. Weekend Macrocell/Smallcell Connections (%)

Table 1. Experiment parameters

Parameter	Setting	Parameter	Setting
Macrocells	19	White noise density	−174 dBm/Hz
Air Protocol	5G NR	Macrocell Coverage	375 m
5G Frequency Range	FR1	Small cell Coverage	50 m
RB Bandwidth	360 kHz	BS Antenna Gain	15 dBi
Modulation Scheme	64QAM	UE Antenna Gain	0 dBi
Bandwidth	100 MHz	Macro BS $P^{rad}_{i,max}$	40 W
Carrier Frequency	3.5 GHz	Small BS $P^{rad}_{i,max}$	1 W
RBs	273	UE $P^{rad}_{j,max}$	0.2 W
Subcarrier spacing	30 kHz		

the proposed ML techniques gain knowledge from the input dataset from multiple previous instances of the geographical area gathered. Using the K-means (Mechanism 1), they suggest the optimal connection based on the existing network. When the DRA comes into play, we can see that the connections to macrocells maximize, whereas the connections to smallcells minimize. This is because the DRA mechanism takes advantage of the existing ML dataset and can accurately relocate resources from relatively empty BSs to BSs that need them the most due to device congestion (see lines 19–30 in Mechanism 2). This is an important achievement, because with this ML technique, we can mitigate the need for acquiring and installing smallcell infrastructures by relying on existing knowledge of the network's datasets.

In the figures above (Fig. 5 and Fig. 6), we observe the usage of the overall RBs available to the network for the three different simulation scenarios (without ML, with ML, and with ML combined with DRA). Studying the weekly and the weekend device

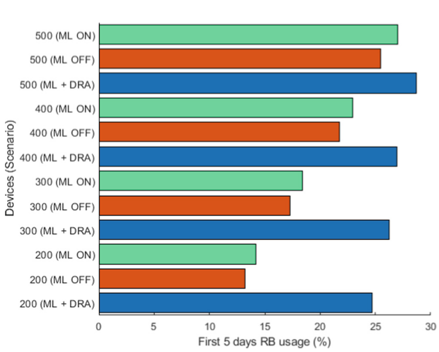

Fig. 5. Weekdays overall RB usage (%)

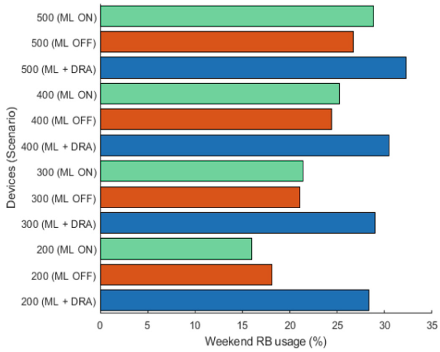

Fig. 6. Weekend overall RB usage (%)

congestion, we conclude that as the number of devices augments, more RBs are needed, since the UE-BS association algorithm relies on the device's RB demands (see (3) and lines 13–33 in Mechanism 2). The more devices in our network, the more resources are needed from the BSs. Since all BSs have a pre-defined number of RBs, the RB usage increases proportionally to the device amount and according the ML techniques applied, the RB usage augments. When DRA is applied, more RBs are being consumed, since this mechanism tries to associate the current device with the best BS available according to the ML output. If this option is not possible, resources will be relocated inside the network infrastructure for the optimal association to be completed successfully (see lines 19–30 in Mechanism 2). Thus, more and more RBs of the macrocells are needed, despite the existence of a small cell infrastructure inside a network, which leads to the ML technique with the DRA being a very cost-effective and efficient solution of the UE-BS association problem inside 5G networks.

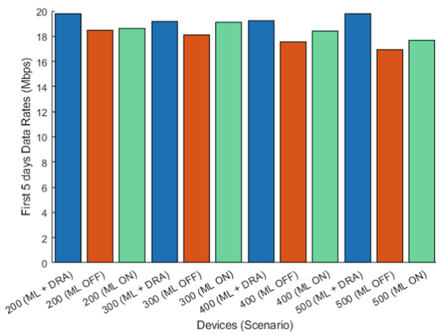

Fig. 7. Weekdays average data rates (Mbps)

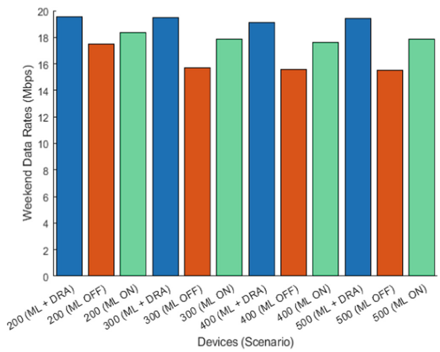

Fig. 8. Weekend average data rates (Mbps)

Figures 7 and 8 show the weekly and weekend average data rates for the connected devices in the network. We conclude that: a) more devices connected lead to lower average data rates, regardless of whether ML techniques were used or not and b) upon applying DRA, the data rates maximize, compared to the previous two scenarios. The first conclusion can be easily justified since more devices connect to the BSs with a predefined amount of RBs allocated to them causing the resources to eventually diminish at a higher rate. As a result, the number of unconnected devices is increased, due to insufficient BS resources. Thus, there must be a compromise at the throughput demands coming from all the devices so that a larger amount of them can be served. Regarding the second conclusion, through DRA, all devices can efficiently be served while scanning the ML dataset ensures that all devices are in fact served by the optimal BS in their area of coverage, providing better signal quality (so, better SINR). Since their optimal BS will be able to cover their RB demands, according to (6), the higher the RB demands and the SINR signal quality, the higher the data rates eventually will end up being (see lines 19–30 in Mechanism 2).

5 Conclusion and Future Work

Studies show that in the future, the number of devices connected in IoT networks will range from 25 up to 50 billion. As a result, IoT infrastructures will require more effective and efficient analysis methods than ever and ML techniques are envisioned to be the solution, accompanied by the coming of 5G networks. In this work, we incorporated the K-means unsupervised learning algorithm inside a 5G network infrastructure to better associate devices with BSs. We used multiple datasets consisting of user distribution (datetime included) in Patras and used K-means to learn the user movement and predict the optimal position for the connection station. Additionally, we proposed a DRA technique, where BSs that are low in usage can lend extra resources to neighboring stations to help tackle user congestion. Simulations revealed that by applying such ML mechanisms inside 5G infrastructures, users can experience significantly higher data rates and extended coverage with minimized interference. The DRA mechanism can mitigate the need for small cell infrastructure and prove a cost-effective solution, due to the resources transferred.

References

1. Rocha Neto, A., et al.: Classifying smart IoT devices for running machine learning algorithms. In: Anais do XLV Seminário Integrado de Software e Hardware, Natal (2018)
2. Polese, M., et al.: Machine Learning at the Edge: A Data-Driven Architecture with Applications to 5G Cellular Networks, Institute of Electrical and Electronics Engineers (IEEE) (2020)
3. Zantalis, F., Koulouras, G., Karabetsos, S., Kandris, D.: A review of machine learning and iot in smart transportation. Future Internet **11**(4), 94 (2019)
4. Dietterich, T.G., Bae, E.: Machine Learning Bias, Statistical Bias, and Statistical Variance of Decision Tree Algorithms, Oregon State University (2008)
5. L'Heureux, A., Grolinger, K., Elyamany, H.F., Capretz, M.A.M.: Machine learning with big data: challenges and approaches. In: Institute of Electrical and Electronics Engineers. (IEEE) Access, vol. 5, pp. 7776–7797 (2017)
6. Ray, S.: A quick review of machine learning algorithms. In: International Conference on Machine Learning, Big Data, Cloud and Parallel Computing. (COMITCon), Faridabad, India, pp. 35–39 (2019)

7. Zhang, H.: The optimality of naive bayes. In: Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference, FLAIRS (2004)
8. Breiman, L.: Random forests. Mach. Learn. **45**, 5–32 (2001)
9. Breiman, L.: Arcing classifiers. Ann. Stat. **26**(3), 801–849 (1998)
10. Freund, Υ., Schapire, R.E.: Experiments with a new boosting algorithm. In: Machine Learning: Proceedings of the Thirteenth International Conference, pp. 148–156 (1996)
11. Balevi, E., Gitlin, R.D.: A clustering algorithm that maximizes throughput in 5G heterogeneous F-RAN networks. In: IEEE International Conference on Communications (ICC), Kansas City, MO, pp. 1–6 (2018)
12. Cabrera, E., Vesilo, R.: An enhanced k-means clustering algorithm with non-orthogonal multiple access (NOMA) for MMC networks. In: 28th International Telecommunication Networks and Applications Conference (ITNAC), Sydney, Australia, pp. 1–8 (2018)
13. Kumar, S., et al.: Handover forecasting in 5G using machine learning. Int. J. Eng. Technol. (UAE) **7**(31), 76–79 (2018)
14. Richter, F., Fehske, A., Fetwweis, G.: Energy efficiency aspects of base station deployment strategies for cellular networks. In: Institute of Electrical and Electronics Engineers (IEEE) 70th Vehicular Technology Conference Fall, Anchorage, AK, (2009) pp. 1–5 (2009)
15. Ghaleb, A.M., Mansoor, A.M., Ahmad, R.: An energy-efficient user-centric approach high-capacity 5G heterogeneous cellular networks. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **9**(1), 405–411 (2018)
16. Holma, H., Toskala, A., (eds.): LTE for UMTS: OFDMA and SC-FDMA Based Radio Access. John Wiley & Sons, pp. 213–257 (2009)
17. 3rd Generation Partnership Project (3GPP). 3GPP TS 36.931, v13.0.0, Radio Frequency (RF) requirements for LTE Pico Node B, (Release 13) (2016)