

# A Multi-Agent Deep Reinforcement Learning Approach for RAN Resource Allocation in O-RAN

Farhad Rezazadeh<sup>\*†</sup>, Lanfranco Zanzi<sup>‡</sup>, Francesco Devoti<sup>‡</sup>, Sergio Barrachina-Muñoz<sup>\*</sup>, Engin Zeydan<sup>\*</sup>,  
Xavier Costa-Pérez<sup>‡¶</sup>, and Josep Mangués-Bafalluy<sup>\*</sup>

<sup>\*</sup>Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain

<sup>†</sup>Universitat Politècnica de Catalunya (UPC), Barcelona, Spain

<sup>‡</sup>NEC Laboratories Europe, Heidelberg, Germany

<sup>¶</sup>i2CAT Foundation and ICREA, Barcelona, Spain

Contact Emails: {name.surname}@cttc.es, {name.surname}@neclab.eu

**Abstract**—Artificial intelligence (AI) and Machine Learning (ML) are considered as key enablers for realizing the full potential of fifth-generation (5G) and beyond mobile networks, particularly in the context of resource management and orchestration. In this demonstration, we consider a fully-fledged 5G mobile network and develop a multi-agent deep reinforcement learning (DRL) framework for RAN resource allocation. By leveraging local monitoring information generated by a shared gNodeB instance (gNB), each DRL agent aims to optimally allocate radio resources concerning service-specific traffic demands belonging to heterogeneous running services. We perform experiments on the deployed testbed in real-time, showing that DRL-based agents can allocate radio resources fairly while improving the overall efficiency of resource utilization and minimizing the risk of over provisioning.

## I. INTRODUCTION

The highly heterogeneous nature of 5G and beyond makes increasingly challenging to meet the performance requirements for emerging use cases. The strict reliability and latency requirements exacerbates the need for intelligent orchestration solutions able to provide resource and cost-efficient services. In this context, network slicing is a key enabler to underpin such challenging scenarios by deploying logically-isolated virtual network instances. Motivated by open architecture and interfaces for intelligent and agile RAN, O-RAN Alliance [1] provides flexibility, interoperable and open interface for virtualization and network innovation. Despite this revolutionary approach, it is still not clear how to efficiently support large network slicing scenarios. Therefore, we address this challenge by proposing a hierarchical architecture for network slice resource orchestration. In particular, given the variable spatio-temporal distribution of mobile traffic demands [2], we envision the dynamic setup of a network of local decision agents (DAs) as virtual software instances co-located within the Near-Real Time RAN Intelligent Controller (Near-RT RIC) premises. These instances are capable of accessing local RAN monitoring information and extracting local knowledge without the need for a centralized entity to make decisions based on the aggregated information.

Our framework, which was initially presented and validated by means of an exhaustive simulation campaign in [3], leverages a distributed learning mechanism and multiple

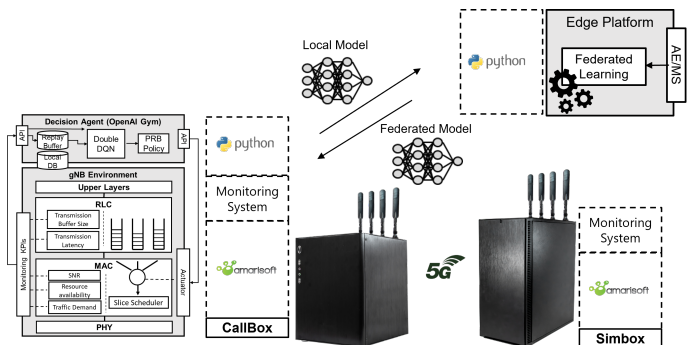


Fig. 1: Federated RAN resource allocation architecture overview.

decision agents that collaboratively specialize their decision policies onto real-time traffic demands, aided by a coordinated exchange of information to avoid the occurrence of conflicting situations. In our approach, we perform resource allocation in a distributed fashion directly at the edge of the network. This brings manifold advantages including *i*) performing radio resource allocation decisions leveraging the most up-to-date information, *ii*) the decision process does not involve a centralized controller, significantly decreasing the need for control information exchange and reducing the overhead towards the core of the network, and *iii*) by allowing information exchange among local DAs, we can establish federated learning schemes to further enrich the capabilities of the DAs. Namely, DAs not only rely on local observations to drive the learning process but can exploit the experience from other (statistically different) RAN nodes to speed up the convergence time and better generalize resource allocation decisions.

## II. TESTBED ARCHITECTURE

We propose a distributed architecture for RAN slice resource orchestration based on DRL, consisting of multiple AI-enabled decision agents that independently take local radio allocation decisions without the need for a centralized control entity. We design and implement the overall framework as shown in Fig. 1. The testbed includes:

**AMARI UE Simbox:** It is capable of simulating tens of

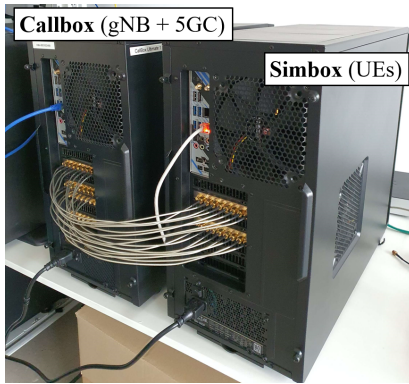


Fig. 2: Testbed deployment.

User equipments (UEs) sharing the same spectrum with different types of traffic within multiple cells. Each UE can be independently configured as a 5G NR device [4].

**AMARI Callbox:** provides Enhanced Packet Core (EPC)/5G Core (5GC) functionalities, including authentication of UEs. Moreover, it implements up to 3 gNodeBs cells enabling functional and performance testing. Thanks to its multi-cell configuration, it is also suitable for handover and reselection tests. The technical specifications of the gNB and the core can be found on the vendor’s website [5].

**Monitoring System:** Exploiting a dedicated API hosted by the gNB, it enables real-time monitoring of multiple KPIs, including the number of connected devices, bandwidth utilization, channel quality, etc., which are stored in a local database.

**Local Decision Agents:** collect and consume local monitoring information from the monitoring system to adjust radio resource allocation policies according to real-time traffic variations. Decision-making is supported by AI algorithms and DRL approaches.

**Federated Learning Layer:** It acts as an aggregation point for the local decision engines deployed at RAN. It collects locally trained (and therefore heterogeneous) decision models and combines them to gain global knowledge about the underlying infrastructure behavior to improve the generalization of the decision process in the agents.

The overall testbed is depicted in Fig. 2.

### III. OPERATIONAL PHASES

A simplified example of our demonstration is available online<sup>1</sup>. This video presents the network performance of different deployed slices from our monitoring dashboard based on Grafana<sup>2</sup>. The possibility of acquiring instantaneous and effective resource allocation in network slicing is precluded due to the heterogeneous and dynamic characteristics of traffic requests. This scenario leverages a Multi-DRL mechanism to alleviate this challenge. Each agent deployed over the Callbox is equipped with a double deep Q-network (DDQN) [3] that,

<sup>1</sup><https://nextcloud.neclab.eu/index.php/s/AwHqH7GYAixmaR>  
Access password: e2WmM9BiLE

<sup>2</sup><https://grafana.com/>



Fig. 3: Graphical User Interface.

after an initial exploration and training phase, achieves the right trade-off between optimal allocated radio resources and traffic demand accommodation, i.e., exploitation. To this aim, the agents receive partial observations from the running services or *slices* (e.g., channel quality, consumed resources, etc.) by interacting with each other and with the underlying physical environment. A reward is calculated as an incentive mechanism based on the actions of all agents, indicating how the agents ought to behave. The designed reward function guarantees adequate performances, impeding over provisioning and leading to fair resource allocation among the three slices.

Fig. 3 illustrates the slice performance of three agents and validates the presented framework by displaying different metrics such as average reward, allocation gap, and allocated radio resources. Furthermore, given the spatio-temporal variation of the traffic demand due to end-user mobility proper of realistic RAN deployments, it is not enough to leverage the geographical locations and related spatial proximity of the gNB to obtain a comprehensive view of traffic demands. Therefore, it is hard for a single AI agent to predict resources for the same slices distributed in different gNB. In this regard, we pursue an experimental approach, and due to the lack of multiple gNB equipments, we deployed three enhanced Mobile Broadband (eMBB) slices with different traffic shapes and supposed each slice is deployed in different gNB and target learning improvement assisted by a federated DRL. Federated Learning allows training ML models across multiple decentralized entities, each agent having access to a limited set of the overall data realization and statistics.

### REFERENCES

- [1] O-RAN Alliance, “O-RAN: Towards an Open and Smart RAN,” Tech. Rep., Oct. 2018.
- [2] D. Bega, M. Gramaglia, M. Fiore, A. Banchs, and X. Costa-Pérez, “AZTEC: Anticipatory Capacity Allocation for Zero-Touch Network Slicing,” in *Proc. IEEE Conf. Compu. Commun.*, Jul. 2020.
- [3] F. Rezazadeh, L. Zanzi, F. Devoti, H. Chergui, X. Costa-Pérez, and C. Verikoukis, “On the Specialization of FDRL Agents for Scalable and Distributed 6G RAN Slicing Orchestration,” *IEEE Transactions on Vehicular Technology*, pp. 1–15, 2022.
- [4] “AMARI UE Simbox Series,” <https://www.amarisoft.com/app/uploads/2021/12/AMARI-UE-Simbox-E-Series.pdf>, 2023, [Online; accessed 2023-01-05].
- [5] “AMARI Callbox Ultimate,” <https://www.amarisoft.com/app/uploads/2022/10/AMARI-Callbox-Ultimate.pdf>, 2023, [Online; accessed 2023-01-05].