

# Intelligence and Learning in O-RAN for Data-Driven NextG Cellular Networks

Leonardo Bonati, Salvatore D'Oro, Michele Polese, Stefano Basagni, and Tommaso Melodia

## ABSTRACT

Next generation (NextG) cellular networks will be natively cloud-based and built on programmable, virtualized, and disaggregated architectures. The separation of control functions from the hardware fabric and the introduction of standardized control interfaces will enable the definition of custom closed-control loops, which will ultimately enable embedded intelligence and real-time analytics, thus effectively realizing the vision of autonomous and self-optimizing networks. This article explores the disaggregated network architecture proposed by the O-RAN Alliance as a key enabler of NextG networks. Within this architectural context, we discuss the potential, the challenges, and the limitations of data-driven optimization approaches to network control over different timescales. We also present the first large-scale integration of O-RAN-compliant software components with an open source full-stack software cellular network. Experiments conducted on Colosseum, the world's largest wireless network emulator, demonstrate closed-loop integration of real-time analytics and control through deep reinforcement learning agents. We also show the feasibility of radio access network (RAN) control through xApps running on the near-real-time RAN intelligent controller to optimize the scheduling policies of coexisting network slices, leveraging the O-RAN open interfaces to collect data at the edge of the network.

## INTRODUCTION

The fifth and sixth generations (5G, 6G) of cellular networks will undoubtedly accelerate the transition from inflexible and monolithic networks to agile, disaggregated architectures based on *softwarization* and *virtualization*, as well as on openness and re-programmability of network components [1]. These novel architectures are expected to become enablers of new functionalities, including the ability to:

- Provide on-demand virtual network slices that, albeit sharing the same physical infrastructure, are tailored to different mobile virtual network operators, network services, and runtime traffic requirements
- Split network functions across multiple software and hardware components, possibly provided by multiple vendors
- Capture and expose key performance indicators (KPIs) and network analytics through open interfaces that are not available in old architectures

- Control the entire network physical infrastructure in real time via third-party software applications and open interfaces

## DISAGGREGATION AND PROGRAMMABILITY IN O-RAN

The O-RAN Alliance — a consortium of industry and academic institutions — is working toward realizing the vision of next generation (NextG) cellular networks, where telecom operators use standardized interfaces to control multi-vendor infrastructures and deliver high-performance services to their subscribers [2]. To achieve this goal, the Alliance proposes an architectural innovation based on two core principles. First, O-RAN embraces and promotes the 3rd Generation Partnership Project (3GPP) functional split, where base station (BS) functionalities are virtualized as network functions and divided across multiple network nodes: central unit (CU), distributed unit (DU), and radio unit (RU) [1]. This facilitates the instantiation and execution of diverse networking processes at different points of the network. Specifically, CUs implement functionalities at the higher layers of the protocol stack operating over larger timescales, while DUs handle time-critical operations at the lower layers. Finally, the RUs manage radio frequency (RF) components and lower physical (PHY) layer parts.

The second core innovation — which is likely to be even more impactful — is the radio access network (RAN) intelligent controller (RIC), a new architectural component that provides a centralized abstraction of the network, allowing operators to implement and deploy custom control plane functions. In both its *non-* and *near-real-time* versions, the RIC facilitates RAN optimization through closed-control loops (i.e., autonomous action and feedback loops between RAN components and their controllers). O-RAN envisions different loops operating at timescales that range from 1 ms (e.g., for real-time control of transmission strategies) to thousands of milliseconds (e.g., for network slicing and traffic forecasting). For instance, the non-real-time RIC performs operations with a time granularity higher than 1 s, such as training of artificial intelligence (AI) and machine learning (ML) models. The near-real-time RIC instead handles procedures at timescales above 10 ms, hosts third-party applications (xApps) that communicate with the CU/DU through standardized open interfaces, and implements intelligence in the RAN through data-driven control loops.

Figure 1 illustrates one of the possible disaggregated deployments specified by O-RAN,

The authors explore the disaggregated network architecture proposed by the O-RAN Alliance as a key enabler of NextG networks. Within this architectural context, they discuss the potential, the challenges, and the limitations of data-driven optimization approaches to network control over different timescales.

This work was partially supported by the U.S. National Science Foundation under Grant CNS-1923789 and the U.S. Office of Naval Research under Grant N00014-20-1-2132.

The authors are with the Institute for the Wireless Internet of Things, Northeastern University.

Digital Object Identifier:  
10.1109/MCOM.101.2001120

While the O-RAN architectural vision is gaining momentum among researchers, the challenges of implementing it for data-driven, open, programmable and virtualized NextG networks are still largely to be dealt with. Important architectural questions are yet to be answered.

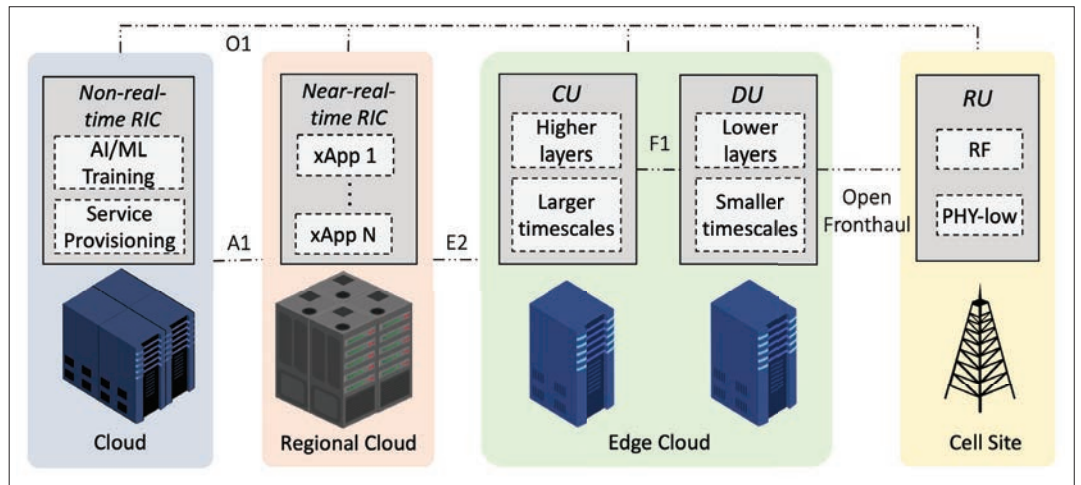


FIGURE 1. O-RAN: An example of disaggregated deployment.

where different network components are connected by open interfaces.

In this deployment (“Scenario B,” deemed the most common [1]), the RICs are deployed in the cloud. They interact with each other via the A1 and O1 interfaces, and control specific parameters of the RAN defined through the so-called service models (SMs). The CU and DU are deployed at the network edge, interconnected through the F1 interface and controlled by the near-real-time RIC via the E2 interface [3]. The RU is located at the operator cell site, and controlled by the DU through the Open Fronthaul interface. Finally, CU, DU, and RU are connected to the non-real-time RIC through the O1 interface for periodic reporting. Other deployment options allow instantiating RICs and CUs in the edge or regional cloud (Scenarios A and C–F, respectively); the DU in the edge cloud (A–D); and the RU at the operator cell site (A–D) or in the cloud cell site, possibly co-located with the DU (E, F) [1].

### CONTRIBUTIONS

While the O-RAN architectural vision is gaining momentum among researchers, the challenges of implementing it for data-driven, open, programmable, and virtualized NextG networks are still largely to be dealt with. Important architectural questions have yet to be answered, including:

- The exact functionalities and parameters to be controlled by each network component
- Where to place network intelligence
- How to validate and train data-driven control loop solutions
- How AI agents can access data and analytics from the RAN while minimizing the overhead of moving them from the RAN to the storage and inference locations

To answer these questions, we provide the following contributions:

- We discuss how data-driven, closed-control loop solutions can be implemented in NextG RANs. We focus on the *opportunities* offered by the O-RAN architecture, including functional split and open interfaces, and on their role in advancing intelligent and programmable networks.
- Different from prior work [1, 4], we investigate the *limitations* of the current O-RAN

specifications and the *challenges* associated with deploying data-driven policies at different nodes of the RAN.

- We discuss how large-scale experimental testbeds will play a key role by providing researchers with heterogeneous and large datasets, critical to the success of data-driven solutions for cellular networks. We focus on the three PAWR platforms (i.e., POWDER [5], COSMOS [6], and AERPAW [7]), and on Colosseum and Arena [8], which can all be used to generate massive datasets under a variety of network configurations and RF conditions.
- We provide the first demonstration of an O-RAN data-driven control loop in a *large-scale experimental testbed using open source, programmable RAN and RIC components*. We deploy O-RAN on the Colosseum network emulator and use it to control multiple network slices instantiated on 4 software-defined radio (SDR) BSs serving 40 SDR user equipments (UEs).
- We develop a set of deep reinforcement learning (DRL) agents as RIC xApps to optimize key performance metrics for different network slices through *data-driven closed-control loops*. Experimental results show that our DRL approach outperforms other control strategies, improving spectral efficiency by up to 20 percent and reducing buffer occupancy by up to 37 percent. We released the DRL agents and the 7 GB dataset used to train them.<sup>1</sup>

The remainder of this article is organized as follows. We first discuss how intelligent control schemes can be embedded in the O-RAN architecture. We then present how experimental testbeds can foster the development of data-driven solutions. Finally, we present our experimental evaluation and draw our conclusions.

### INTELLIGENT WIRELESS ARCHITECTURES

Openness, programmability, and disaggregation are key enablers of data-driven applications. However, they are only the first step toward the seamless integration of AI- and ML-based control loops in cellular networks. Typically, data-driven approaches involve several steps, ranging from

<sup>1</sup> <https://git.io/JiFPR>

Control and learning objective	Scale	Input data	Timescale	Architecture	Challenges and limitations
Policies, models, slicing	> 1000 devices	Infrastructure-level KPIs	Non-real-time > 1 s		Orchestration of very many near-real-time RICs and CUs/DUs/RUs
User Session Management e.g., load balancing, handover	> 100 devices	CU-level KPIs e.g., number of sessions, PDCP traffic	Near-real-time 10-1000 ms		Process streams from multiple CUs and sessions
Medium Access Management e.g., scheduling policy, RAN slicing	> 100 devices	MAC-level KPIs e.g., PRB utilization, buffering	Near-real-time 10-1000 ms		Operate at small time scales, make decisions involving several DUs/UEs
Radio Management e.g., resource scheduling, beamforming	~10 devices	MAC/PHY-level KPIs e.g., PRB utilization, channel estimation	Real-time < 10 ms		Deployment of AI/ML models at the DU is not supported
Device DL/UL Management e.g., modulation, interference, blockage detection	1 device	I/Q samples	Real-time < 1 ms		Require device- and/or RU-level standardization

FIGURE 2. Learning-based closed-control loops in an O-RAN architecture.

data collection and processing, to training, model deployment, and closed-loop control and testing.

This section illustrates how O-RAN is steering 5G deployments to bring intelligence to the network by defining a practical architecture for the swift execution of data-driven operations, and discusses extensions to control procedures not currently considered by O-RAN.

**Data Handling and Training Procedures.** The effectiveness of data-driven approaches heavily depends on how data is handled, starting from data collection and aggregation at the RAN (where data is generated) to the point where it is processed for model training and inference. However, collecting and moving large amounts of data might result in significant overhead and latency costs. Hence, data-driven architectures must cope with trade-offs between centralized approaches — providing a comprehensive view of the state of the network at the cost of overhead and latency — and distributed ones — operating at the edge only, gathering data from a small number of sources while enjoying low latency [9].

In this context, the O-RAN ML specifications introduce standardized interfaces (e.g., O1) to collect and distribute data across the entire infrastructure as well as operational guidelines for the deployment of ML and AI solutions in the network [10]. These include practical considerations on how, where, and when models can be trained, tested, and eventually deployed in the network. First, AI/ML models are made available to operators via a marketplace system similar to that of the well-established network function virtualization (NFV) management and orchestration (MANO) architecture, where models are stored in a catalog together with details on their control objectives, required resources, and expected inputs and outputs. Second, data-driven solutions must be trained and validated offline to avoid causing inefficiencies — or even outages — to the RAN. Indeed, since AI/ML techniques usually rely on a randomized initialization, O-RAN requires all ML models to be trained and validated offline before their deployment [10]. As we discuss next, albeit shielding the network from unwanted behavior, this requirement also limits the effectiveness of such approaches, especially the online ones. Online AI/ML techniques could still be used in O-RAN-compliant architectures by allowing models to be trained with offline data in the non-real-time RIC, and then perform online learning in the

near-real-time RIC. The smaller timescale of the control loops of the latter would in fact allow the online training pipeline to be fed with data collected in real time.

**Control Loops.** Figure 2 portrays how intelligence can be embedded at different layers and entities of a disaggregated cellular network together with the challenges and limitations of doing so. Each closed-control loop optimizes RAN parameters and operations by running at different timescales, with different number of UEs, and using different sources for the input data. The O-RAN Alliance is also looking into how to standardize the data-driven workflows for these control loops. As of this writing, O-RAN only considers non- and near-real-time loops, while real-time loops are left for future studies. Figure 2 also depicts the additional inference timescale below 1 ms to process raw I/Q samples and perform AI-driven PHY layer tasks, currently not part of O-RAN as it would require device- and/or RU-level standardization.

To better highlight the potential and limitations of the approach proposed by O-RAN, in the following we analyze each control loop individually, highlighting the role of each network component. Finally, we discuss how the current O-RAN architecture can be extended to realize the control loops and applications illustrated in Fig. 2.

### NON-REAL-TIME CONTROL LOOP

The O-RAN Alliance defines *non-real-time* any control loop that operates on a *timescale of at least 1 s*. As shown in Fig. 2, this involves coordination between the *non-real-time* and *near-real-time* RIC through the A1 interface. This control loop manages the orchestration of resources at the infrastructure level, making decisions and applying policies that impact thousands of devices. These actions can be performed using data-driven optimization algorithms processing data from multiple sources, and inference models deployed on the non-real-time RIC itself.

Practical examples of non-real-time data-driven control include instantiating and orchestrating network slices, as well as selecting which pre-trained inference models in the catalog should be deployed to accomplish operator intents, and deciding in which near-real-time RIC these models should be executed. Said decisions can be made according to a variety of factors, ranging from computational resources and data availability to

Data-driven approaches are aimed at autonomously managing the network requiring little to no human intervention. Training and testing algorithms and data-driven closed-control loop policies require large amounts of data gathered in diverse scenarios, with varying traffic patterns, requirements, and user behaviors so that the resulting policy is effective when deployed in real networks.

minimum performance requirements to comply with service level agreements. Moreover, since the non-real-time RIC is endowed with service management and orchestration capabilities, this control loop can also handle the association between the near-real-time RIC and the DUs/CUs. This is particularly useful in virtualized systems where DUs and CUs are dynamically instantiated on demand to match the requests and load of the RAN. However, non-real-time loops are challenging to actuate in practice because of the very many interactions among the non-real-time RIC and the network elements, which require tight coordination, data collection, and orchestration capabilities.

### NEAR-REAL-TIME CONTROL LOOPS

*Near-real-time* control loops operate on a time-scale between 10 ms and 1 s. As shown in Fig. 2, they run between the near-real-time RIC and two components of the next generation node bases (gNBs): the CU and the DU. Because one near-real-time RIC is associated with multiple gNBs, these control loops can make decisions affecting up to thousands of UEs, using user-session aggregated data and medium access control (MAC)/PHY layer KPIs. ML-based algorithms are implemented as external applications (i.e., xApps) and are deployed on the near-real-time RIC to deliver specific services such as inference, classification, and prediction pipelines to optimize the per-user quality of experience, controlling load balancing and handover processes, or the scheduling and beamforming design. Challenges of near-real-time control loops include the need to promptly make decisions in a matter of tens or hundreds of milliseconds for each of the several CUs and DUs controlled by the RIC.

### REAL-TIME CONTROL LOOPS

A crucial component of the operations of a cellular network involves actions at a *sub-10 ms, or even sub-millisecond, timescale*. In O-RAN, these operations are labeled *real-time control loops*, and mainly concern interactions between elements in the DU. Control loops at a similar timescale could also be envisioned to operate between the DU and the RU, or at the UEs. However, as deploying ML solutions at the DU is not currently supported, these loops are left for future extensions of the O-RAN specifications.

Finally, data-driven approaches at the lower layers of the protocol stack or at the device (i.e., involving sub-millisecond timescales) are extremely powerful and can be used for data-driven scheduling decisions [11] and for feedback-less detection of PHY layer parameters (e.g., modulation and coding scheme, and interference recognition) [12]. Overall, the fact that device-/RU-level standardization is required for sub-millisecond loops makes it very challenging to realize them in practice, thus limiting their applicability.

### OPEN WIRELESS DATA FACTORIES

Data-driven approaches are aimed at autonomously managing the network requiring little to no human intervention. Training and testing algorithms and data-driven closed-control loop policies require large amounts of data gathered in diverse scenarios, with varying traffic patterns,

requirements, and user behaviors so that the resulting policy is effective when deployed in real networks.

Access to the massive amounts of data needed for training, however, is usually a privilege that only telecom operators enjoy. Due to privacy and competition concerns, operators seldom share such data openly with the research community. As a consequence, researchers and practitioners are often constrained to rely on datasets collected in small laboratory setups, which seldom capture the variety and scale of real cellular deployments. In the context of intelligent networking for NextG cellular systems, large-scale wireless testbeds are needed for developing, training, and testing new data-driven solutions, serving as *open wireless data factories* for the community. Such open platforms would facilitate massive data collection in realistic and diverse wireless deployments [1].

The city-scale platforms of the U.S. National Science Foundation PAWR program promise to be a valuable tool to provide the community with the desired diversity of scenarios and scale. The program currently supports three open testbeds representative of a variety of wireless use cases, ranging from state-of-the-art SDRs and massive multiple-input multiple-output (MIMO) communications (POWDER, in Salt Lake City, Utah [5]), to ultra-high-capacity and low-latency wireless networks (COSMOS, in New York City [6]), and to aerial wireless communications (AER-PAW, in the Research Triangle of North Carolina [7]). All three platforms provide users with data generation and analysis tools [1]. Arena is a SDR ceiling testbed that allows study of MIMO, cellular, and Internet of Things (IoT) applications with up to 64 antennas deployed in an  $8 \times 8$  grid in an office space [8].

Another instrument for wireless research at scale is Colosseum, the world's largest wireless network emulator with hardware in the loop. Colosseum includes 128 compute nodes, called standard radio nodes (SRNs), equipped with USRP X310 SDRs that can be used to run generic protocol stacks. These are connected in a mesh topology through 128 additional USRPs X310 of the massive channel emulator (MCHEM) for emulating realistic RF scenarios. The wireless channel between each pair of devices is modeled through complex-valued RF filter taps. In this way, scenarios are able to capture effects such as path loss, multi-path, and fading as if the SDRs were operating in a real RF environment. Colosseum is also equipped with an edge data center, with 900 TB of storage and the capability of processing RF data at a rate of 52 TB/s, enabling massive data collection and testing of ML algorithms on heterogeneous networks.

### USE CASE: SCHEDULING CONTROL IN SLICED 5G NETWORKS THROUGH THE O-RAN RIC

This section showcases an example of a data-driven closed-loop control implemented using the O-RAN Software Community near-real-time RIC and an open cellular stack on Colosseum (Fig. 3). We demonstrate the feasibility of a closed-control loop where DRL agents running in xApps on the near-real-time RIC select the best-performing scheduling policy for each RAN slice.



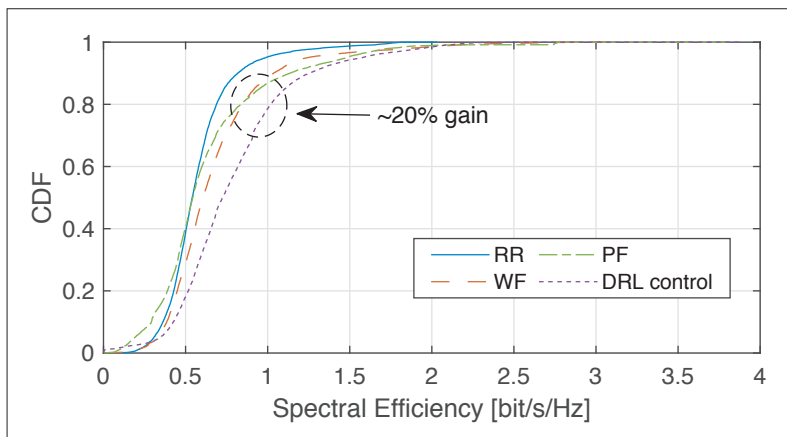


FIGURE 4. Downlink spectral efficiency of the URLLC slice for different scheduling policies and with DRL control.

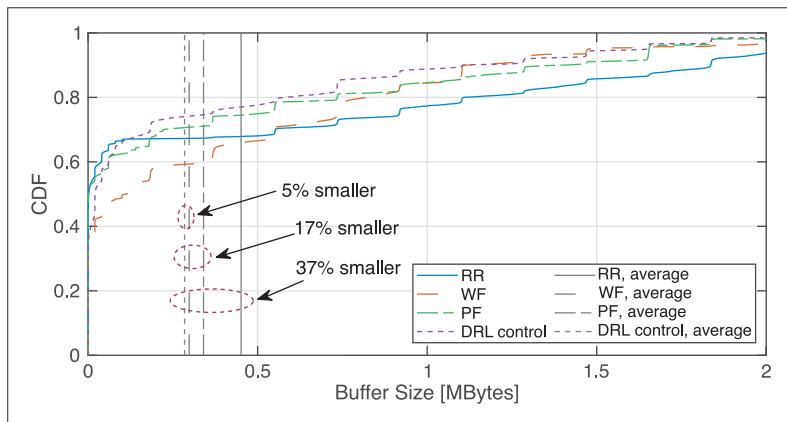


FIGURE 5. Downlink buffer size of the URLLC slice for different scheduling policies and with DRL control.

of 89 hours of experiments on Colosseum. Each DRL agent has been trained via the proximal policy optimization (PPO) algorithm to manage a single slice for fine-grained and flexible control of the whole cellular network. Agents have been trained under network configurations obtained by varying the distance between BSs and UEs and the mobility of the UEs. Testing has been performed in the most challenging setup, which includes the random mobility of the UEs. Although the training is performed with the same topology configuration, we notice that our agents are topology-independent, as each of them controls a single slice for a given BS. Specifically, agents process the performance metrics received by the BS they are controlling — which possibly expresses the performance of several UEs — through an encoder. This allows them to cast the dimensionality of the input data to a fixed size and to process it regardless of the number of active UEs of the slice. As a consequence, the DRL agents do not need to be aware of the number of UEs and BSs in the network, which makes our approach general and scalable. Through the RIC Indication messages sent via the O-RAN E2 interface (Fig. 3), the agent is fed real-time performance measurements of the slice it controls. These messages generate an overhead of 72 B/UE. Data goes through an encoder for dimensionality reduction and is then used by the agent to identify the state of the system. The agent uses a fully connected neural network with 5 layers and 30 neurons each to determine the best scheduling policy for the corresponding slice.

This policy is then signaled to the corresponding BS through RIC Control messages sent via the E2 interface. The reward of the agents depends on the specific slice and the corresponding KPI requirements. Specifically, eMBB and MTC agents have been trained to maximize the throughput of UEs; the URLLC agent has been trained to minimize latency by allocating resources (i.e., PRBs) as quickly as possible. To comply with O-RAN directives, we have trained the DRL agents offline in the non-real-time RIC, which also performs the initial data collection and deploys the model in the near-real-time RIC. We have then tested them on the emulated Colosseum scenario.

**Experimental Results.** Figure 4 shows the cumulative distribution function (CDF) of the downlink spectral efficiency of the eMBB slice. We compare the performance of the network when DRL agents dynamically select the best scheduling strategy among RR, PF, and WF against the case where scheduling strategies are fixed over time. Our results clearly indicate that data-driven optimization outperforms fixed policies by delivering gains in spectral efficiency that are up to 20 percent higher than that of the best performing static policy. This is due to the fact that eMBB traffic requires high data rates, and DRL agents are capable of dynamically adapting scheduling decisions to the current network state and traffic demand.

Figure 5 shows the CDF of the downlink buffer size for the URLLC slice under different scheduling policies. Low buffer size indicates timely data delivery to requesting UEs; higher buffer size results in higher latency due to packets waiting in the queue. Results show that DRL agents serve the UEs faster than the static policies, resulting in lower latency. In particular, the average buffer size of the URLLC slice when using DRL control is 37, 5, and 17 percent smaller than that of the RR, WF, and PF scheduling policies, respectively. The DRL agent also significantly outperforms the WF policy between the 50th and 90th percentiles.

Figure 6 depicts how often DRL agents select specific scheduling policies as a function of the number of PRBs of each slice. The bigger the circle, the higher the probability of selecting a given policy. We observe that MTC and eMBB DRL agents select WF 99 percent of the time. Also, eMBB agents select RR 4 percent of the time when only a few PRBs are allocated to the slice. On the contrary, URLLC DRL agents are likely to select both PF and WF scheduling policies even when more PRBs are available. These results show that adapting control strategies to current network state and traffic requirements is essential to achieve remarkable performance improvements (Fig. 4 and Fig. 5). DRL agents dynamically select the best performing scheduling strategy based on available resources and current network state, providing performance gains simply unattainable with static scheduling policies.

## CONCLUSIONS

In this article we provide a path to and a demonstration of the feasibility of integrating closed-control loops in cellular networks. We first review key enablers, namely, virtualization, disaggregation, openness, and reprogrammability of NextG cellular networks using O-RAN as an exemplary technology. We then discuss which data-driven control

loops can be implemented, their timescale, and whether the current O-RAN architecture supports them. We finally show how large-scale experimental testbeds can be used to develop and validate data-driven algorithms by deploying a DRL-based O-RAN RIC on Colosseum. Our results show that using closed-control loops can provide a strong foundation toward the full realization of future generation, data-driven, autonomous, and self-optimizing cellular networks.

## REFERENCES

- [1] L. Bonati *et al.*, "Open, Programmable, and Virtualized 5G Networks: State-of-the-Art and the Road Ahead," *Computer Networks*, vol. 182, Dec. 2020, pp. 1–28.
- [2] O-RAN WG 1, "O-RAN Architecture Description – v2.00," Tech. Spec., 2020.
- [3] O-RAN WG 3, "O-RAN Near-Real-Time RAN Intelligent Controller Architecture & E2 General Aspects and Principles – v1.01," Tech. Spec., 2020.
- [4] S. Niknam *et al.*, "Intelligent O-RAN for Beyond 5G and 6G Wireless Networks," arXiv:2005.08374 [eess.SP], May 2020.
- [5] J. Breen *et al.*, "POWDER: Platform for Open Wireless Data-driven Experimental Research," *Proc. ACM WINTeCH*, London, U.K., Sept. 2020.
- [6] D. Raychaudhuri *et al.*, "Challenge: COSMOS: A City-Scale Programmable Testbed for Experimentation with Advanced Wireless," *Proc. ACM MobiCom*, London, U.K., Sept. 2020.
- [7] M. Sichertiu *et al.*, "AERPAW Emulation Overview," *Proc. ACM WINTeCH*, London, U.K., Sept. 2020.
- [8] L. Bertizzolo *et al.*, "Arena: A 64-Antenna SDR-Based Ceiling Grid Testing Platform for Sub-6 GHz 5G-and-Beyond Radio Spectrum Research," *Computer Networks*, vol. 181, Nov. 2020.
- [9] M. Polese *et al.*, "Machine Learning at the Edge: A Data-Driven Architecture with Applications to 5G Cellular Networks," *IEEE Trans. Mobile Computing*, June 2020, pp. 1–16.
- [10] O-RAN WG 2, "O-RAN AI/ML Workflow Description and Requirements – v1.01," Tech. Spec., 2020.
- [11] S. Chinchali *et al.*, "Cellular Network Traffic Scheduling with Deep Reinforcement Learning," *AAAI*, 2018, pp. 766–74.
- [12] T. O’Shea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," *Proc. AAAI*, vol. 3, no. 4, Oct. 2017, pp. 563–75.
- [13] L. Bonati *et al.*, "CellIOS: Zero-Touch Softwarized Open Cellular Networks," *Computer Networks*, vol. 180, Oct. 2020, pp. 1–13.
- [14] S. D’Oro, F. Restuccia, and T. Melodia, "The Slice Is Served: Enforcing Radio Access Network Slicing in Virtualized 5G Systems," *Proc. IEEE INFOCOM*, Paris, France, May 2019.
- [15] S. D’Oro *et al.*, "SI-EDGE: Network Slicing at the Edge," *Proc. ACM Mobihoc*, Virtual Event, Oct. 2020.

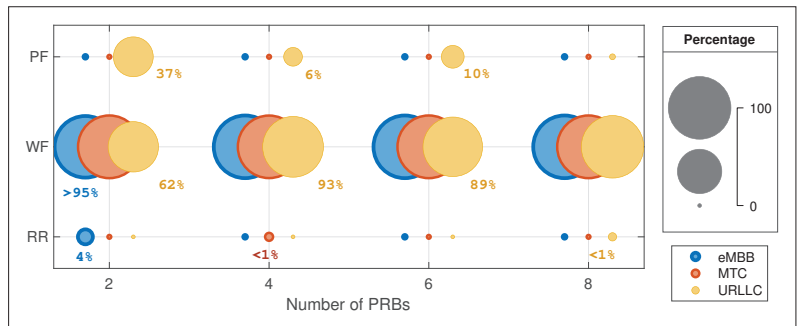


FIGURE 6. DRL action selection distribution vs. number of slice PRBs. Values > 99 percent (big circles) or < 0.5 percent (small circles) are omitted.

## BIOGRAPHIES

LEONARDO BONATI [S'19] (l.bonati@northeastern.edu) is a Ph.D. candidate at Northeastern University. He received his M.S. in telecommunication engineering from the University of Padova, Italy, in 2016. His research focuses on softwarized NextG systems.

SALVATORE D'ORO [M'17] (s.doro@northeastern.edu) is a research assistant professor with the Institute for the Wireless Internet of Things at Northeastern University. He received his Ph.D. from the University of Catania in 2015. He serves on the Technical Program Committee of IEEE INFOCOM. His research focuses on optimization and learning for NextG systems.

MICHELE POLESE [M'20] (m.polese@northeastern.edu) is a research scientist at Northeastern University. He obtained his Ph.D. from the University of Padova in 2020, where he was also a postdoctoral researcher and adjunct professor. His research focuses on architectures for wireless networks.

STEFANO BASAGNI [SM'06] (s.basagni@northeastern.edu) received a Ph.D. in electrical engineering from the University of Texas at Dallas in 2001 and a Ph.D. in computer science from the University of Milano, Italy, in 1998. He is with the Institute for the Wireless Internet of Things and a professor in the ECE Department at Northeastern University. His research concerns mobile networks and wireless communications systems.

TOMMASO MELODIA [F'18] (t.melodia@northeastern.edu) received a Ph.D. in electrical and computer engineering from the Georgia Institute of Technology in 2007. He is the William Lincoln Smith Professor at Northeastern University, the director of the Institute for the Wireless Internet of Things, and the director of research for the PAWR Project Office. His research focuses on wireless networked systems.