Data Analytics in the 5G Radio Access Network and its Applicability to Fixed Wireless Access

O. Sallent, J. Pérez-Romero, R. Ferrús, R. Agustí
Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
[sallent, jorperez,ferrus,ramon]@tsc.upc.edu

***Abstract*—This paper discusses the exploitation of data analytics for supporting the operation of Radio Resource Management (RRM) in the Next Generation Radio Access Network (NG-RAN), analysing the need for standardised and open frameworks and the specific characteristics of the NG-RAN that impact on the design of data analytics solutions. As a practical example of this concept, a simple but illustrative use case is presented, which intends to enhance the Radio Admission Control (RAC) function with data analytics information in scenarios with both Fixed Wireless Access (FWA) users and mobile users. A functional framework for the realization of the proposed solution is described and the performance of the RAC algorithm is evaluated under different scenario conditions to better delimitate its potential.**

Keywords—5G; Data Analytics; Radio Admission Control; Fixed Wireless Access

I. Introduction

Along with supporting new business opportunities, data analytics –understood as the pursuit of extracting meaning from raw data– will be key to getting 5G networks efficiently rolled out and operated. Tightly integrated with Artificial Intelligence (AI) technologies, the challenge is to efficiently handle the big amount of data originated within a mobile network and turn it into value by gaining insight and understanding data structures and relationships, extracting exploitable knowledge and deriving successful decision-making. The long term vision, expected to fertilize in the 5G ecosystem, is to enable largely autonomous networks which are able to self-configure, self-monitor, self-heal and self-optimize with minimum human intervention [1].

Data analytics are a major differentiator allowing operators to leverage knowledge from data. Operators need to see a 360-degree view of customers and networks. Data analytics suite brings to operators intelligent and self-learning software that continuously optimizes towards its goals, embracing a wide array of analytics to use standalone or combined. For example, customer data analytics allows operators to increase customer satisfaction to gain and maintain a loyal subscriber base, while network data analytics enables deep insight in the functioning of the network and connected devices that can be exploited for agile network corrective actions and optimization. In this respect, and specifically for the Next Generation Radio Access Network (NG-RAN), functions such as Radio Network Planning (RNP) and Self-Organizing Networks (SON) are likely to undergo significant transformation though the introduction of data analytics and closed-loop automation. For instance, the RNP process will benefit from the capability to process live network measurements and extract data features and models (e.g. network spatial-temporal traffic patterns along with the corresponding prediction models, clustering models for identifying cells with high interaction) to be integrated in the planning process loop. Similarly, complex SON use cases gaining more relevance in 5G such as massive MIMO parameters configuration can benefit from Machine Learning (ML) algorithms trained with big data analytics. In addition to the exploitation of AI/ML and data analytics in the management and operation of the NG-RAN, Radio Resource Management (RRM) algorithms embedded within the NG-RAN nodes are also within the scope of this transformation. In this regard, given that the new features introduced in the 5G New Radio (NR) interface enables for a much deeper time/frequency/space granularity compared to LTE, RRM complexity grows considerably and new approaches become necessary. For example, scheduling for 5G massive MIMO networks by conventional optimization algorithms may become impractical, so ML with massive offline training can be useful to approximate optimal solutions [2].

The potential high impact of data analytics in 5G networks has raised a lot of research interest as reflected by different papers. Some of them, such as [3][4][5], establish general frameworks for the application of data analytics in 5G. Focusing on the RAN, there are different papers that have proposed and assessed the performance of analytics-based techniques and algorithms. A significant number of them are related to self-healing, covering aspects such as anomaly detection [6], [7], diagnosis [8] or alarm prediction [9]. Some other works have dealt with different aspects of self-optimization, such as tilt and power adjustment based on identification of highly interfering cells [10], interference management [11], self-configuration of neighbour cell lists [12] and coordination of SON functions [13]. With respect to the considered methodologies, most of the abovementioned works apply ML tools to process the data, such as K-means clustering [6][11], random forest prediction [9], Self-Organizing Maps [8] or regression models [10][13]. Above works have also considered different types of data sources, such as call detail records [6][7], mobile traces [8], counters [9][11] or User Equipment (UE) measurement reports [10][13]. The state-of-the-art analysis reveals that the main focus has been placed on the design and evaluation of specific solutions for on-top-of-the-RAN operations (i.e. management and optimization processes). However, it is worth noting that little attention has been placed so far in the exploitation of data analytics for supporting the RRM operation within the NG-RAN nodes.

In this context, this paper develops a simple but illustrative use case considering a 5G MNO (Mobile Network Operator) that provides mobile and Fixed Wireless Access (FWA) services and intends to exploit the knowledge about the FWA users and data measurements collected from their service usage in order to perform smarter Radio Admission Control (RAC) decisions for guaranteed bit rate services. The paper assesses the performance of this RRM function and discusses its practical realization in terms of network architecture aspects and procedures.

The rest of the paper is organised as follows. Section II provides a general discussion on the suitability to gather measurements and implement data analytics in the 5G RAN. Section III presents the considered use case that involves RAC for scenarios with FWA users. The detailed description of the solution is presented in Section IV, followed by a performance evaluation in Section V. Section VI concludes the paper.

II. Data Analytics in NG-RAN

A. *Need for standardised and open data analytics frameworks*

In order to spur innovation and facilitate adoption of data analytics solutions, standardised frameworks are necessary. In this direction, the latest 3GPP 5G specifications (Release 15) have already delineated a network function (NF) called NetWork Data Analytics Function (NWDAF) [14] intended to provide data analytics information services to other NFs within the 5G Core (5GC). In this regard, a number of use cases involving the operation of the NWDAF in the context of network automation have been analysed in [15] and normative work is on-going to specify the necessary data to expose to NWDAF and the necessary NWDAF outputs (e.g. statistics, predictions) in order to support the use cases. The approach to be followed in the NG-RAN concerning the support of data analytics services is at a much more incipient stage, being one of the open points under consideration the need and feasibility of introducing a logical entity/function for RAN centric data collection and utilization [16]. Leveraging and complementing the 3GPP work, another initiative currently working on defining open interfaces at RAN level is Open RAN Alliance, which has elaborated a high-level reference architecture [17] with support for, among others, RAN data analytics.

B. *Specifics of data analytics in the RAN*

The inherent peculiarities of the wireless mobile environment have to be accounted in the design of data analytics and AI-based mechanisms. Specifically, the intrinsic randomness associated to e.g. propagation and traffic conditions or the dynamicity due to mobility and interference conditions require careful consideration when collecting, processing and exploiting measurements and statistics derived from them.

One of the fundamental assumptions behind many supervised ML algorithms is that the characteristics of the data used to design a system will remain the same once the system is deployed. However, this important assumption is often violated in practice, for example, because of the non-stationarity of the environment, and as a result, a system's accuracy may suffer significantly. Specifically, supervised learning such as regression and classification learn an input-output dependency from input-output paired training samples so that test output for unseen test input can be accurately estimated (i.e., learning machines can generalize to unseen test data from training data). If the training and test data do not follow the same probability distribution, standard supervised learning algorithms suffer significant estimation bias. Consequently, data analytics frameworks in the RAN should include mechanisms to ensure robustness against non-stationarity. In this respect, the "reliability tester" concept presented in [18] could be a possible solution.

Another important issue to consider in ML is the problem of imbalanced learning [19], in which an algorithm has to learn from datasets where some situations may be extremely underrepresented. This can be particularly critical in the RAN when trying to detect situations that rarely occur in the real network, e.g. anomalous cell behaviors, anomalous traffic levels, faults, etc. and therefore they can hardly be represented in the datasets used to train a ML algorithm for detecting those situations. To handle this problem, ML algorithms need to incorporate mechanisms to efficiently transform vast amounts of raw data into useful information and representative knowledge.

C. *Value/complexity trade-off*

Huge expectations are placed in AI-empowered RANs. To what extent data analytics tightly integrated with AI flourish in the future 5G ecosystem depends on how well the value/complexity trade-off is resolved. The value has to do with the improvement achieved by the considered technique/algorithm in specific use cases. The complexity has to do with its practical implementation, balancing aspects such as design feasibility (e.g., real-time operation requirements, existence of open standards) and required resources (e.g. computing, storage). In this respect, the authors believe that the identification of simple (though sound) use cases that admit simple solutions (though providing high gain) are worthy contributions to the progress in this research area. In this way, the role that data analytics and AI will realistically play in 5G networks can be better delimitated.

III. Use Case: Data-Driven RAC with FWA users

Fixed Wireless Access (FWA), also referred to as fixed wireless broadband, is a way for homes and businesses to connect to the internet over the air using mobile network technology, rather than a physical connection through traditional fibre-optic or copper wiring, which is costly and time-consuming to deploy, maintain and upgrade. While FWA services are already a reality today by using existing 4G technology [20], it is predicted that 5G will lead to an explosion in FWA services [21] because of the much higher internet speeds and reduced latency achievable. Thus, 5G FWA services are anticipated to offer a user experience well on a par with that achieved in fibre-line connections, this being particularly attractive in urban, suburban or rural areas where fibre access is difficult or expensive to implement [22]. Indeed, many mobile operators are expecting 5G FWA services to be a key driver in early 5G network adoption for its ability to solve the 'last mile' issues and the roll out of these services is expected to start well ahead of 5G's predicted 2020 mobile rollout [20].

In this context, this paper proposes an application of data analytics in the RAC function that controls the assignment of radio resources for Guaranteed Bit Rate (GBR) services in a deployment scenario where the NG-RAN is used to deliver both FWA and mobile services. The solution is built upon the exploitation of two aspects. First, it is considered that the RAC algorithm embedded in the NG-RAN nodes is able to discriminate between stationary and mobile UEs when performing GBR services' admission decisions. Second, the RAC decision-making logic is designed to be fed with data analytics information that characterises the spectral efficiency achieved at cell and at UE level for stationary users. Such data analytics information is computed from measurements collected from the NG-RAN nodes and assumed to be available in real-time for supporting the RAC decisions

IV. Solution Description

Fig. 1 depicts the network architectural view for the proposed data analytics-enhanced RAC solution. The figure represents a gNB, i.e. the NG-RAN node supporting 5G NR, decomposed into a Central Unit (CU) and a Distributed Unit (DU) [23]. Considering the current functional split defined in 3GPP Release 15, the gNB CU hosts the upper layers of the radio interface protocol stack, i.e. Packet Data Convergence Protocol (PDCP) and above, while the gNB DU hosts the lower layers, i.e. Radio Link Control (RLC) and below. Similar to the approach being considered in 3GPP for handling data analytics services in the 5G core network (i.e. NWDAF [14]), a new functional entity called RAN Data Analytics Function (RANDAF) is proposed here to serve as a collector of measurements from NG-RAN nodes and as a provider of data analytics services for enhanced RRM functionality. The different components and interactions of the proposed solution are described in the following.

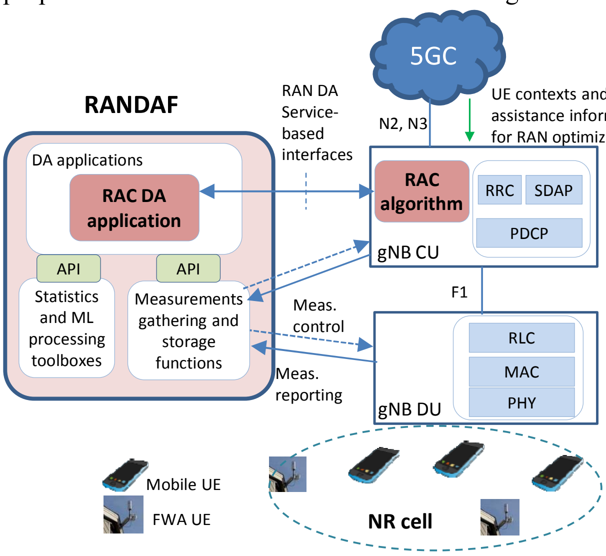

Fig. 1. Data analytics enhanced Radio Admission Control (RAC) for cell deployments with FWA and mobile users

A. RAC algorithm

The RAC algorithm is hosted in the gNB-CU and is triggered whenever a new GBR Data Radio Bearer (DRB) has to be established in one of its handled cells. It has to decide whether the new DRB can be admitted or not. In general terms, the RAC needs to ensure that the cell will have sufficient Resource Blocks (RB) to fulfil the GBR requirements of the new DRB as well as those of the already admitted DRBs. For this reason, a key issue of the RAC process is to properly predict the amount or RBs that will be needed by the new DRB. This is challenging because, due to the random behaviour of the radio channel and the experienced propagation conditions, the bit rate requirements of a DRB cannot be deterministically mapped to a number of required RBs, but they need to be statistically estimated. While the usual approach in existing RAC schemes (see e.g. [25]) is to perform this estimation for a new UE based on measurements of all users in a cell, the proposed solution allows the RAN algorithm to incorporate UE-specific data analytics information provided by the RANDAF.

Specifically, the proposed RAC accepts a new incoming DRB with requirement *GBR* if the following condition holds:

$$\rho + \Delta\rho \le \rho_{\max} \qquad (1)$$

where ρ is the actual number of RBs consumed by the GBR DRBs already admitted in the cell, obtained as an average over a certain time window, ρ_{max} is the maximum number of RBs to be used by GBR DRBs and Δρ is the estimation of the RBs required by the new GBR DRB. This estimation involves information provided by the RANDAF and distinguishes the cases of FWA and mobile UEs. In particular, for FWA UEs, Δρ is defined as $GBR/(S_{UE,i}\cdot B)$, where B is the RB bandwidth and $S_{UE,i}$ is the average spectral efficiency associated to the i-th FWA UE. Instead, in case of mobile UEs, Δρ is defined as $GBR/(S_{cell}\cdot B)$, where S_{cell} is the average spectral efficiency associated to the whole cell. Clearly, $S_{UE,i}$ could substantially differ from S_{cell} (e.g., $S_{UE,i}$>>S_{cell} if the FWA UE is close to a base station covering a wide area).

The values of $S_{UE,i}$ and S_{cell} are obtained by the gNB-CU from the RANDAF, as explained in the following sub-section. In turn, the distinction between FWA and mobile UEs within the gNB-CU can be done thanks to the information provided from the 5GC as part of the so called Core Network (CN) assistance information for RAN optimization [14]. In particular, the CN assistance information contains, among others, the "expected UE mobility" field used to indicate whether a UE is expected to be stationary or mobile. Such knowledge may be derived e.g. from statistical information collected at the 5GC or directly from subscription information.

B. RAN Data Analytics Function

As depicted in Fig. 1, the RANDAF is conceived as an execution platform for Data Analytics (DA) applications. A particular DA application may be programmed to deliver general-purpose data analytics information that can be exploited by several RRM/SON functions or just be designed to support the operation of a specific RRM/SON algorithm. Through Application Programming Interfaces (APIs), the DA applications leverage the following capabilities within the RANDAF:

- Measurements gathering and storage functions: These functions are used to configure the measurement collection tasks and gather the measurements retrieved (e.g. measurements files extraction) or received (e.g. stream-based monitoring) from NG-RAN nodes in accordance with DA applications' needs.
- Statistics and ML processing toolboxes: This is the collection of libraries with data processing functions for the DA applications to manipulate and generate the necessary statistics from the collected data or to extract

behavioural models through the use of ML functions (e.g. classification, predictors, clustering, etc.).

The RAC algorithm described in section IV.A is supported at the RANDAF by the RAC DA application, responsible of computing and storing the values of the average spectral efficiency S_{cell} for each cell and $S_{UE,i}$ for each i-th FWA UE. For that purpose, it is assumed that the RAC DA application is aware of the list of stationary UEs (which include the FWA UEs) in each cell. This can be done via provisioning mechanisms at management level (e.g. the list of FWA UEs is configured by the network operator) or automatically generated from information provided at runtime by the gNB to the DA application (e.g., when a FWA UE is first registered in the network, its serving gNB informs the RAC DA application about the presence of a new "stationary" user).

In the cells for which the RAC DA application provides support to the RAC decisions, the RAC DA application triggers the measurements gathering and storage function at the RANDAF for starting the collection of the measurements needed to compute S_{cell} and $S_{UE,i}$. Specifically, the Channel State Information (CSI)-reports sent by the active UEs of the cell to the gNB-DU are collected, which include [24]:

- Channel Quality Indicator (CQI): Each CQI is mapped to a channel coding rate r and modulation scheme with m bits/symbol that can be used by the UE.
- Rank Indicator (RI): It indicates the suitable number of transmission layers for downlink transmission in case of spatial multiplexing.

From each CSI-report of each UE, the RAC DA application gets one sample of r, m, RI and estimates the spectral efficiency as:

$$S = RI \cdot m \cdot r \quad (b/s/Hz) \tag{2}$$

Then, the average value of S_{cell} is estimated from all the spectral efficiency samples obtained from all the UEs of the cell as:

$$S_{cell} = \frac{1}{K}\sum_{k=1}^{K} S(k) \tag{3}$$

where $S(k)$ denotes the k-th sample of spectral efficiency and K is the total number of collected samples.

Similarly, the estimation of $S_{UE,i}$ is done considering only the spectral efficiency samples obtained from the i-th FWA UE as follows:

$$S_{UE,i} = \frac{1}{K_i}\sum_{k=1}^{K_i} S(k,i) \tag{4}$$

where $S(k,i)$ denotes the k-th sample of spectral efficiency of the i-th FWA UE and K_i the total number of collected samples for this UE.

The number of samples K and K_i to be collected could be set in accordance with the desired accuracy (e.g. defined in terms of a confidence interval width) of the estimated average values with respect to their real values. For example, in case of a desired 95% confidence interval width W for estimating S_{cell}, the required number of samples would be $K \approx 16 \cdot \sigma^2/W^2$ where σ is the standard deviation of the spectral efficiency in the cell [26].

The resulting values of S_{cell} and $S_{UE,i}$ are stored and provided upon request to the RAC algorithm at the gNB-CU. For that purpose, the interaction between the RAC DA application and the gNB-CU is conceived based on the definition of a service-based interface in line with the service-based architecture adopted for the 5GC. Specifically, whenever the RAC function is executed for a new DRB of the i-th UE, the RAC will query the RAC DA application asking for $S_{UE,i}$ if the GBR DRB request is associated with a stationary user or S_{cell} if it is associated with a mobile user.

It is worth mentioning that the stored values of S_{cell} and $S_{UE,i}$ will be valid as long as the cell's conditions do not change (i.e., remain stationary). In practice, the system can face modifications in the operating conditions of a cell, such as changes in the spatial traffic distribution, modifications in the configuration parameters of the cell (e.g., transmit power, azimuth, etc.), modifications in the neighbour cells leading to different interference patterns, changes in the environment due to the construction of new buildings, etc. Therefore, the RAC DA application needs to incorporate functions to detect relevant changes that affect the validity of S_{cell} and $S_{UE,i}$ so as to gather again the measurements and regenerate the statistics.

V. PERFORMANCE EVALUATION

In order to assess the performance of the proposed approach, let us consider a NR macrocell in 3.5 GHz deployed in a rural scenario characterised with the parameters listed in Table I. The macrocell provides service to a household with a FWA subscription that includes GBR services (e.g. high-resolution video streaming services) as well as to mobile subscribers camping within the cell coverage.

The FWA subscriber is assumed to be equipped with a residential gateway (RG) installed in the household that connects to the gNB as a FWA UE. Inside the household, different devices (e.g. TV sets, laptops, tablets, smartphones, etc.) are attached to the RG by means of Wi-Fi or wired solutions (e.g. optical fibre, power line communications). Accordingly, the traffic generation model assumes that the FWA UE generates GBR sessions associated to the devices inside the household following a Poisson arrival process with rate λ sessions/s. The duration of the sessions is exponentially distributed with average T (s). Each session requires establishing a DRB with the GBR value indicated in Table I. Then, the results are presented for different values of the average GBR offered load, given by $L=\lambda \cdot T \cdot GBR$ (Mb/s).

The results consider two possible locations of the FWA UE within the cell coverage:

- Location 1 (close to the gNB): The household is located at 100m of the gNB, experiencing a total propagation loss of 85 dB. The average spectral efficiency S_{UE} observed in this location is 8.8 b/s/Hz.
- Location 2 (far from the gNB): The household is located at 1km of the gNB, experiencing a total propagation loss of 128 dB and an average spectral efficiency S_{UE}=2.5 b/s/Hz.

It is also assumed that mobile subscribers only activate non-GBR services, which make use of the remaining capacity of the cell that is not used by the GBR services. Therefore, and in order to ensure that the cell will always have some remaining capacity to properly serve non-GBR services, the

RAC regulates the total RB consumption of GBR services by setting ρ_{max} to 60% of the total RBs of the cell (see Table I). In this way, at least 40% of the RBs remain available for non-GBR services.

Regarding the average spectral efficiency value S_{cell}, when the mobile UEs are uniformly distributed around the gNB within a radius R km and the FWA household is in location 1 with an average offered load L=20 Mb/s, it is obtained that S_{cell}={6.3, 4.7, 4.2} b/s/Hz for R={1,2,3} km, respectively, and a 95% confidence interval width W=0.02. In turn, when the FWA household is in location 2, the resulting values are S_{cell}={5.3, 3.8, 3.2} b/s/Hz.

TABLE I. SCENARIO PARAMETERS

Parameter	Value
Cell radius R	Varied: 1 km, 2 km, 3 km
Path loss and shadowing model	Rural macrocell model of [27] with gNB antenna height 35m, UE height 1.5m and minimum distance of 10m between UE and gNB.
Base station antenna gain	5 dB
Frequency	3.6 GHz
Total transmitted power	43 dBm
Cell bandwidth	20 MHz (N_{RB}=106 RBs with subcarrier spacing 15 kHz)
UE noise figure	9 dB
Spectral efficiency model	Model in section A.1 of [28] with maximum spectral efficiency 8.8 b/s/Hz.
Required bit rate of GBR services	Varied: 5 Mb/s, 10 Mb/s, 20 Mb/s
ρ_{max}	$0.6 \cdot N_{RB}$=63.6

Fig. 2 presents the blocking probability experienced by GBR services, i.e. the probability that the RAC rejects an incoming GBR DRB. Results are presented as a function of the average GBR offered load L for the case when the FWA household is in location 1, the cell radius is R=3 km and the required GBR is 10 Mb/s. The figure compares the results of the proposed approach with a reference approach that considers instead the average cell behaviour, i.e. it applies the condition (1) with $\Delta\rho=GBR/(S_{cell} \cdot B)$. It is observed that the proposed approach achieves substantially lower blocking than the reference scheme. The reason is that, in location 1, $S_{UE,i}$ is higher than S_{cell} and thus the reference approach overestimates the term $\Delta\rho$ with respect to the real RB consumption of the DRBs generated by the FWA UE. Therefore, it leads to some unnecessary rejections that can be avoided by the proposed approach thanks to its more accurate estimation of $\Delta\rho$.

For the case that the FWA household is in location 2, where $S_{cell} > S_{UE,i}$, the situation reverses since the reference approach underestimates the term $\Delta\rho$ with respect to the real RB consumption of the GBR DRBs of the FWA UE. Therefore, it makes incorrect admission decisions that lead to exceeding the maximum capacity of GBR users established by ρ_{max}. As a result of these wrong admissions, less capacity remains for non-GBR services. This is depicted in Fig. 3, which shows, on the one hand, the throughput achieved by non-GBR services when the household is in location 2 and, on the other hand, the probability that the GBR services consume more than the limit of ρ_{max} RBs. It can be observed that, thanks to the better estimation of the term $\Delta\rho$ carried out by the proposed approach, the RAC allows better controlling the cell capacity left for non-GBR services and thus increasing the bit rate available for these services. Instead, the reference approach, leads to a high probability of exceeding the limit of ρ_{max} RBs. Obviously, this better control of the capacity is at the expense of the GBR services that experience a worse blocking probability, as seen in Fig. 4.

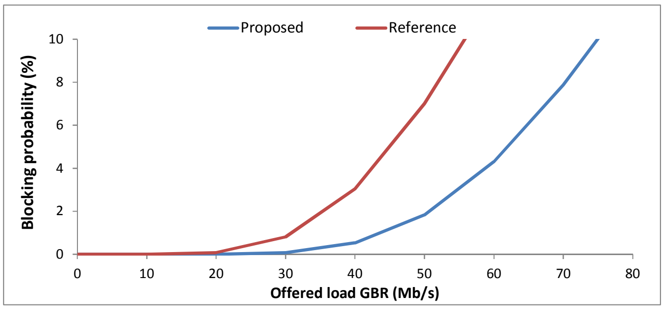

Fig. 2. Blocking probability experienced by GBR services assuming a household with FWA in location 1.

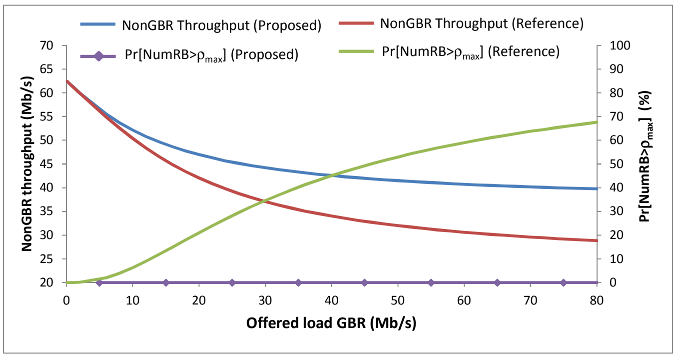

Fig. 3. Throughput of non-GBR services and probability that GBR services consume more than ρ_{max} RBs as a function of the offered GBR load assuming household with FWA in location 2.

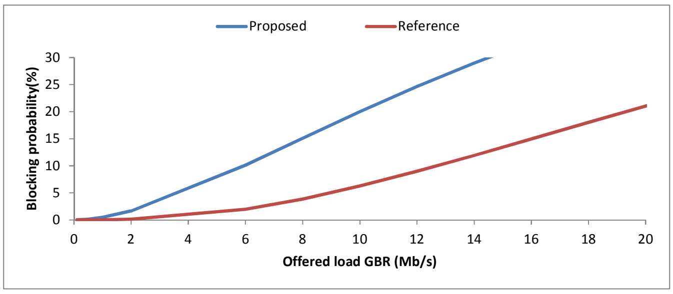

Fig. 4. Blocking probability experienced by GBR services assuming a household with FWA in location 2.

In order to better delimitate the interest of the proposed RAC solution, a wide range of different scenarios reflecting different propagation conditions and different requirements of the GBR services needs to be analysed. In this respect, Table II presents the obtained maximum GBR capacity, measured as the maximum offered load of GBR services to ensure a maximum blocking probability of 2%, for different values of the GBR requirement and for different cell radius. Results consider that the FWA household is in location 1 and compare the capacity achieved with the proposed approach and the reference scheme, showing the gain in percentage in parenthesis. Results reveal that the highest gains are obtained for high GBR values because they lead to higher values of the term $\Delta\rho$ considered in the admission condition (1). Therefore, the overestimation of $\Delta\rho$ that occurs when taking the average cell behaviour has a more significant impact in terms of the maximum users that can be admitted. On the contrary, for low GBR values, the term Δρ represents a smaller contribution to the admission condition (1) and, therefore, errors in the estimation of this term have less impact in terms of capacity.

Regarding the impact of the propagation conditions, analysed through the different radius R in Table II, it is observed that the highest gains are achieved for large values of R. The reason in this case is that, since the average spectral efficiency of the cell decreases with R, it becomes less representative of the spectral efficiency experienced by the FWA UE at location 1, leading to higher overestimation of Δρ and to lower capacity. Based on these results, it is concluded that the proposed RAC solution finds high interest in cells providing wide coverage area and serving high capacity-demanding FWA users.

TABLE II. MAXIMUM GBR CAPACITY (MB/S) FOR DIFFERENT SCENARIOS

	GBR=5 Mb/s		GBR=10 Mb/s		GBR=20 Mb/s	
R	Ref.	Proposed	Ref.	Proposed	Ref.	Proposed
1 km	61.5	66 (+4.5%)	43	51 (+18%)	22	33 (+50%)
2 km	61.5	66 (+4.5%)	43	51 (+18%)	22	33 (+50%)
3 km	61.5	66 (+4.5%)	36.5	51 (+40%)	12.1	33 (+175%)

VI. CONCLUSIONS

Data analytics is expected to be a key capability to efficiently operate 5G networks and turn into value the big amount of data available within a mobile network. In this direction, this paper has discussed the applicability of exploiting data analytics for supporting the operation of the RRM algorithms embedded within the NG-RAN nodes. After outlining the need for standardised and open data analytics frameworks and elaborating on the special characteristics of the NG-RAN that will impact on the design of data analytics solutions, the paper has developed a simple use case for the exploitation of data analytics in the Radio Admission Control algorithm in scenarios where the cell capacity is shared among Fixed Wireless Access users and mobile users. A specific solution has been presented that exploits the knowledge about the FWA users to perform a more accurate estimation of their resource consumption to be used when making admission decisions. The solution has been described both at architectural and procedural level and it has been evaluated under different scenario conditions to better delimitate its relevance. It has been concluded that the main interest of the proposed solution arises in cells providing wide coverage area and serving high capacity-demanding FWA users.

ACKNOWLEDGEMENT

This work has been supported by the EU funded H2020 5G-PPP project 5G ESSENCE under grant agreement 761592 and by the Spanish Research Council and FEDER funds under SONAR 5G grant (ref. TEC2017-82651-R).

REFERENCES

[1] A. Imran, A. Zoha, A. Abu-Dayya, “Challenges in 5G: How to Empower SON with Big Data for Enabling 5G”, IEEE Network, November/December, 2014, pp. 27.33.

[2] R. Agrawal, “Machine Learning for 5G RAN”, Nokia, 2018, https://www.itu.int/en/ITU-T/Workshops-and-Seminars/20180807/Documents/Ran_ML.pdf, Retrieved March 2019.

[3] C-L. I, Y. Liu, S. Han, S. Wang, G. Liu, “On Big Data Analytics for Greener and Softer RAN”, IEEE Access, August, 2015.

[4] Q. Yan, W. Chen, H.V. Poor, “Big Data Driven Wireless Communications: A Human-in-the-Loop Pushing Technique for 5G Systems”, IEEE Wireless Communicatiuons, February, 2018.

[5] S. Han, C-L I, G Li, S. Wang, Q. Sun, “Big Data Enabled Mobile Network Design for 5G and Beyond”, IEEE Comms. Mag, Sept. 2017.

[6] M.S. Parwez, D. B. Rawat, M. Garuba, “Big Data-Analytics for User-Activity Analysis and User-Anomaly Detection in Mobile Wireless Network”, IEEE Trans. on Industrial Informatics, Vol. 13, No. 4, Aug., 2017.

[7] B. Hussain, Q. Du, P. Ren, “Semi-supervised learning based big data-driven anomaly detection in mobile wireless networks”, China Communications, Vol. 15, No. 4, 2018

[8] A. Gómez-Andrades, R. Barco, P. Muñoz, I. Serrano, “Data Analytics for Diagnosing the RF Condition in Self-Organizing Networks”, IEEE Transactions on Mobile Computing, Vol. 16, No. 6, June, 2017.

[9] S. Holmbacka, et al. “Alarm Prediction in LTE Networks”, Int. Conf. on Telecomms. (ICT), 2018.

[10] P. Chiu, J. Reunanen, R. Luostari, H. Holma, “Big Data Analytics for 4.9G and 5G Mobile Network Optimization”, IEEE VTC Spring, 2017.

[11] J. Gao, X. Cheng, L. Xu, H. Ye, “An Inteference Management Algorithm using Big Data Analytics in LTE Cellular Networks”, Int. Symp. on Comms. and Inf. Technologies (ISCIT), 2016

[12] C-L Lee, W-S Su, K-A. Tang, W-I. Chao, “Design of handover self-optimization using big data analytics”, Asia-Pacific Network Operations and Management Symposium, 2014

[13] J. Moysen, M. García-Lozano, S. Ruiz, L. Giupponi, “Conflict Resolution in Mobile Networks: A Self-Coordination Framework Based on Non-Dominated Solutions and Machine Learning for Data Analytics”, IEEE Computational Intelligence Magazine, May, 2018.

[14] 3GPP TS 23.501 v15.2.0, “System Architecture for the 5G System (Release 15)”, June, 2018.

[15] 3GPP TR 23.791 v16.0.0, “Study of Enablers for Network Automation for 5G (Release 16)”, December, 2018.

[16] 3GPP Study Item, “Study on RAN-centric Data Collection and Utilization for NR”, 3GPP TSG RAN Meeting #80, La Jolla, USA, June 11 - 14, 2018”

[17] Open RAN Alliance, “O-RAN: Towards an Open and Smart RAN”, White Paper, October 2018

[18] F. Bouali, O. Sallent, J. Perez-Romero, R. Agustí, “Exploiting Knowledge Management for Supporting Multi-Band Spectrum Selection in Non-Stationary Environments”, IEEE Transactions on Wireless Communications, Vol. 12, No. 12, December, 2013.

[19] H. He, E. A. García, “Learning from Imbalanced Data”, IEEE Trans. on Knowledge and Data Engineering, Vol. 21, No. 9, September, 2009.

[20] J. Mundy, “What is Fixed Wireless Broadband?” https://5g.co.uk/guides/what-is-fixed-wireless-broadband/

[21] Ericsson Mobility Report, June 2018.

[22] 5G Fixed Wireless Internet – What Is It, Why Is It Good And Where Can You Get It? https://www.lightningbroadband.com.au /technology/5g-fixed-wireless-broadband/

[23] 3GPP TS 38.401 v15.2.0, “NG-RAN; Architecture description (Release 15)”, June, 2018.

[24] E. Dahlman, S. Parkvall, J. Sköld, *5G NR. The next generation wireless access technology*, Elsevier, 2018.

[25] R. Kwan, R. Arnott, M. Kubota, “On Radio Admission Control for LTE Systems”, IEEE 72nd Vehicular Technology Conference (VTC Fall, Ottawa, Canada, September, 2010.

[26] M. Bonamente, “Statistics and Analysis of Scientific Data”, Springer, 2013..

[27] 3GPP TR 38.901 v15.0.0, “Study on channel model for frequencies from 0.5 to 100 GHz (Release 15)”, June, 2018.

[28] 3GPP TR 36.942 v15.0.0, “Radio Frequency (RF) system scenarios (Release 15)”, June, 2018.