



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/224315/>

Version: Accepted Version

Proceedings Paper:

Qazzaz, M.M.H., Kułacz, Ł, Kliks, A. et al. (2024) Machine Learning-based xApp for Dynamic Resource Allocation in O-RAN Networks. In: 2024 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN). 2024 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN), 05-08 May 2024, Stockholm, Sweden. IEEE, pp. 492-497. ISBN: 979-8-3503-4320-5.

<https://doi.org/10.1109/icmlcn59089.2024.10625184>

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Machine Learning-based xApp for Dynamic Resource Allocation in O-RAN Networks

Mohammed M. H. Qazzaz^{1,2} Łukasz Kułacz^{3,4} Adrian Kliks^{3,4}
Syed A. Zaidi¹ Marcin Dryjanski⁴ Des McLernon¹

¹ School of Electronic and Electrical Engineering, University of Leeds, Leeds, UK

² College of Electronics Engineering, Ninevah University, Mosul, Iraq

³ Institute of Radiocommunications, Poznan University of Technology, Poznan, Poland

⁴ Rimedo Labs, Poznan, Poland

Corresponding Author: Mohammed M. H. Qazzaz (ml14mmh@leeds.ac.uk)

Abstract—The disaggregated, distributed and virtualised implementation of radio access networks allows for dynamic resource allocation. These attributes can be realised by virtue of the Open Radio Access Networks (O-RAN) architecture. In this article, we tackle the issue of dynamic resource allocation using a data-driven approach by employing Machine Learning (ML). We present an xApp-based implementation for the proposed ML algorithm. The core aim of this work is to optimise resource allocation and fulfil Service Level Specifications (SLS). This is accomplished by dynamically adjusting the allocation of Physical Resource Blocks (PRBs) based on traffic demand and Quality of Service (QoS) requirements. The proposed ML model effectively selects the best allocation policy for each base station and enhances the performance of scheduler functionality in O-RAN - Distributed Unit (O-DU). We show that an xApp implementing the Random Forest Classifier can yield high (85%) performance accuracy for optimal policy selection. This can be attained using the O-RAN instance state input parameters over a short training duration.

Index Terms—5G, RAN, O-RAN, Resource Allocation, xApp, AI/ML

I. INTRODUCTION

A. Motivation

The traditional cellular network model, relying upon a single vendor's proprietary hardware and software solutions, limits flexibility, increases costs and hampers innovation [1]. Open RAN (O-RAN), also known as Open Radio Access Network, has revolutionised the telecommunications industry by decoupling hardware and software components, fostering an open and interoperable ecosystem. This transformative approach aims to mitigate traditional network limits by enhancing network adaptability, reducing costs, and promoting innovation. Additionally, the integration of artificial intelligence (AI) and machine learning (ML) capabilities within O-RAN holds tremendous promise for optimising network performance, automating operations, and enabling intelligent decision-making [2], [3].

In that context, traditionally regarded as monolithic and immutable “black-box” systems, cellular networks are transitioning towards more flexible, software-based open architectures following the O-RAN paradigm. This paradigm promotes openness, virtualisation, and programmability of RAN functionalities and components, enabling data-driven intelligent

control loops for cellular systems [4]. O-RAN empowers network operators to support bespoke services on shared physical infrastructures and dynamically reconfigure them based on network conditions and user demand. The O-RAN Alliance, a notable standardisation body, develops specifications to apply O-RAN principles to prevailing radio access technologies, including 3rd Generation Partnership Project (3GPP) LTE and 5G networks. O-RAN encompasses RAN Intelligent Controllers (RICs) that operate at different timescales, enabling data-driven applications. These applications optimise network performance by leveraging live data received from the RAN through standardised and open interfaces [5].

Within the O-RAN framework, RICs and open interfaces play a pivotal role in enabling AI and ML capabilities. The near Real-Time RIC (near-RT RIC) connects to RAN elements through the E2 interface, facilitating control loops operating between 10ms and 1s. On the other hand, the non-Real-Time (non-RT RIC) is integrated into Service Management and Orchestration (SMO) frameworks, operating at timescales larger than 1s and connecting to near-RT RIC through the A1 interface. As shown in Fig. 1, these RICs adopt a centralised perspective enabled by data pipelines that continuously flow and aggregate Key Performance Measurements (KPMs) involving the network infrastructure's status, such as user count, load, throughput, and resource utilisation, along with additional context information via RICs interfaces with RAN entities. By processing and analysing this worthy information, both RICs employ AI/ML algorithms to implement control policies and actions within the RAN based on the established policies. This paradigm presents data-driven, closed-loop control, enabling automated optimisation of network and load balancing, scheduling policies, handovers, and other essential functionalities [6].

The integration of AI and ML within O-RAN brings numerous advantages. Firstly, these technologies enhance network optimisation by analysing extensive data collected from various network elements. AI/ML algorithms identify patterns and anomalies, predict traffic patterns, and optimise resource allocation for improved efficiency and user experience. Secondly, AI and ML capabilities enable intelligent automation of network operations and routine tasks and minimise human in-

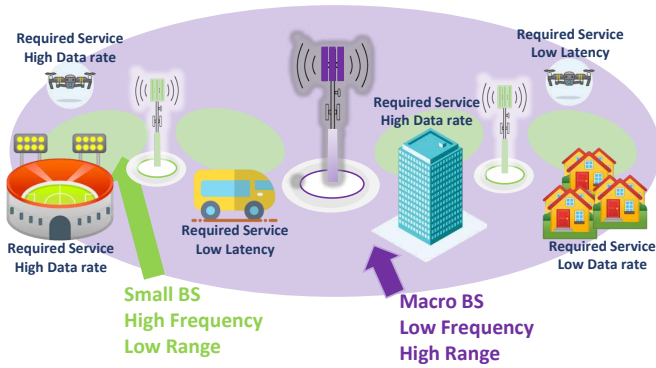


Fig. 2. A typical architecture of a 5G network where small and macro cells are deployed to serve multiple user demands

II. DYNAMIC RESOURCE ALLOCATION IN 5G NETWORKS AND ML

A. Resource Allocation in 5G Networks

Dynamic resource allocation is an important aspect of cellular networks as it enables the efficient utilisation of network resources. The concept of Network Slicing presents an evolutionary approach that exceeds the traditional idea of a singular “best network”. Instead, it defines the creation of multiple logical networks, known as slices, layered upon the physical infrastructure. These slices are designed to cater to different requirements, effectively offering a variety of specialised services to meet a wide range of needs.

In the context of 5G networks, as shown in Figure 2, a specific feature is the ability to categorise user demands based on QoS flows. Each 5G cell operates with specified radio resources, making it essential to allocate them efficiently to network slices to fulfil Service Level Agreements (SLAs) as well as meet individual QoS flow criteria. A dynamic strategy to tackle this challenge involves splitting the resources in response to real-time traffic demands. This adaptive allocation approach ensures optimal resource utilisation and enhances general network performance.

Splitting the PRBs between the associated users as per their required QoS is done inside the RAN controllers as per O-RAN Alliance specifications, where xApps and rApps are needed to perform multiple policies as per the mobile network operators (MNOs). rApps and xApps are software-specialised applications installed on the non-RT and the near-RT RIC, respectively. These applications serve as software components that boost the functionality of the RICs. A remarkable aspect of these apps is their potential to be developed by third-party vendors, independent of the RIC vendor, as O-RAN encourages a more open and cooperative ecosystem.

B. Near Real Time RIC Implementation

The Near-RT RIC, where the xApp is installed, controls the allocation of Physical Resource Blocks (PRBs) that should be allocated to the different network slices to meet their SLA requirements. To optimise the network functionality inside the

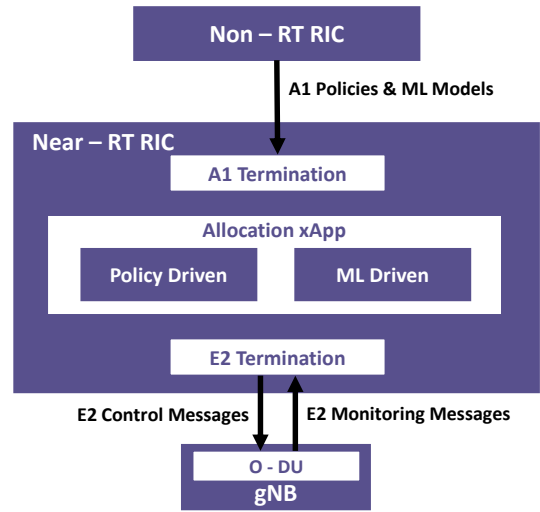


Fig. 3. RIC and xApp Deployment

Near-RT RIC, monitoring the SLA-defined requirements of the networks is mandatory within the accepted time (10ms to 1s). Multiple required parameters need to be observed continuously, such as DL/UL UE throughput, Radio Resource Utilization, DL/UL total available PRB, Number of Active UEs, and other parameters via the E2 interface from the O-RAN central unit (O-CU) and the O-RAN Distributed unit (O-DU).

On the other hand, to enable the Near-RT RIC to obtain full closed-loop control of the O-CU/O-DU via the E2 interface, multiple actions could be taken by changing different network parameters inside CU/DU such as the Min&Max PRB Policy Ratio and the dedicated PRB Policy Ratio as well as the radio resources splitting policy between users. As shown in Figure 3, the Near-RT RIC sends the E2 commands to the O-DU, where the Medium Access Control (MAC) scheduler operates, to specify the policy for slice management. The policies by which the xApp is controlled are provided through the A1 interface. The A1-P messages are implemented based on O-RAN alliance specifications to comply with the standard. Each policy performs a different way of base station resource allocation when PRBs are allocated so as to provide all the needed slices with the optimal number of PRBs.

The xApp is triggered either by event occasions when the network status changes due to a handover action or configuration change or by a constantly triggered timer which the network administrator could select to monitor the network periodically and solve any emerging issues.

III. RESOURCE ALLOCATION ML-BASED XAPP

A. System Model

Including ML capabilities within the dynamic resource allocation in O-RAN networks brings significant advantages. ML can play a crucial role in optimising resource allocation by analysing vast amounts of data and identifying complex patterns and correlations.

Based on the provided details, the focus of this article lies in the development and deployment of an ML-based xApp to optimise the allocation of PRBs within each base station and select the most suitable resource allocation policy for the network slices. An ML model must be trained in the SMO framework or directly in the non-RT RIC. The well-trained model can be deployed in the near-RT RIC, while updates and enrichment information of the ML model are transmitted from the non-RT RIC to the near-RT RIC through the A1 interface.

To build the xApp and ML model, we have built a Het-Net 5G scenario through extreme programming with Python programming comprising a high-power macro base station covering an area of 20 km × 20 km and a small cell in an area of 10 km × 10 km. Two separated frequency bands, 800 MHz and 2 GHz, with channel bandwidths of 5 MHz and 10 MHz, respectively, are chosen. Voice and Mobile Broadband (MBB) users are randomly deployed within the simulated area, with varying numbers and locations. Multiple resource allocation policies are explored in the near-RT RIC and xApp to determine how available resources should be managed and allocated to the associated users to meet the required data rates of the users.

Model Parameters:

- Network: 5G Heterogeneous Network (HetNet)
- Base Stations: One Macro Base Station (MBS) and One Small Base Station (SBS)
- Coverage Area: 20 km × 20 km for MBS and 10 km × 10 km for SBS
- Frequency Bands: 800 MHz and 2 GHz
- Channel Bandwidth: 5 MHz and 10 MHz
- Type of UEs: Voice and Mobile Broadband (MBB)
- Minimum QoS Requirement: 250 Kbps for Voice UEs and 3 Gbps for MBB UEs

To address the importance of the dynamic resource allocation adjustment, four different resource allocation policies are proposed. These policies are observed through extensive simulations to comprehensively assess their impact on network performance. When distributing the total allocated resources for a base station, which is denoted as $PRBs$, among the associated voice users, represented by N_{voice} , and the associated MBB users, denoted as N_{mbb} , users will collectively utilize these resources, represented by PRB_{user} according to one of the following policies:

- *Equal Allocation Policy*: Under this policy, an equal split of PRBs between voice and MBB users will occur, disregarding their respective QoS requirements as described in (1). This policy is suitable when the number of users is relatively low.

$$PRB_{user} = \frac{PRBs}{N_{voice} + N_{mbb}} \quad (1)$$

- *Voice Priority Allocation Policy*: This policy allocates more resources (M times) to voice users within each cell type to ensure a higher priority for network access. It

is an ideal selection for scenarios where voice services are important. The allocated $PRBs$ for each voice user denoted by PRB_{voice} , and for each MBB user, PRB_{mbb} , is represented in (2)

$$PRB_{mbb} = \frac{PRBs}{(M * N_{voice}) + N_{mbb}} \quad (2)$$

$$PRB_{voice} = M * PRB_{mbb}$$

- *MBB Priority Allocation Policy*: Alternatively, this policy allocates more resources to MBB users (K times) within each cell type compared to voice users as defined in (3). This approach prioritises broadband services and is valuable when dealing with high data traffic requirements.

$$PRB_{voice} = \frac{PRBs}{N_{voice} + (K * N_{mbb})} \quad (3)$$

$$PRB_{mbb} = K * PRB_{voice}$$

- *Dedicated Resources Reservation Policy*: This policy reserves a specific portion of the resources for voice users and another portion for MBB users within cells. This ensures dedicated resources for essential services, even during peak network usage and saves the remaining resources for other services. For voice users connected to macro base stations, the resource allocation policy involves dividing a reserved portion, denoted as α , from the total available $PRBs$ equally as represented in (4). On the other hand, voice users connected to small base stations will also receive an equivalent allocation from the voice users connected to macro base stations. The allocation of resources for MBB users, which is defined in (5), follows a different approach. These resources are distributed based on a selected portion, denoted as β , from the total $PRBs$ regardless of whether they are connected to macro or small base stations.

In macro base stations

$$PRB_{voice} = \frac{PRBs}{\alpha * N_{voice}} \quad (4)$$

And in any kind of base stations

$$PRB_{mbb} = \frac{PRBs}{\beta * N_{mbb}} \quad (5)$$

B. Results and Discussion

The performance and effectiveness of the mentioned resource allocation policies, in conjunction with the ML model, have been thoroughly investigated and evaluated. We investigate the impact of each resource allocation policy on network performance, taking into account a diverse deployment of associated users with different types. To evaluate the effect of different policies on network performance, the network outage, a critical performance metric, has been considerably analysed. The measurement of network outage was achieved by determining the percentage of users who failed to get the minimum QoS requirements out of the total number of the associated users within the cell, disregarding the user types. Multiple

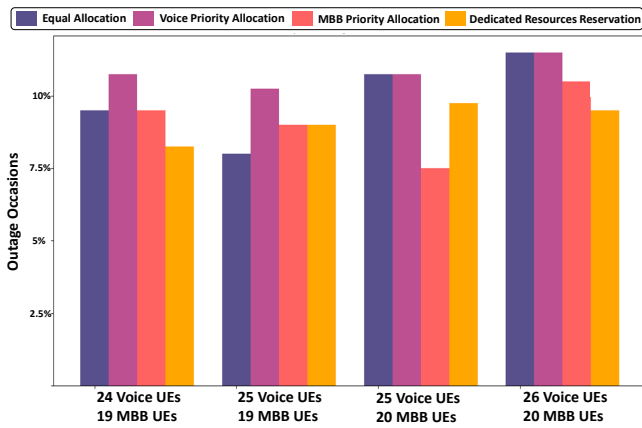


Fig. 4. Small Cell Network Outage under Different Policies

simulations have been conducted, encompassing various user deployment scenarios within the network, to accurately capture the system outage for each configuration and policy setting.

The simulations produced valuable insights into the behaviour of our 5G simulated network. To emphasise the variance in network performance under different resource allocation policies, we considered high QoS requirements for each user in our simulation. We observed that, even under similar conditions concerning user types and numbers, the implementation of each of the four available policies led to varying outage events. Figure 4 presents the outage occurrences when different policies are applied to particular network configurations in a small cell. In particular, the Dedicated Resources Reservation policy yielded the least user outages when there were 24 voice users and 19 MBB users. However, when an additional voice user is associated with the same cell, implementing the same policy does not yield optimal performance, whilst the Equal Allocation policy is shown to be more effective. Similar observations in performance were noticed when cells had more users of different types. These impacts emphasise the need for dynamic selection of policy based on the network conditions and the number and types of users associated with each cell.

Accordingly, out of that, we proceeded to develop an ML model trained using the comprehensive and detailed information collected from the network to select the optimal performance policy for each network condition reasonably. We aim to enhance the overall act of resource allocation in 5G networks and lead to improving network performance, user satisfaction, and overall Quality of Experience (QoE).

To train the ML model, comprehensive simulations were conducted to generate datasets consisting of a large number of input pairs. These input pairs included the count of the users from both voice and MBB user types, along with corresponding outage information and network configuration. The objective of the classification problem is to determine the best allocation policy for each network configuration, as certain policies may result in suboptimal network performance.

After careful evaluation of various classification algorithms,



Fig. 5. ML Training Time

multiple algorithms show an excellent performance in terms of classification accuracy during the model validation, i.e. more than 85%. However, some of the algorithms, as demonstrated in Figure 5, indicate a different training period which might be considered as an essential factor, especially in the case of repeating the training periodically or performing the online training event-based. That is why the Random Forest classifier was chosen as the ML algorithm for this study.

The utilisation of the Random Forest classifier as the ML algorithm, along with the dataset collection and simulations, contributes to the effectiveness and reliability of the proposed resource allocation framework.

Applying our trained model presents a dynamic allocation mechanism that can intelligently select the most suitable policy based on the specific configuration and user demands. Fig.6 illustrates the selected policy for each network configuration and user count. A slight variation in the number of associated users from different user types can significantly impact the occurrence of network outages. Choosing the best performance policy will ensure the minimum outage at each cell state.

Furthermore, it is worth mentioning that within the macro cell observations, the MBB Priority Allocation policy rarely emerged as the best allocation choice, while a similar remark was observed with the Voice Priority Allocation policy in the small cells. In the case of the MBB Priority Allocation policy, it ensures that all cells, regardless of their kind, allocate resources primarily to MBB users, no matter the number of connected users. While this approach guarantees a prioritised service for MBB users, it may lead to a higher outage for voice users and system outages at the end. The same action happens with the Voice Priority Allocation policy within the small cell results, where using it leads to higher outages than using the other available policies.

The integration of ML within the xApp enables a responsive and efficient resource allocation process. By accurately estimating the current traffic demands and dynamically adapting

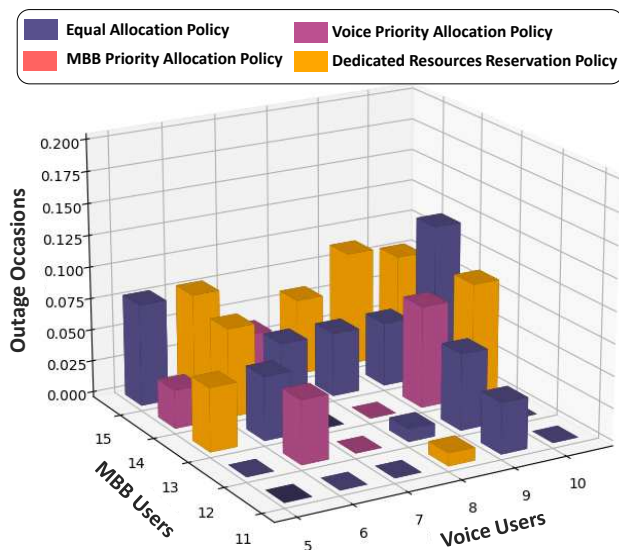


Fig. 6. Macro Cell Network Outage

the allocation of PRBs, ML algorithms facilitate the efficient utilisation of network resources and enhance the overall network performance. This approach ensures that the network slices receive the appropriate allocation of resources, leading to improved user experience, optimised network efficiency, and enhanced fulfilment of SLAs and QoS requirements. By leveraging historical data and real-time monitoring, ML algorithms optimise the allocation of PRBs to meet the dynamic traffic demands and ensure the best possible network performance. The utilisation of ML algorithms within the xApp empowers O-RAN networks to achieve efficient and intelligent resource management, contributing to enhanced user experience and network efficiency.

IV. CONCLUSION

In conclusion, the integration of AI/ML capabilities within the O-RAN brings considerable advantages to the telecommunications industry. This paper presents an xApp-based implementation for dynamic resource allocation using ML within O-RAN architecture. A Het-Net 5G network has been built to explore the consequences and observations of resource allocation policies. Macro and small 5G cells deployed with different frequency bands and bandwidths. Four different policies have been implemented at multiple network configurations, and the user outage was chosen as a performance metric. Our simulation experiments show that under similar network conditions, the implementation of each of the four policies led to different outage events. Therefore, an ML-based xApp has been developed to enable a dynamic selection for resource allocation policies. The proposed ML model effectively picks the best allocation policy for each base station, optimising resource allocation and achieving Service Level Specifications (SLS). With a high accuracy rate of 85% for policy classification, the Random Forest Classifier displays its potential to enhance scheduler functionality within the O-DU. Leveraging input

parameters from the O-RAN instance conditions, the xApp performs rapid convergence to changing network conditions. This study highlights the transformative capability of ML in advancing network efficiency and performance, paving the way for more resilient and adaptable telecommunication networks.

ACKNOWLEDGMENT

This research was funded by EP/X040518/1 EPSRC CHED-DAR and was partly funded by UKRI Grant EP/X039161/1 and MSCA Horizon EU Grant 101086218.

REFERENCES

- [1] Leonardo Bonati, Michele Polese, Salvatore D'Oro, Stefano Basagni, and Tommaso Melodia. Open, programmable, and virtualized 5g networks: State-of-the-art and the road ahead. *Computer Networks*, 182:107516, 2020.
- [2] Paul H Masur, Jeffrey H Reed, and Nishith K Tripathi. Artificial intelligence in open-radio access network. *IEEE Aerospace and Electronic Systems Magazine*, 37(9):6–15, 2022.
- [3] Jean-Christophe Plantin. The geopolitical hijacking of open networking: the case of open ran. *European Journal of Communication*, 36(4):404–417, 2021.
- [4] Dariusz Wypiór, Mirosław Klinkowski, and Igor Michalski. Open ran—radio access network evolution, benefits and market trends. *Applied Sciences*, 12(1):408, 2022.
- [5] Leonardo Bonati, Salvatore D'Oro, Michele Polese, Stefano Basagni, and Tommaso Melodia. Intelligence and learning in o-ran for data-driven nextg cellular networks. *IEEE Communications Magazine*, 59(10):21–27, 2021.
- [6] M Dryjanski and R Lundberg. The o-ran whitepaper; overview, architecture, and traffic steering use case. *Overview Architecture and Traffic Steering Use Case*, 2021.
- [7] Open RAN Alliance. O-ran-wg1-o-ran architecture description-v01.00.00. *Technical Specification*, 2020.
- [8] Michele Polese, Leonardo Bonati, Salvatore D'oro, Stefano Basagni, and Tommaso Melodia. Understanding o-ran: Architecture, interfaces, algorithms, security, and research challenges. *IEEE Communications Surveys & Tutorials*, 2023.
- [9] Mohammed M. H. Qazzaz, Syed A Zaidi, Des McLernon, Abdelaziz Salama, and Aubida A Al-Hameed. Low complexity online rl enabled uav trajectory planning considering connectivity and obstacle avoidance constraints. In *2023 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, pages 82–89. IEEE, 2023.
- [10] Michele Polese, Leonardo Bonati, Salvatore D'Oro, Stefano Basagni, and Tommaso Melodia. Colo-ran: Developing machine learning-based xapps for open ran closed-loop control on programmable experimental platforms. *IEEE Transactions on Mobile Computing*, 2022.
- [11] Daniel Dik and Michael Stübert Berger. Transport security considerations for the open-ran fronthaul. In *2021 IEEE 4th 5G World Forum (5GWF)*, pages 253–258. IEEE, 2021.
- [12] Han Zhang, Hao Zhou, and Melike Erol-Kantarci. Team learning-based resource allocation for open radio access network (o-ran). In *ICC 2022-IEEE International Conference on Communications*, pages 4938–4943. IEEE, 2022.
- [13] Cezary Adamczyk and Adrian Kliks. Conflict mitigation framework and conflict detection in o-ran near-rt ric. *IEEE Communications Magazine*, 2023.
- [14] Abdelaziz Salama, Syed Ali Zaidi, Des McLernon, and Mohammed M. H. Qazzaz. Flcc: Efficient distributed federated learning on iomt over csmalca. In *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, pages 1–6, 2023.
- [15] Amardip Kumar Singh and Kim Khoa Nguyen. Mcoranfcd: Communication efficient federated learning in open ran. In *2022 14th IFIP Wireless and Mobile Networking Conference (WMNC)*, pages 15–22. IEEE, 2022.
- [16] Rimedo Labs. Quality of Service-based Resource Allocator xApp (QRA-xApp). Technical specification, RIMEDO Labs, 2022.