

Data Analytics Architectural Framework for Smarter Radio Resource Management in 5G Radio Access Networks

R. Ferrús, O. Sallent, J. Pérez-Romero
Universitat Politècnica de Catalunya (UPC)

Abstract— Along with supporting new business opportunities, data analytics –understood as the pursuit of extracting meaning from raw data– will be key to getting 5G Radio Access Networks (RANs) efficiently rolled out and operated. The exploitation of data analytics services within the RAN itself, that is, to empower the operation of the RRM algorithms executed within the RAN nodes, is an incipient area of interest that broadens the more well-established applicability of using data analytics for on-top-of-the-RAN operations. In this respect, this article is an attempt to put a spotlight on this topic by motivating a plausible data analytics architectural framework for data-driven, smarter RRM algorithms.

INTRODUCTION

Tightly integrated with Artificial Intelligence (AI) technologies, data analytics is anticipated to be a game-changer for network operators at all levels, ranging from business, service and network management levels (e.g. customer care, service assurance) down to the level of driving the operation of specific functions embedded within the network nodes (e.g. traffic routing).

This is especially relevant for the Next Generation Radio Access Network (NG-RAN) of 5G systems, which is set to undergo a significant transformation through the introduction of data analytics to cope with its increasing complexity and scale. Indeed, the exploitation of data analytics in 5G networks has raised much research interest as reflected by different papers. Some of them, such as [1][2], establish general frameworks for the application of data analytics in 5G systems. Focusing on the RAN, several papers have 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 [3], diagnosis [4] or alarm prediction [5]. Some other works have dealt with self-optimization aspects, such as tilt and power adjustment [6], interference management, self-configuration of neighbour cell lists [7] and coordination of Self-Organizing Network (SON) functions [8]. With respect to the considered methodologies, most of the abovementioned works apply Machine Learning (ML) tools to process the data, such as K-means clustering [3], random forest prediction [5], Self-Organizing Maps [4] or regression models [6][8]. These works have also considered different types of data sources, such as call detail records [3], mobile traces [4], counters [5] or User Equipment (UE) measurement reports [6][8].

In this context, the state-of-the-art analysis reveals that the main focus so far has been placed on the design and evaluation of specific solutions that extract and exploit data analytics information for on-top-of-the-RAN operations, that is, within the scope of the Operations, Administration and Maintenance (OAM) systems. However, it is worth noting that little attention has been placed to date to the exploitation of data analytics to drive the operation of the Radio Resource Management (RRM) algorithms executed within the NG-RAN nodes (such as admission control, mobility control, scheduling, inter-cell interference coordination, etc.), with only a few works starting to appear in this area (e.g. [9], [10]).

To contribute to fill this gap and consistently with recent initiatives arising from the standardisation domain (e.g. [11], [12]), this paper describes a novel functional framework for the extraction and delivery of RAN data analytics information services intended to enable smarter, data-driven RRM functions within the RAN nodes, that is, enhanced RRM algorithms by exploiting the extracted RAN data analytics information. While the computation of basic statistics from RAN measurements is already used in conventional RRM, the proposed architecture attempts to decouple the processing and delivery of the RAN data analytics services from the RRM functionality that consumes them. In this way, a much richer set of data analytics information can be produced, thanks to the use of specialized platforms, separate from the RAN nodes, for big data collection and knowledge extraction via a variety of ML processing techniques. Furthermore, the proposed approach can spur innovation in multi-vendor scenarios where the RRM software and RAN data analytics services may be delivered by different providers.

The rest of the paper is organised as follows. First, we set the stage with an overview of the Third Generation Partnership Project (3GPP) work regarding data analytics support in 5G systems. Next, we describe the proposed functional framework. This is followed by illustrative assessments of potential RRM improvements based on real measurements collected in a live 4G/LTE network. Finally, concluding remarks are drawn.

DATA ANALYTICS IN 3GPP 5G SYSTEMS

While still in a preliminary stage, 3GPP is addressing the specification of the network functions and management services intended to become the building blocks of data analytics solutions within 5G systems. In particular, in the latest approved 5G specifications (Release 15), 3GPP has introduced two

functional entities, one at network level called Network Data Analytics Function (NWDAF) and another at management level called Management Data Analytics Function (MDAF), both conceived to provide data analytics services. These analytics services can consist of either statistical information of the past or predictive information of the future. In addition to NWDAF and MDAF, 3GPP capabilities for tracing subscriber/UE activity and Core Network (CN) assistance information for RAN optimization are also anticipated to become key pillars of the data analytics solutions. A schematic representation of these building blocks in the context of a 5G network is depicted in Fig. 1. A more detailed description of each block follows.

As part of the 5G Core network (5GC), the NWDAF (“A” tag in Fig. 1) is conceived as a network function (NF) aimed to provide network analytics information to other NFs of the 5GC. Hence, through a service-based oriented interface called N_{nwdaF} , NFs can request/cancel subscription and have access to network analytics information. Under Release 15, the supported analytics information is limited only to the load level of a slice instance. Accordingly, only network functions such as the Policy Control Function (PCF) and the Network Slice Selection Function (NSSF) have been identified in such release as potential consumers of the slice load level information: e.g. the PCF may use it for QoS parameters selection while the NSSF may use it for slice selection. Under Release 16, new use cases have been studied in 3GPP TR 23.791 and architecture enhancements are being introduced at the normative stage in 3GPP TS 23.288 to extend the data analytics services delivered by the NWDAF for e.g. customized mobility management, among others. However, Release 16 work on NWDAF improvements only focuses on the exploitation of the analytics within the 5GC NFs, leaving a further analysis of the exploitation of the data analytics services within the NG-RAN to be addressed in future Release 17 under the topic of RAN centric data collection enhancements [11].

As part of the 3GPP management system, the MDAF (“B” tag in Fig. 1) is the component in charge of providing management data analytics services (MDAS). For example, the MDAF can retrieve OAM data from the NFs and produce management analytics information used to recommend appropriate management actions to the network operator. The MDAF produced analytics can also be made available and consumed by other management functions such as SON functions executed within the OAM systems, referred to as centralized SON (C-SON).

Tracing capabilities (“C” tag in Fig. 1) are also being improved within 5G systems (e.g. new trace reporting methods for non-file-based tracing are being studied in 3GPP TR 28.806). They can be used to obtain detailed information of subscriber/UE activity at session level from the network functions handling them. Tracing capabilities also include the so called Minimization of Drive Tests (MDT) feature, which allows for the collection of immediate or logged measurements directly from UEs. Tracing information can be exploited for multiple purposes such as troubleshooting or network performance optimization.

Last but not least, given that persistent data on

subscribers/UEs (e.g. subscriber information) is only held within the 5GC, 3GPP specifications allow the 5GC to provide assistance information to the NG-RAN (“D” tag in Fig. 1) about the subscriber/UE behaviour at the time it connects to the RAN (i.e. when the “UE Context” is activated). This includes information such as “Expected UE mobility” (i.e. whether the UE is expected to be stationary or mobile) or “Expected UE activity behaviour” (i.e. expected pattern of connected/idle states changes).

Compliant and complementary to the 3GPP specifications, it is worth noting that Open-RAN Alliance (O-RAN) is working to establish the foundation for building the NG-RAN based on open hardware and software reference designs as well as embedding data analytics and ML technologies. As part of the O-RAN architecture [12], a new interface between the RAN and the OAM systems is being specified (named A1 interface) whose purpose is the delivery of policy-based guidance, trained ML models (it is assumed that ML algorithms can be used within the RAN to implement some near real time RRM functions) and *enrichment information* that can be used to optimize RRM under certain conditions. While the concept of supporting *enrichment information services* is only drafted and no specifications have been produced yet in this regard, we believe that the solution proposed in this paper is well aligned with this view.

FUNCTIONAL FRAMEWORK FOR DATA ANALYTICS IN THE NG-RAN

Consistently with the 3GPP and O-RAN work previously outlined, this section develops a proposal that extends the applicability of data analytics into the NG-RAN. In 5G, the NG-RAN is a multi-RAT access network that integrates gNB nodes, which provide access via the NR interface, and ng-eNB nodes, for LTE access. Both gNB and ng-eNB nodes are connected to the 5GC via NG interfaces (N2, N3) and may be interconnected with each other via Xn interfaces. A functional split into a Central Unit (CU), for upper radio interface protocol layer processing, and a Distributed Unit (DU), for lower protocol layer processing, is supported using standardised interfaces (F1 for the gNB split and W1 for the ng-eNB split). On the radio interface side, multiple spatially distributed Transmit Reception Points (TRPs) with beamforming capabilities can be used. Fig. 2 provides an illustrative view of the NG-RAN architecture and introduces the elements that guide the following discussion about the potential allocation of data analytics capabilities within the NG-RAN.

As depicted in Fig. 2, RRM is a core functionality inside NG-RAN nodes to handle the necessary decision-making logic for efficient radio interface utilization. RRM functions are in charge of, among others, admission control and radio bearers’ configuration, mobility control and time/frequency radio resource allocation (i.e. scheduling). While the scope of the RRM functions in 5G NR is similar to LTE, there are important differences in their realization because of the new features coming with 5G NR such as network slicing, beamforming, bandwidth parts (BWPs), low-latency resource allocation and intra-cell/inter-beam mobility. Besides, tightly coupled with the RRM functions, NG-RAN nodes can also embed

SON functionality, known as Distributed SON (D-SON) in contrast to the C-SON at OAM level, to automatically configure and adjust, via control signalling, different operational parameters such as cell neighbour relations and mobility load balancing and robustness optimization settings. Currently, the operation of RRM and D-SON algorithms within the RAN nodes is mainly driven by (1) *configuration parameters* set from the OAM systems, which span from basic cell configuration settings (e.g. cell identifiers, channel configuration) to very detailed aspects of specific features (e.g. scheduler configuration, BWPs configuration); (2) *UE context information*, which includes UE radio capabilities, QoS flows parameters as well as other information that could be provided by the 5GC (e.g. expected UE activity behaviour, expected UE mobility); and (3) *measurements* performed locally within the NG-RAN nodes or reported by the UEs. In this regard, a rich variety of measurements and measurement reporting mechanisms are defined by 3GPP for OAM purposes as well as for supporting the operation of the network (e.g. UE measurements reporting for mobility control). A summary of the standardised measurements types for 5G NR in Release 15 is given in Table 1, pointing out the 3GPP specs where the measurements are defined.

A. RAN data analytics functions

As laid out in Fig. 2, the proposed functional framework consists of a distributed architecture of two types of components, named RAN Data Analytics Function (RANDAF) and Local Data Analytics Function (LDAF), interacting via a service-based control interface for the exchange of data analytics information within the RAN, denoted as Analytics Information Service (AIS).

The RANDAF is conceived as a data analytics capability separate from the RAN nodes and equipped with the necessary storage and computation resources to properly cope with the large spatial dimension (e.g. multi-node/multi-cell) and temporal dimension (e.g. long time series of collected data) of the data analytics within the NG-RAN. Internally, the RANDAF serves as an execution platform where multiple Data Analytics (DA) applications can be executed to produce the different sets of data analytics information, discussed later in this paper. To do their job, DA applications may exploit different Application Programming Interfaces (APIs) to get access to:

-Measurements gathering and storage functions, used to configure the measurement collection tasks and gather the measurements from NG-RAN nodes in accordance with DA applications' needs. The collection of OAM information can use the performance management (PM) and fault management (FM) services as defined in 3GPP TS 28.532, 28.550 and 28.545 as well as the management services for collecting subscriber/UE tracing and MDT reports. File-based and stream-based reporting methods are envisioned. In addition to OAM data, RAN nodes may also provide behavior data related to individual UEs or UE groups (e.g. UE Radio Resource Control [RRC] states) as well as other non-OAM metrics (e.g. number of connected UEs per spatial and temporal dimensions) that could be necessary for the RANDAF to compute the RAN data analytics.

-Statistics and ML processing toolboxes, as the set of data processing libraries for DA applications to manipulate and generate the necessary statistics from the collected data or to extract behavioral models through the use of ML functions (e.g. classification, predictors, clustering).

On the other hand, the LDAF is conceived as a lightweight data analytics capability with reduced resource requirements (e.g. reduced storage, no need of specialized processing hardware) embedded inside RAN nodes to process the measurements locally available and produce analytics information (e.g. non-OAM metrics of the RAN node) that could be directly exploited by the co-located RRM/D-SON algorithms. Besides, such information could be also retrieved by the RANDAF for aggregation or further processing. Note that this would reduce the amount of raw measurements needed by the RANDAF to produce similar information in case that LDAF was not in place. The scale in terms of the physical resources allocated for the data analytics capability and the scope of the analytics produced by the LDAF should be adapted to fit the needs of the different types of RAN nodes (e.g. CU or DU units) and RAN deployments (e.g. centralized or distributed).

The information exchange between the RANDAF and LDAF is to be supported via the AIS interface in the form of data analytics information services. These services would include subscription services to request specific reports along with the associated information delivery services for data exchange (e.g. push and/or pull transfer methods). Indeed, the implementation of the AIS interface could be based on the same design principles adopted by 3GPP for the service-based interfaces in the 5GC (3GPP TS 29.501).

B. RAN data analytics information

Three outstanding categories of RAN data analytics information that can be extracted by RANDAF/LDAF entities and exploited for smarter RRM/D-SON are:

-Radio resource utilization analytics: It could include volume (e.g. load level in terms of utilized radio resource blocks) and efficiency (e.g. spectral efficiency) indicators about the radio resources utilization. It could cover different granularities (e.g. per cell, per beam, per UE, per network slice, per QoS class) and be reported for specific entities or groups of entities (e.g. a specific UE in a given cell, a network slice in a group of cells). Radio resource utilization values could be statistical information or predictions. Different validity periods for the analytics (e.g. short-term/long-term, specific time intervals) with different levels of certainty/confidence could be possible. For instance, the knowledge of the load level for a network slice that spans several cells in different gNBs can be used to enforce admission control decisions in each gNB so that aggregated resource constraints at network slice level can be obeyed (e.g. maximum authorized radio load).

-Mobility analytics: It could include information about the UE locations and mobility patterns with different granularities (e.g. per control beam, per areas defined with geographical coordinates). It could also include the time that the UE moves to the locations and the duration that stays in each of them. For instance, the awareness that a UE is likely going to follow a given

trajectory or remain stationary at a given position can be exploited for proper reservation and allocation of the radio resources via admission control, mobility control and congestion control. In beamforming settings, this information can also be exploited to feed beam selection decisions and mobility between beams. **-Communications activity analytics:** It could include communication pattern parameters (e.g. frequency of RRC connection requests, dwell time distribution in RRC connected/inactive/idle states). It could also include Key Performance Indicators (KPIs) per UE (e.g. number of successful/failed handovers). For instance, the knowledge that a particular UE experiences KPIs lower than expected with regard to e.g. retainability (e.g. higher number of failed handovers than other UEs) could be used for mobility robustness optimization.

C. Further considerations

Conceptually, the RANDAF resembles the NWDAF entity defined by 3GPP for the 5GC. However, there are several reasons from a practical/implementation perspective that advocate for a RAN-domain specific entity. First, the scope of the delivered analytics services differs substantially between the RAN and 5GC domains (e.g. location resolution of meters required for the RAN domain in contrast to cell-level resolution in the 5GC). Second, a separate, local implementation for the RAN can improve responsiveness and scalability since faster decisions and shorter time scale operation can be facilitated. Similarly, a reduction of the amount of information that needs to be moved from the cell edge to central locations for data analytics processing can be achieved. Last but not least, the separation between the NWDAF and RANDAF may also be necessary from a business perspective when service delivery involves several stakeholders, as it could be the case of RAN sharing scenarios with different service providers in charge of the RAN and the 5GC.

From the management plane perspective, the proposed RANDAF entity can be also seen as a potential realization of the MDAF component as devised in the 3GPP management system architecture (3GPP TS 28.550) for the RAN domain. In this respect, as illustrated in Fig.2, the RANDAF would serve as a producer of RAN MDAS for the RAN OAM systems such as C-SON applications.

POTENTIAL RRM IMPROVEMENTS

Illustrative assessments of potential RRM improvements resulting from the extraction by the RANDAF/LDAF of RAN data analytics information and the exploitation of this information by the RAN nodes are provided in this section. The assessments take as a reference a measurements dataset collected in 31 cells of an LTE live network deployed in an urban area of a European city. The cells use 20 and 10 MHz channel bandwidths and are equipped with 2 transmission antennas. The dataset covers a period of 2 weeks with measurements taken every 15 minutes.

Case A. Prototype trajectories for enhanced handover

UE mobility patterns within a given region can be characterized in terms of *prototype trajectories*, which could form part of the mobility data analytics

information made available by the RANDAF. Following the methodology presented in [13], the detection and characterization of the prototype trajectories can be done at the RANDAF using UE location information available in a given area during a long period of time (obtained from e.g. GNSS coordinates reported by UEs via RRC signaling or from any other location functionality that might be in place), processing it to identify UE trajectories and then applying a clustering method (e.g. Self-Organizing Map, K-means) to extract the prototype trajectories. Each prototype trajectory is characterized by a vector of coordinates (i.e. the trajectory) and a percentage of hits reflecting its relevance.

Through the proposed A_{IS} interface, RAN nodes could inquire or be notified by the RANDAF about the existence of prototype trajectories within their coverage area and retrieve them. Once uploaded in a RAN node, prototype trajectories can be exploited for mobility prediction of individual UEs. In particular, by matching the location information reported by a UE with the prototype trajectories, it is possible for the serving RAN node to determine the likelihood that the UE is following a given prototype trajectory (see details in [13]). A high value of likelihood for a given trajectory means that future positions of the UE can be accurately predicted.

This mobility prediction capability can then be leveraged in multiple ways. For instance, handovers can be anticipated without waiting for the signal to drop below a threshold and ping pong effects can be mitigated. Also, in the case of high resource demanding UEs, a resource reservation procedure could be triggered ahead of time to ensure the availability of resources when the handover is eventually executed, avoiding service degradation. To illustrate this latter case in terms of achievable performance gains, let us consider a scenario with Guaranteed Bit Rate (GBR) services operated under strict admission control policies to avoid service degradation. Under such scenario, Fig. 3 shows the reduction of the dropping probability achieved by a handover algorithm that reserves the resources for GBR traffic of a UE heading a target cell T_p seconds in advance, compared to the case in which there is no such anticipation.

Results in Fig. 3 are given for two GBR session durations T_d (120s and 300s) and two anticipation times T_p (8s and 12s which, applying the methodology of [13] in a urban scenario with 30 km/h, appear as plausible values to anticipate handovers with likelihood of around 0.9). To determine the Physical Resource Block (PRB) requirements of GBR traffic, a data rate of 10 Mb/s is considered along with the spectral efficiency characteristics of the cells included in the measurements dataset. In this respect, estimations are provided for three individual cells #1, #2, #3 representative of, respectively, high, medium and low average spectral efficiency values (shown in the horizontal axis) as well as when considering the average spectral efficiency of all the cells in the dataset. Results show that a substantial dropping probability reduction is obtained (e.g. up to 36% on average for all cells when $T_p=12s$ and $T_d=300s$). Besides, the reduction increases with the anticipation time T_p and for smaller session durations, because this facilitates

the successful execution of the resource reservation before the handover.

Case B. UE radio resource utilization patterns for enhanced admission control

Seeking the characterization of subscriber/UE behaviour, the RANDAF can serve as the entity responsible to keep track and extract patterns related to radio resource utilization, mobility and communications activity for individual UEs within the RAN. Such patterns can be stored in the RANDAF, which would act as a central repository for the RAN nodes, given that those do not retain UE-specific data of UEs without an active RRC connection. In this way, upon a new UE connection activation attempt, the serving RAN node could retrieve the corresponding UE pattern from the RANDAF and use it as an additional input for RRM decision-making.

This can be especially relevant for UEs exhibiting stationary or nomadic mobility patterns, that is, connecting to the RAN in most cases from one or a limited set of locations. For these UEs, a detailed characterization of the past utilization of the RAN can be very helpful to guide future RRM decisions given the less variability of the propagation conditions along successive connections. The RAN could acquire the knowledge that a given UE is nomadic or stationary in different ways, such as based on the processing of the location information at the RANDAF (as considered in Case A) or based on subscription information provided from 5GC.

In this context, let us consider that the UE radio resource utilization pattern produced by the RANDAF includes a detailed characterization of the spectral efficiency distribution experienced by the UE when connecting from specific locations. This characterization can be derived from the CSI measurements collected at the DU for scheduling purposes. CSI measurements include so called Channel Quality Indicator (CQI) values, which indicate the most suitable modulation and coding scheme (MCS) supported by a UE given its propagation and interference conditions and thus they can be mapped to a spectral efficiency value. CQI measurements are typically reported by the UEs on a short-term time scale (milliseconds) to feed the scheduling process. On this basis, an LDAF capability at the DU could be used to temporarily store the reported CQI measurements and derive the distribution of the spectral efficiency on a per cell and per UE basis. The resulting characterization could later be retrieved by the RANDAF to compute longer-term statistics and to keep track of the characterization of specific individual UEs showing stationary or nomadic mobility conditions.

To illustrate how the spectral efficiency per UE can be exploited, let us consider the same scenario as in Case A for GBR services and assume now the presence of stationary UEs such as Fixed Wireless Access (FWA) subscribers. If no prior information on a particular UE is available at the RAN node when a GBR service is requested, admission control can only be carried out based on an estimation of the average cell spectral efficiency. However, due to the high variability of spectral efficiency, this may differ considerably from the actual conditions, leading to bad

estimations of the amount of required radio resources. Therefore, having a more accurate estimation of the spectral efficiency is key to improve admission control.

To illustrate the potential improvement, let us consider the presence of a FWA UE in each cell of the scenario placed at a location where the spectral efficiency equals the 95th percentile of the cell spectral efficiency. Following the admission control algorithm presented in [14], Fig. 4 depicts the average capacity gain achievable by exploiting the spectral efficiency per UE to estimate the PRB requirements of FWA UEs with respect to the case where the estimation is done using the average spectral efficiency. The capacity here is accounted as the maximum offered load of GBR traffic that can be supported for a blocking probability not higher than 2%. Results consider two GBR values, 5 Mb/s and 10 Mb/s, and two fractions of PRBs in the cell devoted to GBR traffic, 30% and 50%. Results for the same example cells of Fig. 3 as well as the total average gain considering the 31 cells of the scenario are presented. It is observed that, for 10 Mb/s and 30% of PRBs for GBR traffic, the average capacity gain across the scenario is 36%, while some specific cells can reach gains up to 80%. It is also noticed that the gain increases with the GBR value and when reducing the fraction of GBR resources, because in these situations the resource consumption is larger, so the accuracy in the resource estimation at admission control becomes more critical.

CONCLUDING REMARKS

The support of data analytics services within the RAN, together with the conception of smarter, data-driven RRM algorithms, arise as a promising area of research and development.

In this regard, this paper has discussed the types and sources of measurements as well as the possible data analytics information categories to enable data-driven RRM algorithms. As a central contribution of the paper, a functional architectural framework has been proposed as an attempt to progress towards the specification of a common reference based on the definition of a service-based control interface for the exchange of RAN data analytics.

Further work on the analysis of data-driven RRM use cases and their impact on the architecture lays ahead, with challenging open issues such as the extraction and exploitation of data analytics information at much shorter time scales (e.g. online learning of short-scale traffic and mobility patterns forecast [15]) and the achievable accuracy of the data analytics information.

ACKNOWLEDGEMENTS

This work has been supported by the Spanish Research Council and FEDER funds under SONAR 5G grant (ref. TEC2017-82651-R).

REFERENCES

- [1] Q. Yan, W. Chen, H.V. Poor, "Big Data Driven Wireless Communications: A Human-in-the-Loop Pushing Technique for 5G Systems", IEEE Wireless Communications, February, 2018.
- [2] 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.

- [3] 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.
- [4] 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.
- [5] S. Holmbacka, et al. "Alarm Prediction in LTE Networks", Int. Conf. on Telecomms. (ICT), 2018.
- [6] P. Chiu, J. Reunanen, R. Luostari, H. Holma, "Big Data Analytics for 4.9G and 5G Mobile Network Optimization", IEEE VTC Spring, 2017.
- [7] 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
- [8] 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.
- [9] E. Pateromichelakis et al., "End-to-End Data Analytics Framework for 5G Architecture", IEEE Access, April 2019.
- [10] F. D. Calabrese et al., "Learning Radio Resource Management in RANs: Framework, Opportunities, and Challenges," in IEEE Communications Magazine, vol. 56, no. 9, pp. 138-145, Sept. 2018.
- [11] Study on RAN-centric data collection and utilization for LTE and NR, document RP-182105, 3GPP, September 2018.
- [12] Open RAN Alliance ORAN-WG2.A1.AP-v01.00, "A1 interface: Application Protocol", October, 2019.
- [13] J. Sánchez-González, J. Pérez-Romero, R. Agustí, O. Sallent, "On Learning Mobility Patterns in Cellular Networks", 12th IFIP Int. Conf. on Artificial Intelligence Applications and Innovations, Thessaloniki, Greece, September, 2016.
- [14] O. Sallent, J. Pérez-Romero, R. Ferrús, R. Agustí, "Data Analytics in the 5G Radio Access Network and its Applicability to Fixed Wireless Access", IEEE 89th Vehicular Technology Conference, Kuala Lumpur, Malaysia, April-May, 2019.
- [15] A. Azari, M. Ozger, C. Cavdar, "Risk-Aware Resource Allocation for URLLC: Challenges and Strategies with Machine Learning", IEEE Communications Magazine, March, 2019.

BIOGRAPHIES

Ramon Ferrús is an associate professor at Universitat Politècnica de Catalunya (UPC). His research interests include system design, network and service management in wireless communications. He is co-author of two books and over 120 papers published in peer-reviewed journals and conferences.

Oriol Sallent is a full professor at UPC. He has participated in a wide range of European projects and contributed to standardisation bodies such as 3GPP, IEEE and ETSI. He has published 200+ papers mostly in IEEE journals and conferences. His research interests include self-organising networks, radio network optimization and QoS provisioning in heterogeneous wireless networks.

Jordi Pérez-Romero is a full professor at UPC. He has been working in the field of radio resource management and network optimization for wireless systems. He has been involved in different European Projects and in projects for private companies. He has published more than 200 papers in international journals and conferences.

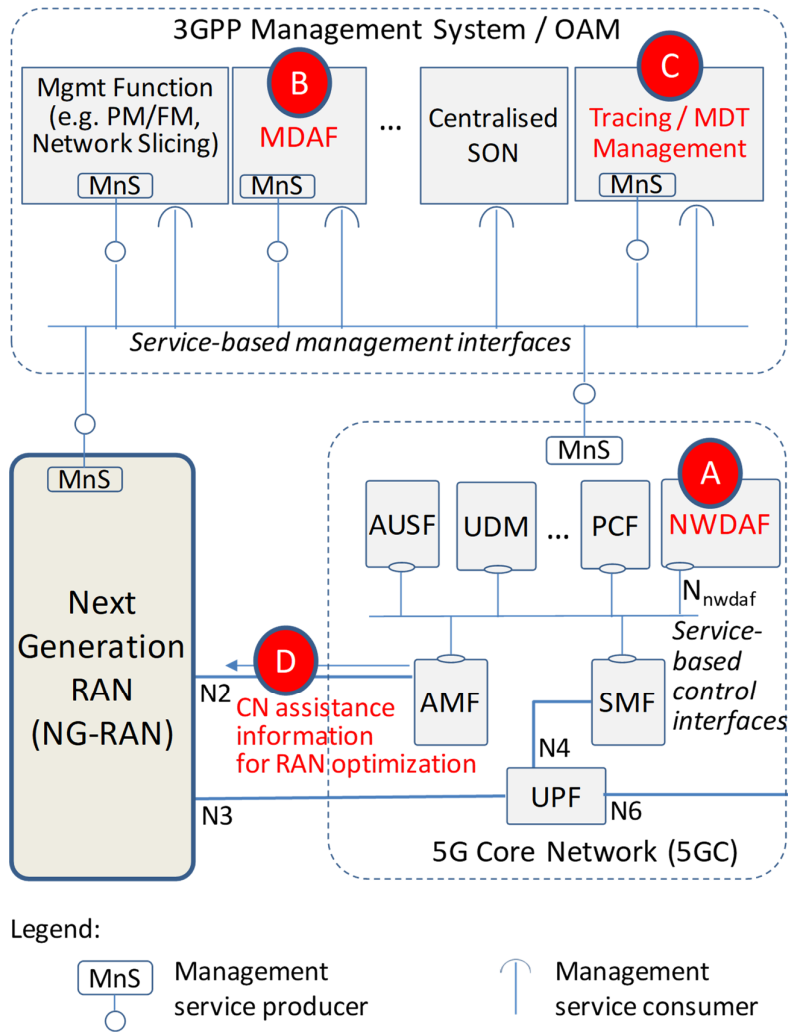


Fig. 1.- Functional framework for management and network data analytics services support in 3GPP 5G systems

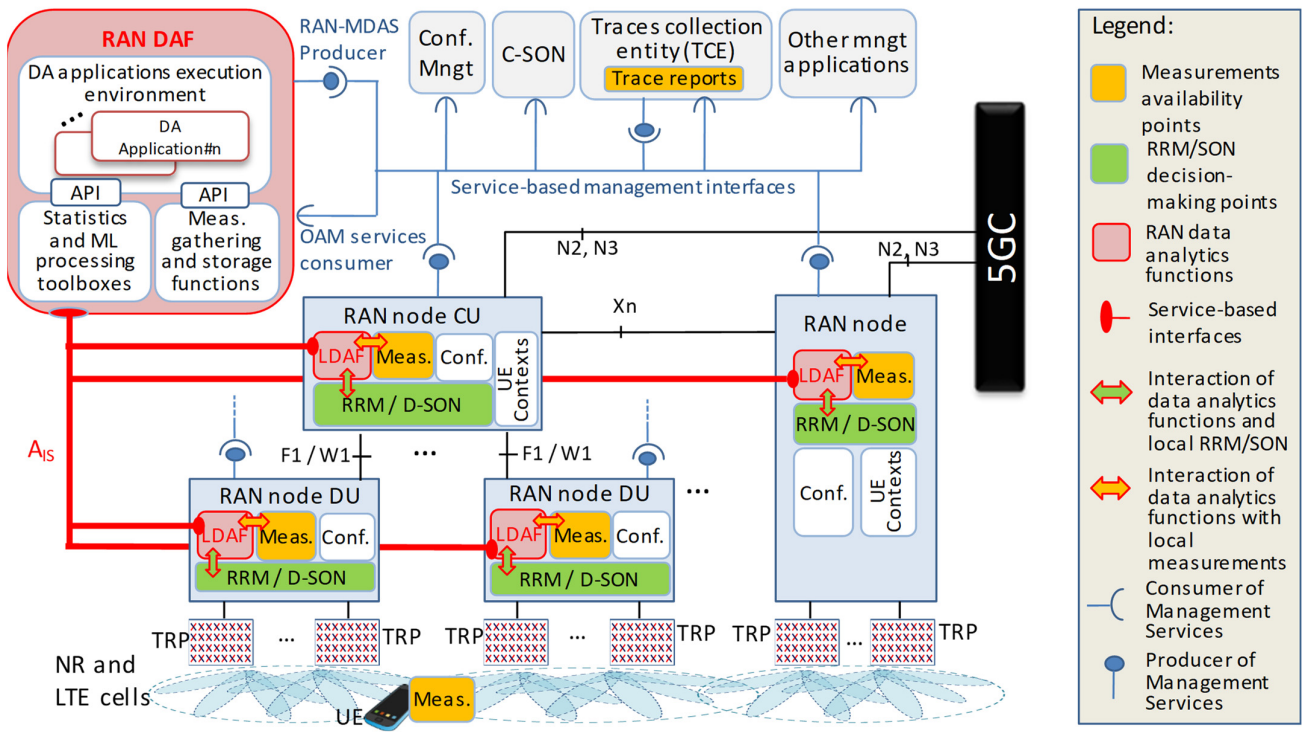


Fig. 2.- Proposed functional framework for the exploitation of data analytics services within the NG-RAN architecture

Table 1. Standardised measurements available within NG-RAN nodes for 5G NR

Nodes	Available Measurements
gNB-CU (Related 3GPP specs: TS 28.552, TS 38.215)	-RRC connection counters -Protocol Data Unit (PDU) session setup counters -Mobility management counters -Data Radio Bearer (DRB) setup management counters -UE context release counters -Packet Delay/Loss/Drop rates -Data volume measurements -Reference Signal Received Power (RSRP) -Reference Signal Received Quality (RSRQ) -Signal-to-Interference and Noise Ratio (SINR) -Timing measurements -GNSS signal reception measurements.
gNB-DU (Related 3GPP specs: TS 28.552, TS 38.214 and TS 38.321)	-Radio resource utilization -UE throughput -Transport Block (TB) related measurements -Packet Delay/Loss/Drop Rates -IP latency measurements -RF measurements statistics (e.g. MCS distribution) -Channel State Information (CSI) -Buffer status reports -Power headroom reports
TCE (Related 3GPP specs: TS 32.423)	-Signalling logs of protocol operation over the NG-RAN interfaces (e.g. Uu, F1, NG, Xn) -UE MDT measurements, including diverse measurements such as RSRP, RSRQ, used QoS class identifiers, UL/DL traffic volume, UL/DL throughput and location of the UE.

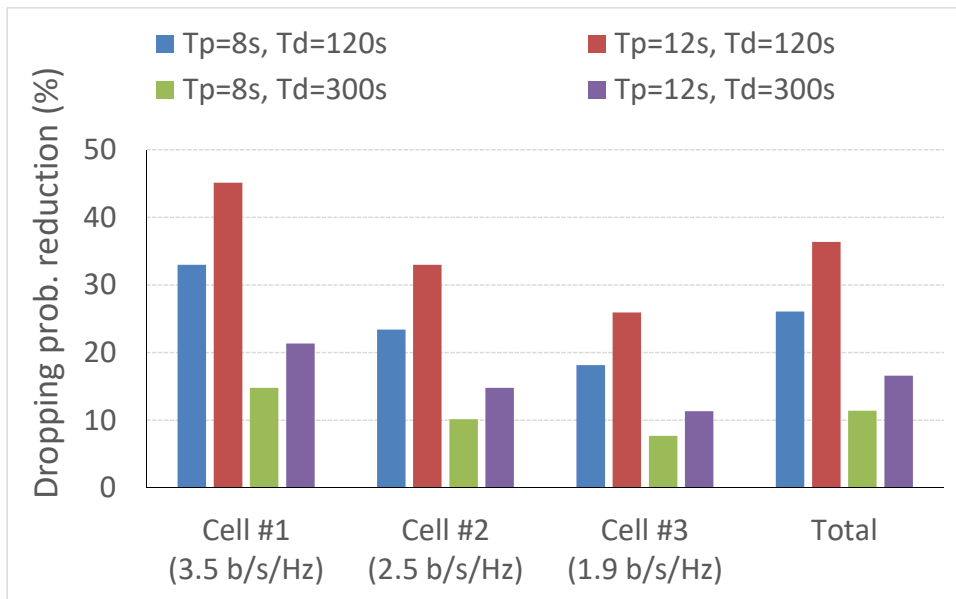


Fig. 3.- Dropping probability reduction achieved by handover prediction

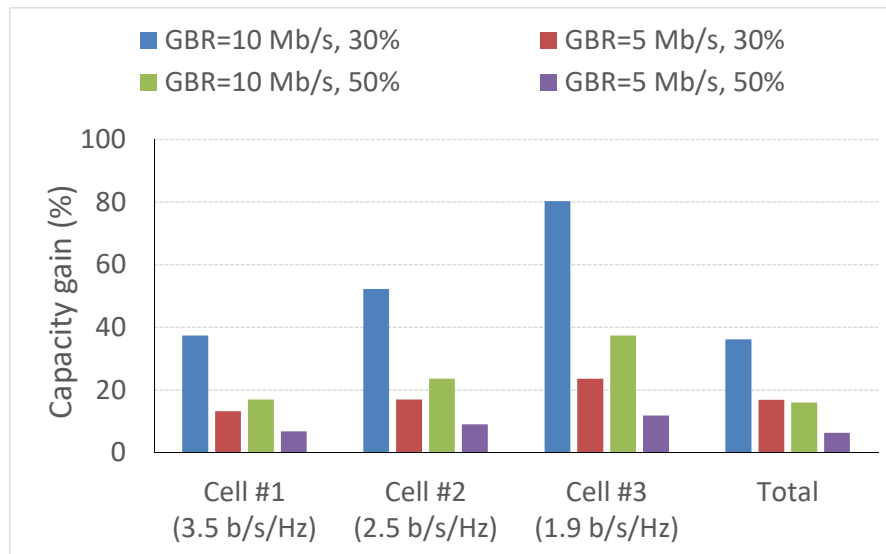


Fig. 4.- Capacity improvement by exploiting the spectral efficiency per UE in the admission control