# Enhancing the Handover Algorithm with an Intelligent Approach in the O-RAN Architecture

**Kleber Vilhena**[1], **Carlos Rocha**[1], **Rafael Veiga**[1], **Lucas Bastos**[1], **Eduardo Cerqueira**[1], **Denis Rosário**[1]

[1]Federal University of Pará (UFPA) - Belém, PA, Brazil

{kleber.vilhena, carlos.almeida.rocha, lucas.bastos}@itec.ufpa.br,

rafael.teixeira.silva@icen.ufpa.br, {cerqueira, denis}@ufpa.br

***Abstract.*** *O-RAN is an architecture that promotes interoperability and openness in 5G Radio Access Networks (RAN) using scheduling, disaggregation, and virtualization. RICs (RAN Intelligence Controllers) offer solutions such as Machine Learning (ML), traffic steering, anomaly detection, and QoS (Quality of Service) support. Novel intelligent handover strategies are critical to the success of 5G or even 6G O-RAN-based networks. This paper proposes and evaluates an intelligent handover algorithm for O-RAN environments. It leverages an LTE testbed featuring O-RAN architecture to assess downlink and uplink performance across various User Equipment (UE) scenarios. The proposed scheme was implemented and tested using ns-O-RAN, an O-RAN system integrated with the NS-3 simulator. Our simulator results demonstrate a throughput and delay enhancement compared to traditional handover methods across various scenarios involving 50 to 100 UEs.*

## 1. Introduction

Traditional Radio Access Networks (RANs) typically consist of closed, monolithic units that integrate all network functions, encompassing the layers of the 4G/5G protocol stack and network interfaces [Lacava et al. 2023b, de Oliveira et al. 2023]. Traditional RANs face challenges, including vendor lock-in, limited configuration options, increased unit prices, coordination difficulties, closed interfaces, and scalability concerns [Linsalata et al. 2024]. The open-RAN (O-RAN) Alliance [O-RAN Alliance 2021] has played a vital role in developing the O-RAN vision by extending the capabilities of 3rd Generation Partnership Project (3GPP) by offering an open and interoperable ecosystem, transforming the telecommunication ecosystem. O-RAN aims to provide greater flexibility, customization, and innovation for network operators, addressing the limitations of traditional RAN architectures [Alavirad et al. 2023]. For instance, O-RAN introduces virtualization and disaggregation of RAN technologies, dividing network functions into white-box software and hardware components connected via open interfaces, allowing network virtualization and programmability [Musa et al. 2023].

The RAN intelligent controller (RIC) optimizes RAN components, serving as a centralized abstraction that analyzes data collected from RAN functions, applies control actions, and enables algorithmic control [Bonati et al. 2021]. In this way, RIC enhances Radio Resource Management (RRM) with data-driven approaches utilizing real-time telemetry from the RAN, processed by ML algorithms for optimization and control of the RAN [Zangooei et al. 2023]. The ML-based RIC control functions aim to solve existing hard-to-solve issues in the RAN domain, such as mobility, scheduling, admission control, and others [Brik et al. 2023] Hence, O-RAN plays an essential role for 5th generation (5G) networks and beyond by enabling the deployment of multi-vendor, interoperable components and programmatically optimized through a centralized abstraction layer [Garcia-Saavedra and Costa-Perez 2021].

Traditional RAN encounter significant challenges because of their inflexible architectural designs. These fundamental issues frequently lead to operational inefficiencies, manifesting as suboptimal network performance, increased latency, and reduced capacity to handle high traffic volumes. O-RAN improved network efficiency and support for a broader range of suppliers and solutions that translate into more efficient network configurations, enabling networks to adapt dynamically to load and environmental changes, thus ensuring seamless handovers. In this way, the O-RAN environment has been an ongoing focus of research and development since it opened new avenues for network operators to explore greater flexibility, customization, and innovation in their operations [Bonati et al. 2022]. For instance, near-real-time (Near-RT RIC) RIC hosts multiple applications to define control actions applied to the RAN model, called extensible applications (xApps). Specifically, a xApp runs on a microservice that receives real-time data from the RAN through the E2 interface (*e.g.*, user, cell, application, and other measurements) and (if necessary) computes and sends back control actions (*e.g.*, handover, load balancing, and others) [Polese et al. 2023].

Numerous papers have extensively investigated the handover topic, with some of the fundamental traditional network applications highlighted [Costa et al. 2020]. Within the O-RAN context, RICs open opportunities to optimize handovers using 3GPP-defined measurement report parameters to better performance [Riccio et al. 2023]. For instance, the xApp collects contextual measurements from the user and network, which feed an ML to act in this scenario [Hamdan et al. 2023]. For instance, using an ML algorithm for user position prediction is possible. Therefore, the O-RAN model could rely on QoS and network metrics to see how this approach can improve the network performance [Bonati et al. 2022]. O-RAN enhances handover management by linking session transfers between cell base stations. Offering improvement for network operations that respond dynamically to changing conditions and user demands [Wang et al. 2021]. ML enables predictive analysis, allowing networks to anticipate handover events. Resource allocation and network traffic management can be optimized. This approach improves the user experience by minimizing interruptions during handovers. Advanced ML schemes in O-RAN systems enable fast and data-driven handover decisions. That approach reduces call drops and improves the user experience [Sun et al. 2020].

This paper proposes an intelligent handover approach for improving mobility support with Quality of Service (QoS) in O-RAN-based networks. Our methodology enriches handover processes by utilizing ML to optimize User Equipment (UE) transition with diverse mobility patterns between network nodes. Our method evaluates critical performance indicators: throughput, Packet Delivery Ratio (PDR), delay, and jitter. These metrics are critical in determining the effectiveness of our ML-driven handover decisions. We aim to identify the optimal eNB connections for UEs in a simulated dynamic network environment. Our approach aims to bolster the network stability of gNBs and enhance the QoS for UEs, leading to a more reliable environment. Simulator results show that our intelligent approach outperforms a traditional O-RAN handover scheme.

The remainder of this paper is structured as follows. Section 2 presents the overview of works that deal with RIC and their main drawbacks; thus, papers that use other O-RAN approaches and handover methods with ML. Section 3 introduces the ML prediction for handover in the O-RAN approach, which will set the parameters to select the best scenario fit. Section 4 discusses the evaluation of the ML approach and obtained results in the throughput, PDR, delay, and jitter. Finally, Section 5 describes this paper's conclusion and presents some future work directions.

## 2. Related Works

Most existing O-RAN works deal with RIC implementation for QoS improvement using different types of communications. For instance, [Bonati et al. 2022] highlighted Open-RAN Gym, an open toolbox for developing O-RAN-compatible ML solutions. The paper demonstrates how two xApps designed with OpenRAN Gym can control a large-scale RAN, enabling ML for the management network to improve QoS. Based on experiments, the O-RAN ns-3 module facilitates modeling a network architecture that complies with O-RAN specifications. It incorporates essential classes such as RIC, mirroring O-RAN's RIC, and ML Near-RT. This work shows that communication between reporting modules that establish connections with reporting nodes simulates, serving as communication terminals with the O-RAN RIC, similar to the E2 Terminators in the O-RAN. Motivated by these limitations, we introduce an ML approach with more UEs in RIC and Near-RT. In this way, we use an SQLite database to store coordinates and packets lost of cells, improving the handover scenario for eNB/gNB towers.

[Baldesi et al. 2022] introduced a Channel-Aware Reactive Mechanism (ChARM), a data-driven O-RAN-compliant 5G-and-beyond networks framework. ChARM operates within O-RAN specifications without requiring modifications to existing 3GPP standards. The paper demonstrates ChARM's performance in spectrum-sharing scenarios, specifically between LTE and Wi-Fi in unlicensed bands. This paper proposes to use the O-RAN in a 5G scenario, which also has its networking methods. Therefore, the O-RAN follows some metrics to use our ML to predict loss packets and mobility prediction (distance of UEs to eNB) provided by the OPEN-RAN Central Unit (O-CU). However, this approach is not oriented towards using ML, with intelligent solutions in handovers, using an SQLite database of the best positions of the UEs, where it can be whenever necessary to reuse the best positions in which the UEs are in the O-RAN.

[Lacava et al. 2023b] The handover management framework enhances the optimization of O-RAN TS (Traffic Steering) based on Q-Learning and a system-level approach. This approach affects the network performance and QoS. The TS scheme can improve overall network performance when traffic preferences are considered. Efficient data flow management is crucial in 5G networks to meet diverse service requirements. To this end, this study utilizes network slicing (NS) and multi-connectivity (MC) technologies to improve data rates for Enhanced Mobile Broadband (eMBB) services and reduce latency for Ultra-reliable Low Latency Communications (uRLLC) services. The traffic steering xApp and discuss the algorithm design to determine the optimal target cells for mobile user handover. In this paper, we use the xApps in our ML application and develop a data-driven mobile user-based traffic steering/handover optimization. The simulation is also based on the eMBB to represent our results accurately with mobile traffic in an O-RAN system.

[Lacava et al. 2023a] presented a portable 5G Non-Stand Alone (NSA) architecture using NS3 for platform flexibility and functional throughput tests with different channel coding and modulation schemes. Some performance tests are shown for 5G use cases but using the 4G network. The authors present performance results regarding throughput for Internet browsing, voice transmission, and streaming for different distances between transmitter and receiver. The code developed for this paper uses the handover Management capability for the TS use case of O-RAN through the handover Management message, allowing the xApp to send an RIC control message with the RAN ID function. The RIC in our code uses their functions to manage and send the handover Management message after the ML in O-RAN results in the best fit of a client with their features. In this way, our simulation aims to use similar to perform conventional handover and ML with a scenario with more UEs.

[Sahbafard et al. 2023] delved into 5G, the latest cellular technology designed to improve data rates and accommodate a range of new applications. We emphasize the necessity for experimental deployments to thoroughly evaluate the performance of 5G, particularly in the context of specific use cases. The authors used a testbed approach to look at the function of Software Defined Network (SDN) management in a minor place. However, we must use our approach with more access points and mobile users to look for a realistic mobile user scenario.

[Gavrilovska et al. 2020] described how to operate an O-RAN and an edge cloud computing. In this way, they mention how improving this in the network management is called C-RAN. This work demonstrates the concepts of an SDN use of scalability. However, in their scenario, the number of users in the simulation was small. Their paper concludes by highlighting metrics such as latency, link capacity of the BBUs, and throughput. Inspired by this metric's uses, our scenario improves and performs a more extensive simulation with more mobile users and more eNBs.

In summary, our handover algorithm with an intelligent approach considers using the O-RAN architecture to improve LTE 4G/5G networks, which is assisted by xAPP's management through multiple UEs acting between the eNBs/gNBs. It considers the presence of mobile UEs, each running applications with different performance requirements. Furthermore, it extends the use of messages in the O-RAN control plane, namely the performance field, allowing a more accurate representation of the network and communication conditions of each UE and the different services running. Furthermore, we introduce a more innovative approach to managing handover messages that significantly reduces the control layer overhead without compromising the UEs' implementation and the accuracy of the intelligent algorithm, nor the performance of xApp running the ML optimization algorithm with O-RAN.

## 3. O-RAN Smart Prediction for Handover Management

A handover approach requires linking the eNB and the UE based on signal quality to ensure seamless connectivity and optimal network performance as users move through different coverage areas. Therefore, this section presents an O-RAN approach based on the network's behaviors for smartness handover. To this end, we introduce the modules and the implementation details for obtaining this behavior by evaluating an ML handover approach.

### 3.1. O-RAN Overview

The O-RAN architecture breaks the classical approach by adopting the principles of disaggregation, openness, virtualization, and programmability, enabling data exposure and analysis and data-driven optimization, closed-loop control, and automation. In this context, handover algorithms could take advantage of the O-RAN architecture since RIC provides the deployment of network controllers and applications for managing, configuring, monitoring, and maintaining radio unit operations. Specifically, near-RT RIC hosts xApps applications, a software application that runs network management services, such as QoS monitoring, resource allocation, connection management, and frequency scanning [Brik et al. 2023]. In this sense, xApps are vital applications operating on the near-RT RIC, comprising multiple micro-services that receive input data from RIC-RAN interfaces and leverage ML for near-RT control. The xApps are integral components facilitating intelligent and automated RAN control, and the development of xApps is technology-agnostic, depending on an O-RAN-defined interface.

Traditional handover processes in mobile networks involve transferring an ongoing call or data session from one cell base station to another as a user moves through the network. This process typically relies on predefined thresholds for signal strength and quality [Lacava et al. 2023b]. Intelligent handovers use ML to analyze information from the UE and eNB to decide who eNB the UE intends to link with, such as user behavior patterns and network conditions.

The xApp receives QoS metrics and position coordinates via the E2 interface, providing additional functionality as output to RAN operations based on ML algorithms. For instance, QoS metrics are crucial indicators for network operators, application developers, and users to assess network performance [Lacava et al. 2023b]. In addition, the position coordinates can be used by ML algorithms for mobility prediction, which could feed the handover algorithm.

Figure 1 shows the O-RAN architecture for orchestrating mobility prediction, which integrates various components that interact to facilitate dynamic network optimization. The O-RAN Cloud is at the core of the architecture, hosting the Near-RT RIC that hosts xApps to leverage ML algorithms for predictive analytics. It also shows the flow of the collected data, starting with collecting client position and network metrics data, which feeds into the orchestration process. These metrics are processed by the Near-RT RIC, enabling it to make informed decisions for mobility management [Lacava et al. 2023a]. The O-RAN Cloud communicates with the eNB, an evolved Node B or base station, facilitating the necessary adjustments to optimize the user or UE experience.

The xApp for mobility prediction has the following steps: data acquisition, pre-processing, feature extraction, and classification. In the data acquisition step, an xApp receives input from the UEs with various network parameters, such as mobility patterns, distance to each tower (km), mean packet loss, and cell load, providing information related to the UE's patterns of network behavior [Hamdan et al. 2023]. This data goes through a pre-processing step, where outliers are removed.

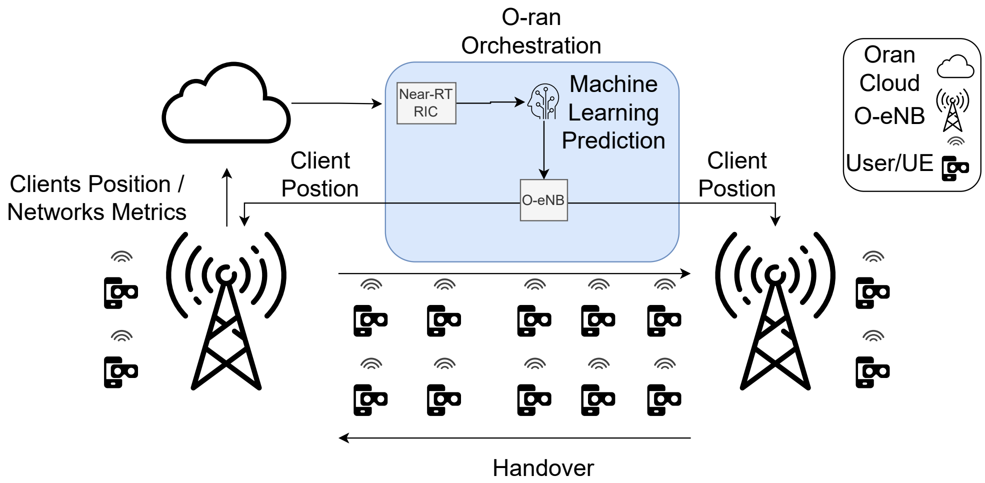


**Figure 1. O-RAN approach for Orchestration mobility prediction**

After pre-processing the data, extracting relevant features from the raw data is necessary. Based on this information, it is possible to identify the network patterns. This information provides an overview of how or when it is necessary to determine the best eNB client and who is best to link with to provide network access. Improvement occurs when comparing each UE with smartphones, VR equipment, or any device capable of receiving packets using a streaming approach.

In the XApp for handover, a set of features is checked against the patterns stored in the database so that it can predict the best fit for whom eNB links each UE. The handover process responds if the UE worsens QoS, while the other eNB can improve the network parameters to link the UE [Riccio et al. 2023].

### 3.2. Prediction Overview

We start the handover prediction by extracting data to train the ML algorithm with some network values. Specifically, the ML algorithm needs historical datasets paired with their handover outcomes to train the estimates of who eNBs to link with each UE. These parameters will be used to train and predict who the best-fit eNB based on the distance to each eNB (km), mean packet loss, and cell load. This way, we use the validation and testing data to appraise the model's predictive prowess and generalizability. Iterative refinement based on performance metrics ensures the model's continuous evolution towards optimal accuracy. Upon satisfactory validation, the ML model is deployed into the O-RAN ecosystem, where real-time decision-making happens, leveraging live data to orchestrate network resources dynamically. The mobility prediction collects network parameters from UEs, including mobility patterns, distance to towers, packet loss, and cell load. After preprocessing to remove outliers, this data informs the prediction process, enhancing handover accuracy and overall network performance.

Figure 2 provides a schematic representation of the data extraction process tailored for an O-RAN approach using ML algorithm for handover management. The process begins with the aggregation of raw data (1), which comprises various UE metrics such as mobility patterns, distances to cell towers (km), mean packet loss, and cell load. These raw data points are then systematically labeled (2), assigning contextual information to each data point, which is essential for identifying their response of who eNB the UE needs to link. In phase (3), the process actively formats the labeled data into a structure that facilitates feature analysis. At that point, the user extraction data contains refined information on the features necessary for the ML handover decision process. The following phase (4) entails a restructuring of data frequency to ensure uniformity and coherence across all data points. This step critically synchronizes data inputs for the subsequent training of the ML model, allowing it to accurately predict handover events based on temporal patterns and network dynamics. Finally, the process reaches the feature extraction phase (5), extracting features such as mobility patterns, proximity to network towers, packet loss rates, and network traffic load. At that point, the process filters the data from the raw dataset. These features are essential to make an ML model that can accurately predict when a handover should occur, optimizing network efficiency and improving user experience. Hence, this extraction and feature processing pipeline enables the creation of a predictive model for network handovers, which is crucial for maintaining seamless connectivity and service quality in mobile networks as part of the O-RAN.

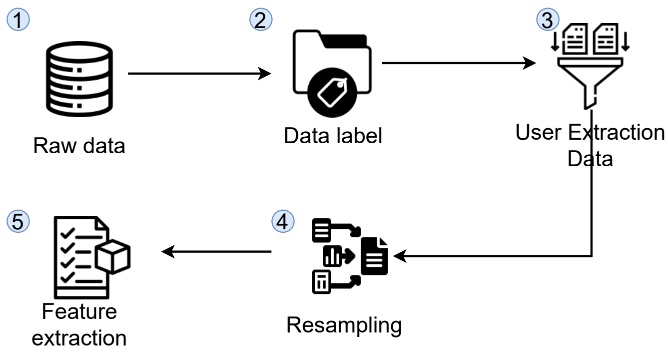


**Figure 2. Data extraction**

In developing our ML, the essential approach is a method that can be generated and progressively trained during its use. For this purpose, we must considered a multi-class Neural Network classification. Specifically, the input layer that receives network features as inputs. Sequential hidden layers incorporate linear functions by applying a transformation that maps inputs to a desirable feature space. This linear function is $y = xA^T + b$, where x is the input, A is the weight matrix, b is the bias vector, and y is the output. After each linear transformation, a ReLU activation function introduces non-linearity to the model, allowing it to capture more complex relationships in the data.

The model output layer transforms the processed features into a probability distribution of the classes using a softmax function. The model is trained end-to-end using cross-entropy loss, which will help the model be adept at multi-class classification in each training epoch. In that way, each training epoch improves the model by minimizing this loss function. Each epoch will set an iterative optimization of weights. Adjusting the weights and biases of the linear transformation helps improve the model's accuracy. The model learns to choose the most suitable NB to link each UE connection by training using the network parameters, such as the distance to each eNB (km), mean packet loss, and cell load. They are collecting these data from the network using the RIC to do a data process for training the ML model with specific data measurements.

The ML train uses a database using the traditional rules of handover manager communications, which provides a complex way of simulating to collect data on the network behavior. The Handover Manager communicates its ML decision regarding the relative loss packet parameters of an UE to both the serving and target radio base stations, indicating the mobile node for transfer. It facilitates the exchange of control messages between the radio base stations, conveying node-specific details. After that, the system triggers the handover process, transitioning the communication path of the mobile node from the serving radio base station to the target radio base station.

## 4. Evaluation

This section explores the experimental evaluation of our intelligent handover scheme for O-RAN scenarios. Our simulation supports the O-RAN module within the NS3 Framework, while the PyTorch library powers our ML component. This pre-trained ML model facilitates transfer decisions based on location and packet loss data. It is worth mentioning that our entire project is open-source and accessible on GitHub[1].

### 4.1. Simulation

NS-O-RAN is an open-source simulation platform that combines a functional 4G/5G protocol stack on NS-3 with an O-RAN-compatible E2 interface. This platform complements WIoT's OpenRAN Gym with a simulator that can enhance data collection and xApp testing, which is a critical step toward enabling efficient generic AI and ML solutions for OpenRAN and 5G/6G systems. NS-O-RAN is designed and implemented to enable the integration of O-RAN software, such as the Near-RT RIC from the O-RAN Software Community, with large-scale 5G simulations based on 3GPP channel models and detailed network modeling. Complete 3GPP RAN protocol stack. This makes it possible to collect key RAN performance metrics (KPMs) in different simulated scenarios and with different applications, such as, multimedia streaming, web browsing, wireless virtual reality, and holograms.

Table 1 outlines the main parameters for orchestrating a dynamic 5G network environment under the ns-O-RAN framework in our simulation. We employ a Random Walk mobility model distributed across a sophisticated network architecture featuring seven NR gNB towers. At the heart of this setup is a central tower, an amalgamation of LTE eNB and NR gNB technologies. The network operates on a 3.5 GHz center frequency and has a 20 MHz bandwidth allocation, reflecting anticipated conditions in a real-world 5G scenario. We evaluate network performance under varying loads by testing with 100, 75, and 50 mobile users. We conducted 33 simulation runs with different randomly generated seeds, and the results include a 95% confidence interval.

NS-O-RAN employs a traditional handover algorithm described in Lacava's 2023 work and incorporated into the NS-3 LTE module [Lacava et al. 2023b]. The method used for the transfer is known as the "traditional power budget algorithm", which regularly monitors the received reference signal Power (RSRP) of a UE's serving cell and neighboring cells so that once the RSRP of a neighboring cell is greater than that of the serving cell, the UE is handed over to that neighboring cell. On the other hand, our intelligent handover algorithm makes use of the ML-like O-RAN reports approach, such as introduced in Section 3. Our approach considers a ML algorithm that uses the reported

[1]https://github.com/KleberVilhena/LTE-NS3

**Table 1. Simulation Parameters**

| Characteristics | Description | Value |
|---|---|---|
| The Central Tower | Macrocell | One LTE eNB + NR gNB |
| Remaining Towers | Macrocell | Six towers |
| Center Frequency | Frequency at which a signal or communication channel | Using 3.5 GHz |
| Bandwidth | Capacity of a communication channel or network to transmit data | Using 20 MHz |
| ISD(m) | Distance between two or more sites in a network | $\approx$ 600 meters |
| Mobile UE | Number of mobile users | 100, 75 and 50 |
| Mobility model | Mobility used by UEs | Random Walk |

locations to calculate the distance between each UE and eNodeB and then performs a handover to a UE if its distance to the neighboring cell is less than that of the serving cell.

We evaluate our algorithm using widely used metrics including delay, jitter, Packet Delivery Ratio (PDR), and throughput. Specifically, jitter and delay are key indicators of network performance. Delay represents the overall time taken by a packet to travel through the network, while jitter pertains to the variability in packet arrival times. Factors such as propagation time, processing time at network devices, and queuing time contribute to delay, which is essential for ensuring efficient and dependable communication, particularly for applications requiring minimal latency [Gavrilovska et al. 2020]. On the other hand, PDR measures the success rate of packet receipt relative to the number sent, highlighting the effectiveness of traffic management in congested networks. Finally, throughput quantifies the volume of data, a network can successfully transmit over time, factors like bandwidth, signal quality, and network congestion play significant roles in a 5G network's data transfer capacity [Zhang et al. 2022, Riyanto et al. 2023]

### 4.2. Results

Figure 3 shows our PDR evaluation, which compares the traditional O-RAN with the O-RAN AI-based, which will show how the AI-based can deliver more packets until the networks are still more stable. The figure uses the number of users to evaluate the scenario with different quantities of clients and the number of packets transmitted simultaneously. We see a similar design in the first 50 clients. However, the traditional design shows a third quartile with low values for which the AI base does not fit, reaching the traditional approach reaching around 67%, while the O-RAN AI-based reach values around 97,5%, which means an approach with average values. The boxplot of 75 and 100 UEs also shows a constant improvement of PDF on our simulated network in the traditional UE, reaching values of around 91,5% and 86%, while the AI-based reaches around 93% and 87%. In that way, the evaluation demonstrates more unstable results in traditional methods with larger third quartiles and standard deviation.

Figure 4 provides a comparative analysis of throughput in Mbps across two Open

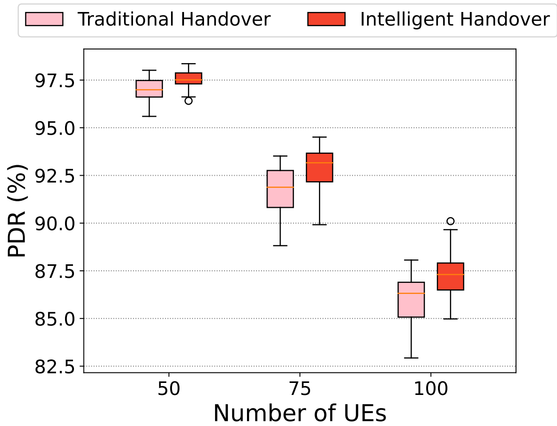


**Figure 3. PDR results for different number of mobile users**

RAN configurations as the number of UEs scales up. The AI-based O-RAN consistently outperforms the Traditional O-RAN, maintaining higher throughput as evidenced by the boxplot medians. With 50 UEs, the AI-based system shows a minor advantage, with throughput values tightly clustered around the 1.40 Mbps mark, while the Traditional system lingers slightly lower. At 75 UEs, the AI-based system's throughput remains above 1.35 Mbps, whereas the Traditional system begins to show a wider spread, indicating less consistency. This trend continues as we reach 100 UEs; the Traditional O-RAN's throughput shows increased variability and a lower median, with the spread below the 1.30 Mbps threshold, highlighting a noticeable decline in performance. In contrast, the AI-based system's throughput demonstrates resilience, with its lower quartile not dipping significantly, suggesting that even with increased load, the AI-based approach sustains superior throughput performance.

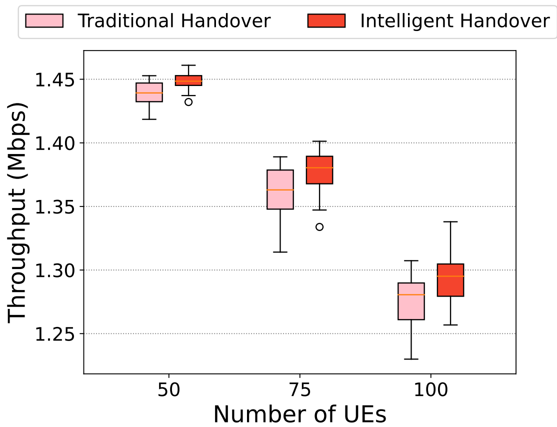


**Figure 4. Throughput results for different number of mobile users**

Our evaluation of delay in the network scenario highlights a comparison between the two approaches based on the number of users as presented in Figure 5. Our O-RAN AI-based approach generally shows a more stable scenario and minor average values. The boxplot evaluation shows how stable and dynamic a scenario with more clients would be than the traditional scenario, having a variety of around 12,6 and 17,4 ms with 75 users, while the AI-based has only 12,4 ms and 15,5 ms on average. That means a quality of signal with a more dynamic application that needs faster delivery of packages. Therefore, the results with 100 users are the same as those with 75 users. However, AI-based delay shows a larger third quartile to low values, which means this approach probably has a ratio of a minimal delay reaching 17,5 ms. In comparison, the traditional approach deviation reaches around 18,80 ms.

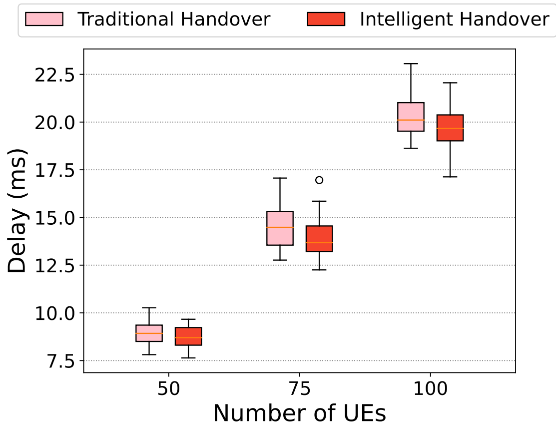


**Figure 5. Delay results for different number of mobile users**

Figure 6 compares jitter, measured in milliseconds, across traditional and AI-based O-RAN systems as the number of UEs increases. The AI-based approach consistently demonstrates lower jitter, with averages of approximately 12.4 ms for 75 users and a maximum jitter value of around 15.5 ms, signifying a stable and reliable network performance. In contrast, the traditional O-RAN exhibits more significant variability, with jitter ranging up to 18.8 ms for the same number of users. This difference is even more pronounced with 100 UEs, where the AI-based system's third quartile suggests a preferable delay profile, rarely exceeding 17.5 ms, compared to its traditional counterpart. These discrete values clearly illustrate the superior capability of the AI-based O-RAN in maintaining lower jitter and, consequently, a higher quality of service in network scenarios with a dense user base.

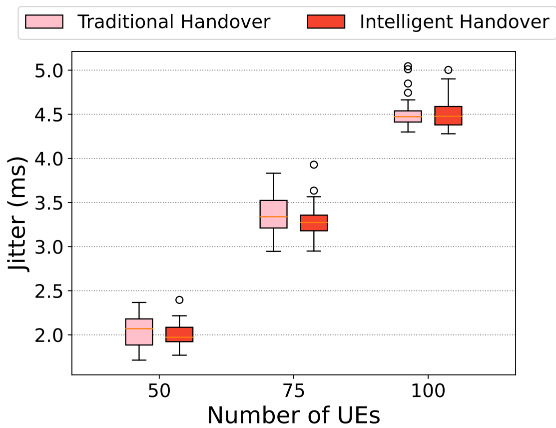


**Figure 6. Jitter results for different number of mobile users**

Figure 7 shows the handover frequency within traditional O-RAN and AI-based O-RAN by the simulation time. In the AI-based O-RAN, the number of handovers remains comparatively lower and more consistent, ranging from approximately 6 to 9 handovers. This consistency underscores a stable network performance, likely due to the AI's predictive capabilities in optimizing handover events. Conversely, the traditional O-RAN shows a broader fluctuation in handover frequency, with numbers spiking to as high as 14 handovers and a noticeable variability at different time stamps.

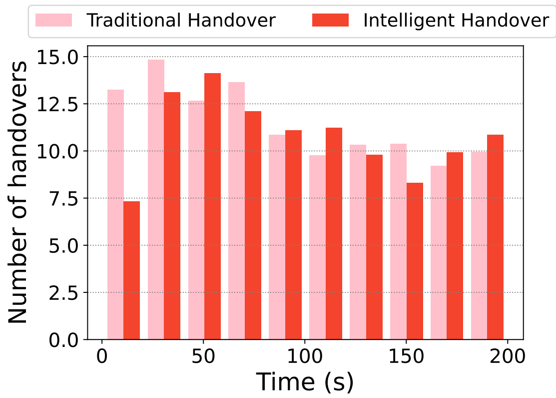


**Figure 7. Number of handovers results by the time of simulation**

## 5. Conclusion and Future Works

This paper evaluates the performance of O-RAN in a handover scenario involving up to one hundred users when the system is configured with both traditional and ML-based handover schemes. Otherwise, with a traditional handover method, our approach offers greater personalization to each xApp within the O-RAN framework, which is enabled by the near real-time RAN Intelligent Controller (RIC). The results reveal a noticeable enhancement in utilizing the O-RAN environment to establish a predictive model for UE positioning. Consequently, the relative improvement in PDR ratio means a better alignment with the client's location across each eNB.

The xApp brings with it considerable computational and programmable functions on RIC. However, the moment when the code calls the ML xApp also shows some issues with the behavior of the network signals. That means some problems for whom UE the eNB will link. The ML does not choose all the links between the UEs and eNBs simultaneously. That demonstrates the complexity of the behavior of the network resources during each handover.

In future works, we will explore innovative ML implementation strategies to enhance throughput and minimize delay. It includes integrating time series and other ML models renowned for their precision, such as Renet or Inception. Thus, enhancing the system could involve scaling up the distance range with simulation metrics that closely resemble real-world conditions, surpassing the constraints found in the literature [Lacava et al. 2023b]. Other approaches include a seamless handover protocol and a selection algorithm optimized with deep reinforcement learning. That can improve the average downlink data rate [Wang et al. 2021]. It entails incorporating more clients and eNBs and extending the distance range. Such enhancements refine the scenario and advance our efforts to provide better QoS support for mobile users in O-RAN environments.

## Acknowledgment

This study was financed in part by CNPq for the project entitle as “INCT Redes de Comunicação e Internet das Coisas Inteligentes (ICoNIoT) ”, and also by the project entitle as “Computação de Borda em um Mundo de IoT Inteligente com Suporte a Latência Ultra Baixa (EdgeIo2T) ”. In addition, This study was financed in part by São Paulo Research Foundation (FAPESP) for the grant 2023/00673-7.

## References

Alavirad, M., Hashmi, U. S., Mansour, M., Esswie, A., Atawia, R., Poitau, G., and Repeta, M. (2023). O-ran architecture, interfaces, and standardization: Study and application to user intelligent admission control. *Frontiers in Communications and Networks*, 4:1127039.

Baldesi, L., Restuccia, F., and Melodia, T. (2022). Charm: Nextg spectrum sharing through data-driven real-time o-ran dynamic control. In *IEEE Conference on Computer Communications (INFOCOM 2022)*, pages 240–249.

Bonati, L., D'Oro, S., Polese, M., Basagni, S., and Melodia, T. (2021). Intelligence and learning in o-ran for data-driven nextg cellular networks. *IEEE Communications Magazine*, 59(10):21–27.

Bonati, L., Polese, M., D'Oro, S., Basagni, S., and Melodia, T. (2022). Openran gym: An open toolbox for data collection and experimentation with ai in o-ran. In *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 518–523. IEEE.

Brik, B., Chergui, H., Zanzi, L., Devoti, F., Ksentini, A., Siddiqui, M. S., Costa-Pérez, X., and Verikoukis, C. (2023). A survey on explainable ai for 6g o-ran: Architecture, use cases, challenges and research directions. *arXiv preprint arXiv:2307.00319*.

Costa, A., Pacheco, L., Rosário, D., Villas, L., Loureiro, A. A., Sargento, S., and Cerqueira, E. (2020). Skipping-based handover algorithm for video distribution over ultra-dense vanet. *Computer Networks*, 176:107252.

de Oliveira, W., Batista Jr, J. O. R., Novais, T., Takashima, S. T., Stange, L. R., Martucci Jr, M., Cugnasca, C. E., and Bressan, G. (2023). Opencare5g: O-ran in private network for digital health applications. *Sensors*, 23(2):1047.

Garcia-Saavedra, A. and Costa-Perez, X. (2021). O-ran: Disrupting the virtualized ran ecosystem. *IEEE Communications Standards Magazine*, 5(4):96–103.

Gavrilovska, L., Rakovic, V., and Denkovski, D. (2020). From cloud ran to open ran. *Wireless Personal Communications*, 113:1523–1539.

Hamdan, M. Q., Lee, H., Triantafyllopoulou, D., Borralho, R., Kose, A., Amiri, E., Mulvey, D., Yu, W., Zitouni, R., Pozza, R., et al. (2023). Recent advances in machine learning for network automation in the o-ran. *Sensors*, 23(21):8792.

Lacava, A., Bordin, M., Polese, M., Sivaraj, R., Zugno, T., Cuomo, F., and Melodia, T. (2023a). ns-o-ran: Simulating o-ran 5g systems in ns-3. In *Proceedings of the 2023 Workshop on ns-3*, pages 35–44.

Lacava, A., Polese, M., Sivaraj, R., Soundrarajan, R., Bhati, B. S., Singh, T., Zugno, T., Cuomo, F., and Melodia, T. (2023b). Programmable and customized intelligence for traffic steering in 5g networks using open ran architectures. *IEEE Transactions on Mobile Computing*.

Linsalata, F., Moro, E., Gjeci, F., Magarini, M., Spagnolini, U., and Capone, A. (2024). Addressing control challenges in vehicular networks through o-ran: a novel architecture and simulation framework. *IEEE Transactions on Vehicular Technology*.

Musa, A. A., Hussaini, A., Qian, C., Guo, Y., and Yu, W. (2023). Open radio access networks for smart iot systems: State of art and future directions. *Future Internet*, 15(12):380.

O-RAN Alliance (2021). O-ran alliance introduces 48 new specifications released since july 2021.

Polese, M., Bonati, L., D'oro, S., Basagni, S., and Melodia, T. (2023). Understanding o-ran: Architecture, interfaces, algorithms, security, and research challenges. *IEEE Communications Surveys & Tutorials*.

Riccio, E. L., Mangipudi, P. K., and McNair, J. (2023). O-ran signaling optimizations for improved iot handover performance in 5g networks. In *Proceedings of the 8th ACM/IEEE Conference on Internet of Things Design and Implementation*, pages 454–455.

Riyanto, Q. T., Astuti, R. P., and Nashiruddin, M. I. (2023). Techno-economic analysis of 5g network using open ran technology at 2.3 ghz frequency for dense urban. In *2023 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, pages 132–137.

Sahbafard, A., Schmidt, R., Kaltenberger, F., Springer, A., and Bernhard, H.-P. (2023). On the performance of an indoor open-source 5g standalone deployment. In *2023 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6.

Sun, Y., Jiang, W., Feng, G., Klaine, P. V., Zhang, L., Imran, M. A., and Liang, Y.-C. (2020). Efficient handover mechanism for radio access network slicing by exploiting distributed learning. *IEEE Transactions on Network and Service Management*, 17(4):2620–2633.

Wang, L., Han, D., Zhang, M., Wang, D., and Zhang, Z. (2021). Deep reinforcement learning-based adaptive handover mechanism for vlc in a hybrid 6g network architecture. *IEEE Access*, 9:87241–87250.

Zangooei, M., Golkarifard, M., Rouili, M., Saha, N., and Boutaba, R. (2023). Flexible ran slicing in open ran with constrained multi-agent reinforcement learning. *IEEE Journal on Selected Areas in Communications*.

Zhang, H., Zhou, H., and Erol-Kantarci, M. (2022). Team learning-based resource allocation for open radio access network (o-ran). In *IEEE International Conference on Communications (ICC)*, pages 4938–4943.