(Article begins on next page)

An RL Approach for Radio Resource Management in the O-RAN Architecture

Federico Mungari
Politecnico di Torino
Turin, Italy
federico.mungari@polito.it

***Abstract*—The new generation mobile network requires flexibility and efficiency in radio link management (RRM), in order to support a wide range of services and applications with diverse target KPI values. In this perspective, the O-RAN Alliance introduces a flexible, intelligent and virtualized RAN architecture (O-RAN), which integrates artificial intelligence models for effective network and radio resource management (RRM). This work leverages an O-RAN platform to develop and assess the performance of an RRM solution based on Reinforcement Learning (RL) and deployed as xApp in the O-RAN ecosystem. The framework receives periodic reports from the O-RAN Distributed Unit (DU) about the network status and dynamically adapts the per-flow resource allocation as well as the modulation and coding scheme to meet the traffic flow KPI requirements.**

***Index Terms*—Virtual radio access networks, O-RAN, reinforcement learning, radio resources managment**

I. Introduction

The virtualization of the Radio Access Network (RAN) has proved to be an enabling technology to serve the wide variety of heterogeneous services and the multitude of connections that the new generation mobile network is required to manage. It allows the support of a wide range of services and applications, while improving both resource scalability and utilization, and facilitating network management [1]. Indeed, traditional RANs exploiting proprietary hardware cannot easily accommodate the continuous network innovations if not at high cost, and do not offer enough resource management granularity. Along these lines, the O-RAN Alliance [1] has introduced an Open RAN (O-RAN) architecture that disaggregates and virtualizes the traditional RAN while defining its internal interfaces. An O-RAN key feature is the deployment of artificial intelligence models embedded in the RAN Intelligent Controller (RIC) with the purpose of an effective management of the RAN node and of its radio resources. In this context, it is critical to define an efficient and fully automated Radio Resource Management (RRM) framework capable of supporting the wide range of Key Performance Indicators (KPIs) in spite of the rapid changes in both network load and channel quality. Conventional RRM techniques are indeed impractical due to the increased system complexity compared to previous generation mobile networks, thus opening the door to ML-based approaches and, above all, RL ones. The latter, unlike supervised/unsupervised approaches, do not require a training dataset, which hardly translates comprehensively the problem complexity.

This work has been partially supported by TIM, through the research contract no. S20AAQIS.

[1]https://www.o-ran.org/

978-1-6654-4108-7/21/$31.00 ©2021 IEEE

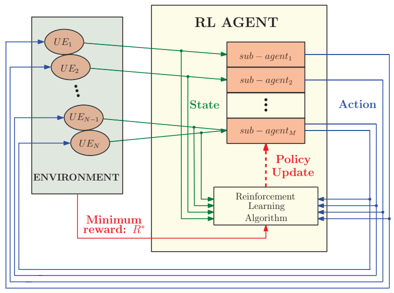

Fig. 1. RL framework architecture.

II. The RL-based framework

This work designs and evaluates a Reinforcement Learning (RL) based dynamic resource controller that leverages on-policy differential semi-gradient State-Action-Reward-State-Action (SARSA) to effectively allocate radio resources. In particular, we focus on a cellular virtual RAN (vRAN) and propose a scheme that manages the available Resource Blocks (RBs) by limiting the maximum number of RBs for which each traffic flow can compete. Further, we identify the corresponding modulation and coding scheme (MCS) that best fit the traffic flows KPIs and the channel quality. The proposed RL solution targets one of the main KPIs identified by 3GPP [2], namely throughput, by minimizing the maximum difference between desired and actual throughput, across all active traffic flows. To the best of our knowledge, no other works have addressed the under assessment problem. The framework draws upon the work recently presented in [3].

The active traffic flows are independently managed, as depicted by Fig. 1. Thus, the RL agent determines the action for each traffic flow based upon the associated gathered contextual information. The latter includes: (i) the experienced signal-to-noise ratio (SNR), signaled to the DU by the user through the Channel Quality Indicator (CQI); (ii) the amount of data in the RLC buffer; (iii) the traffic flow target throughput; (iv) the overall target throughput, aggregated over all active traffic

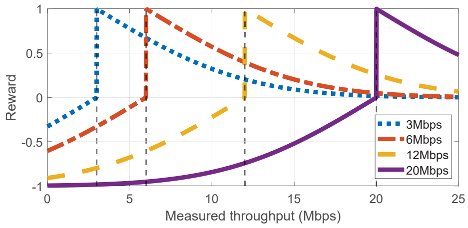

Fig. 2. Reward signal.

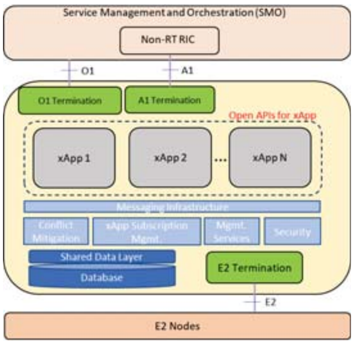

Fig. 3. The RAN Intelligent Controller (RIC).

flows. We remark that we adopt a differential semi-gradient SARSA method, which is an on-line policy approach that resorts to function approximation to deal with uncountable state spaces. Specifically, the deployed framework implements tile coding. The action value function assessment is updated through semi-gradient descent on the basis of temporal difference errors. The latter encompasses state-action-reward-state-action trajectories, ensuring that the agent avoids taking high-risk actions. Furthermore, the developed framework swiftly adapts its policy to non-stationary environments, thus proving to be particularly suitable for dynamic scenarios such as mobile user traffic connections.

III. Reward Signal

To meet the traffic flows KPI requirements, the decisions made by the RL agent must ensure that the throughput experienced by each traffic flow matches the target one. Furthermore, the measured throughput should not significantly exceed the target value, not to waste resources and possibly degrade the performance of other traffic flows. The reward function, which is depicted in Figure 2, is therefore defined as follows:

$$R^{(i)} = \begin{cases} -\mathrm{erf}\left(K(\tau_t^{(i)} - \tau_m^{(i)})\right) & \tau_m^{(i)} \leq \tau_t^{(i)} \\ 1 + \mathrm{erf}\left(K(\tau_t^{(i)} - \tau_m^{(i)})\right) & \tau_m^{(i)} > \tau_t^{(i)} \end{cases} \quad (1)$$

where K represents a scaling factor, while $\tau_t^{(i)}$ and $\tau_m^{(i)}$ are, respectively, the target and the measured throughput for the i-th traffic flow. The framework aims at maximizing the minimum reward, $R^{\star} = \min_i R^{(i)}$.

IV. System Architecture and Implementation

The RL-based framework has been integrated in an O-RAN platform. The deployed testbed provides an LTE RAN running OpenAirInterface Software-Defined Radio (SDR) eNB interfaced with an Open5GS Evolved Packet Core (EPC). The eNB has been implemented on an ETTUS Universal Software Radio Peripheral (USRP) B210 board and offers a 20-Mhz bandwidth and LTE connectivity to three real mobile terminals. Without loss of generality, only the downlink traffic scenario has been considered. Finally, the RL framework is deployed as xApp in the near-real-time RIC (near-RT RIC) of the O-RAN platform. The near-RT RIC is in charge of near-real-time RRM decisions (i.e., at a timescale ranging between 10 ms and 1 s), unlike the non-real-time RIC, which operates at larger timescale. As illustrated by the diagram in Fig. 3, the exchange of information between xApps and the RAN takes place through a messaging infrastructure which, in the employed environment, is a message broker NATS.

The deployed framework, outlined in Fig. 1, has the peculiarity of being very flexible, as it is possible to adapt its architecture and its functioning based on the environment it operates in, e.g., the number of traffic flows. At first, the periodicity with which the framework processes information on the network and updates its decisions can be freely selected. More precisely, the resource controller monitors every time-slot the status of the network and that of the users, while renews its decisions each T time-slots. The choice of T is a trade-off between computation complexity and performance. Furthermore, it is possible to create multiple policy instances, referred to as sub-agents, each of which independently and sequentially serves a subset of users, parallelizing the overall near-RT RIC workload. This approach further improves the framework scalability, as the number of sub-agent increases, thus facing the main RRM implementation challenge identified by its operating timescale.

V. Future Work

Our future work includes testing the framework at large-scale, in heterogeneous and non-stationary scenarios, taking advantage of the scalability and flexibility offered by the O-RAN platform. This will allow us to fully assess the performance of the deployed algorithm within the O-RAN architecture.

Acknowledgment

I wish to thank my advisor PhD. advisor, Prof. Carla Fabiana Chiasserini, for the supervision of this work, and Corrado Puligheddu for his helpful comments.

References

[1] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, “Cloud ran for mobile networks—a technology overview,” *IEEE Communications Surveys Tutorials*, vol. 17, no. 1, pp. 405–426, 2015.

[2] *3GPP TS 23.501 V16.3.0 Technical Specification Group Services and System Aspects; System architecture for the 5G System (5GS); Stage 2, (Release 16)*, 3GPP, 12 2019.

[3] S. Tripathi, C. Puligheddu, and C. F. Chiasserini, “An rl approach to radio resource management in heterogeneous virtual rans,” in *2021 16th Annual Conference on Wireless On-demand Network Systems and Services Conference (WONS)*, 2021, pp. 1–8.