Research Plan

Department of Electrical and Computer Engineering

University of Oulu

Provisional title:
Dynamic Resource Allocation in O-RAN Architecture using Network Slicing

Candidate name
Mojdeh Karbalaee Motalleb

Supervisor name:
Dr. Onel Lopez

Abstract

Open radio access network (O-RAN) is the next generation of RAN systems introduced by the O-RAN Alliance to increase the flexibility and the openness of the network systems and simultaneously decrease CAPEX and OPEX of mobile operators. The O-RAN system combines and takes advantage of Cloud RAN (C-RAN) and the Virtual RAN (VRAN). O-RAN separates RAN into three different units, namely Radio Unit (O-RU), Distributed Unit (O-DU), and Central Unit (O-CU). In this research, we study the problem of baseband resource allocation and virtual network function (VNF) activation in O-RAN architecture based on their service priority for different types of 5G services including enhanced mobile broadband (eMBB), ultra-reliable low latency communications (uRLLC) and massive Machine Type Communications (mMTC) services using intelligent RAN slicing. Network slicing is a promising solution for supporting such heterogeneous and challenging services. Deploying network slicing, the isolation of different types of services in O-DU, O-CU, and user plane function (UPF) is performed. The limited fronthaul capacity, the restriction of end-to-end delay and the reliability of the services are considered at the same time.

The optimization of baseband resources includes O-RU assignment, the assignment of physical resource block (PRB), and power allocation. The main problem is a mixed-integer non-linear programming problem that is tremendously difficult. Since the problem has a two-time scale, it can be broken into two layers. On the large-time scale, the problem of obtaining the optimal number of VNF, the placement of VNFs, and the assignment of PRB to the slices are considered. On the small-time scale, there is a problem of obtaining the optimal power of UEs, and the PRB and O-RU assignment. To solve the problem, we propose deep reinforcement learning, the deep deterministic policy gradient (DDPG) method, and convex optimization.

1 Introduction

O-RAN, as the integration and expansion of C-RAN and xRAN, or C-RAN and vRAN is expected to be a key technology in 5G networks to enhance the RAN performance extensively. The core idea of C-RAN is to divide the radio remote head (RRU) from the baseband unit (BBU). Also, several BBUs are placed together to create the BBU-pool, providing unified baseband signal processing with powerful computing capabilities. On the other hand, xRAN technology has three fundamental features. The control plane is decoupled from the user plane. Besides, a modular eNB software stack is built to operate on common-off-the-shelf (COTS) hardware. Moreover, open north-bound and south-bound interfaces are introduced. O-RAN separate RAN into three different units, namely Radio Unit (O-RU), Distributed Unit (O-DU), and Central Unit (O-CU). O-RU is a logical node that contains RF and lowers PHY. Moreover, the O-DU expresses another logical node that includes higher PHY, MAC, and RLC. In addition, the O-CU depicts the logical node contains two parts, which are the O-CU user plane (O-CU-UP) and O-CU control plane (O-CU-CP). O-CU-UP hosts PDCP-UP and SDAP, and O-CU-CP hosts PDCP-CP and RRC. O-DU and O-CU are connected via an open and well-defined interface F1. Moreover, O-DU is connected to a radio unit (O-RU) with an open fronthaul interface. The architecture of O-RAN contains other principal logical nodes called Orchestration and Automation, RAN Intelligent Controller (RIC)- Near Real-Time and O-Cloud. One of the necessities of the new generation of wireless networks is its intelligence. Based on the requirement of an intelligent wireless network, O-RAN offers machine learning techniques. The two logical nodes RIC-Non Real-Time (which is placed in Orchestration and Automation node) and RIC- Near Real Time, implement the algorithms for network intelligence [1–7].

One of the goals of the next generation of mobile systems is to strictly meet the stringent QoS demands of the different services introduced in 5G , i.e., eMBB, URLLC, and mMTC. Network slicing is a promising solution to obtain the QoS for each type of service. Network slicing contains RAN slicing, core slicing, and both of them together. Efficient slicing of RAN is still challenging due to time-varying network circumstances. Network slicing creates considerable complexity in system performance and makes the conventional mathematical methods insufficient to model the network. Therefore, it motivates researchers to realize network slicing using machine learning strategies such as deep learning and deep reinforcement learning. So, the network can reach the best control policy according to its experience in the past time slots. As mentioned above, network slicing, which contains RAN slicing, core slicing, and the whole, is considered a hot topic for vendors. In addition, the O-RAN system seems to be the next RAN architecture. So, this motivated researchers to study the dynamic RAN slicing in O-RAN architecture to achieve the desired QoS for each service type. In addition, according to the high complexity of RAN slicing and incapability of the traditional model-based methods, dynamic machine learning becomes the best solution for these problems. Generally, there are two kinds of deep reinforcement learning methods for solving dynamic problems. The first class is value-based such as Deep-Q-Network, and the second class is policy-based problems such as DDPG. The first class can only solve the discrete action space problems, which are integer programming. The second class can solve by searching the optimal policy, which is the actor-critic method. Here, we use the second method to solve our problems [8–10].

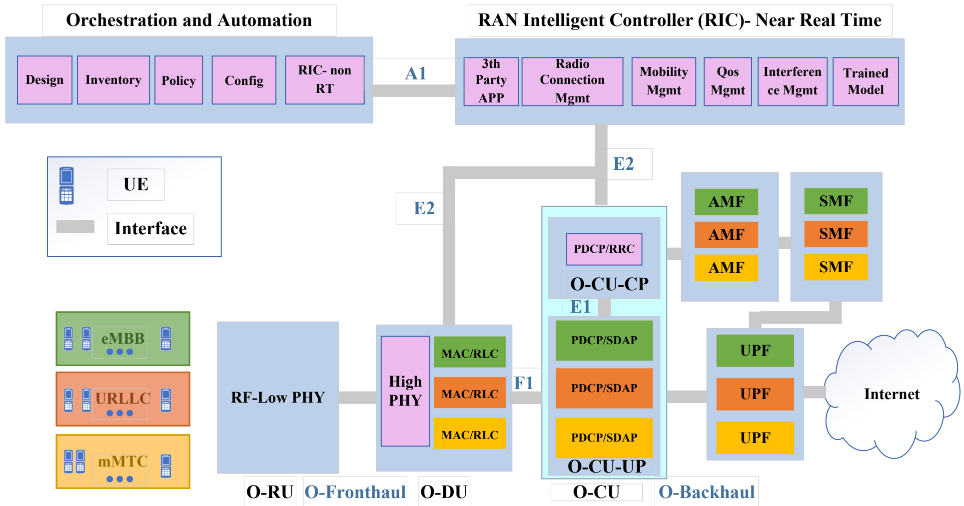

Figure 1: Network sliced ORAN system

1.1 Related Works

In [10], the authors proposes a new intelligent RAN slicing method with two-layered control granularity, which aims at maximizing both the long-term QoS of services and spectrum efficiency (SE) of slices. In this paper, the authors design a RAN slicing strategy with multiple time and resource granularities to accommodate the time-varying network conditions and diverse QoS requirements. Here, the DDPG method is applied for the lower level and deep Q-network is applied for upper level.

In paper [11], an actor-critic methods and a continuous model-free deep reinforcement learning (DRL) method is applied to minimize energy consumption and virtual network function (VNF) instantiation cost. Multiplexing eMBB and uRLLC services on the same RAN and sharing the resources of these services is challenging, and many researchers pay attention to this issue. In [12, 13], the problem of resource allocation in the coexistence of URLLC and eMBB services is considered based on their QoS. Paper [8], represents a joint convex method and deep reinforcement learning to schedule both eMBB and URLLC services. Here we have two time-scale, a two-phase-framework, including eMBB resource allocation and URLLC scheduling phases. In [14], the problem of power minimization for uRLLC and eMBB services is presented for non-orthogonal multiple access (NOMA) and orthogonal multiple access (OMA). In [15], the authors proposed to allocate RAN resources for the network slicing system in coexistence of eMBB and URLLC services. The system guarantees the latency, the service rate, and the maintenance of reliability.

1.2 Main Contributions

In this research, as depicted in Figure 1, we aim at solving the problem of dynamic network slicing in the O-RAN system. Here, we study a single-cell downlink situation involving multiple network slices that share existing radio sources. The main contributions of this research are summarized as follow:

- The system has a two-time scale and should be solved in two layers. On a large-time scale, the problem of obtaining an optimal number of VNFs and the VNF placement is performed. Also, the assignment of PRB to the slices is obtained. On a small-time scale, the problem of power allocation, O-RU association and the assignment of PRB to UE in each slice is applied.

- In this research, we consider three types of services that requires specific QoS: URLLC, which requires low latency and high reliability; eMBB, which needs high data rate; and mMTC, that are massive devices with short packets and low power devices. The problem of RAN slicing for different types of services is studied in this research.

- In this research, we consider an intelligent resource allocation in the O-RAN architecture for the different scenarios of services. Since, the conventional models cannot perform well in the system because of the different QoS for each service type, and the complexity of dynamic RAN slicing. So we must switch to machine learning and dynamic methods and find the most suitable approach. A dynamic strategy of resource allocation is required to solve this problem to achieve the specific QoS for each type of service in the O-RAN architecture. In the small-time scale, the actor-critic algorithm such as DDPG is applied to the system that is based on control policy search. This method directly searches the optimal control policy by estimating of the gradient with respect to the parameters of the control policy. In the large-time scale, the deep reinforcement learning is implemented which is a value based algorithm.

- Also, another goal of this project is to apply transfer learning to the system and enhance the performance and the convergence of the methods.

2 System Model

In this section, first, we present the system model. Then, we obtain achievable data rates and delays for the downlink (DL) of the ORAN system. Afterward, we discuss about assignment of physical data center resources. Finally, the main problem is expressed.

2.1 System Model

Suppose we have three service types includes eMBB, URLLC, and mMTC. Assume we have S_1, S_2, and S_3 different applications for the first, second and third service type, respectively ($S = S_1+S_2+S_3$). Assume we have S preallocated slices serving these S services; There are S_1 slices for the first service type (eMBB), S_2 slices for the second service type (URLLC), and S_3 slices for the third service type (mMTC). Each Service $s_j \in \{1, 2, ..., S_j\}$ consists of U_{s_j} request from the single-antenna UEs which require certain QoS to be able to use the requested program($j \in \{1, 2, 3\}$ indicate service type). There are different application request which fall into one of these service categories. Each application request requires specific QoS. Based on the request for the application and QoS, UE may be admitted and allocated to the resources. Assume each slice $s_j \in \{1, 2, ..., S_j\}$, $j \in \{1, 2, 3\}$ consists of K_{s_j}, $j \in \{1, 2, 3\}$ pre-allocated physical resource blocks (PRBs) that obtained in the large-time scale, M_s^d VNFs for the processing of O-DU, M_s^c VNFs for the processing of O-CU-UP and M_s^u VNFs for the processing of UPF.

Also, each VNF instance is running on the virtual machine (VM) that are using resources from the data centers. Each VM, requires enough resources of CPU, memory, storage and network bandwidth.

In addition, there are R multi-antenna RU that are shared between slices. Each RU $r \in \{1, 2, ..., R\}$ has J antenna for transmitting and receiving data. Moreover, all RUs, have access to PRBs.

2.2 The Achievable Rate

The SNR of i^{th} UE requesting served at slice s on PRB k is obtained from

$$\rho^k_{r,u(s,i)} = \frac{|p^k_{r,u(s,i)} \mathbf{h}^{H\,k}_{r,u(s,i)} \mathbf{w}^k_{r,u(s,i)} g^r_{u(s,i)}|^2}{BN_0}, \tag{1}$$

where $p^k_{r,u(s,i)}$ represents the transmission power from o-RU r to i^{th} UE served at slice s on PRB k. $\mathbf{h}^k_{r,u(s,i)} \in \mathbb{C}^J$ is the vector of channel gain of a wireless link from r^{th} RU to the i^{th} UE in s^{th} slice. In addition, $\mathbf{w}^k_{r,u(s,i)} \in \mathbb{C}^J$ depicts the transmit beamforming vector from r^{th} RU to the i^{th} UE in s^{th} slice that is the zero forcing beamforming vector to minimize the interference which is indicated as below

$$\mathbf{w}^k_{r,u(s,i)} = \mathbf{h}^k_{r,u(s,i)} (\mathbf{h}^{H\,k}_{r,u(s,i)} \mathbf{h}^k_{r,u(s,i)})^{-1} \tag{2}$$

Moreover, $g^r_{u(s,i)} \in \{0, 1\}$ is a binary variable that illustrates whether RU r is mapped to the i^{th} UE allocate to s^{th} slice or not. Also, BN_0 denotes the power of Gaussian additive noise. Also, $a_{u(s,i)} \in \{0, 1\}$ is a binary variable to depict user admission. $e^k_{r,u(s,i)}$ is the binary variable to show whether the k^{th} PRB is allocated to the UE i in slice s, assigned to r^{th} o-RU.

The achievable data rate for the i^{th} UE request in the s_1^{th} application of service type 1 (eMBB) can be written as $\mathcal{R}^e_{u(s_1,i)}$.

$$\begin{aligned} \mathcal{R}^{e,r}_{u(s_1,i)} &= \sum_{k=1}^{K_{s_1}} B \log_2(1 + \rho^k_{r,u(s_1,i)}) a_{u(s_1,i)} e^k_{r,u(s_1,i)}, \\ \mathcal{R}^e_{u(s_1,i)} &= \sum_{r=1}^{R} \mathcal{R}^{e,r}_{u(s_1,i)} \end{aligned} \tag{3}$$

where B is the bandwidth of system. $\mathcal{R}^{e,r}_{u(s_1,i)}$ is the achievable rate of each RU r to UE i in slice s_1. Since the blocklength in URLLC and mMTC is finite, the achievable data rate for the i^{th} UE request in the s_j^{th}, $j \in \{2,3\}$ application of service type 2 (URLLC) and 3 (mMTC) is not achieved from Shannon Capacity formula. So, for the short packet transmission the achievable data rate is approximated from follow

$$\begin{aligned}
\mathcal{R}^{\mathfrak{u},r}_{u(s_2,i)} &= \sum_{k=1}^{K_{s_2}} B(\log_2(1+\rho^k_{u(s_2,i)}) - \zeta^k_{u(s_2,i)})\beta^k_{u(s_2,i)}, \quad \mathfrak{u} \in \{u, m\} \\
\mathcal{R}^{\mathfrak{u}}_{u(s_1,i)} &= \sum_{r=1}^{R} \mathcal{R}^{e,r}_{u(s_2,i)}
\end{aligned} \tag{4}$$

Where $\beta^k_{u(s_2,i)} = a_{u(s_2,i)} e^k_{u(s_2,i)}$ and $\zeta^k_{u(s_2,i)} = log_2(e)Q^{-1}(\epsilon)\sqrt{\frac{C^k_{u(s_2,i)}}{N^k_{u(s_2,i)}}})$ Where, ϵ is the transmission probability, Q^{-1} is the inverse of Q- function (Gaussian), $C^k_{u(s_2,i)} = 1 - \frac{1}{(1+\rho^k_{u(s_2,i)})}$ depicts the channel dispersion of UE i at slice s_2, experiencing PRB k and $N^k_{u(s_2,i)}$ represents the blocklength of it. $\mathcal{R}^{e,r}_{u(s_1,i)}$ is the achievable rate of each RU r to UE i in slice s_2.

2.3 Mean Delay

In this part, the end to end mean delay for a service is obtained. Suppose the mean total delay is depicted as T_{tot}.

$$\begin{aligned}
T_{tot} &= T_{process} + T_{transmission} + T_{propagation} \\
T_{process} &= T_{RU} + T_{DU} + T_{CU} + T_{UPF} \\
T_{transmission} &= T_{front} + T_{mid} + T_{back} + T_{trans2net} \\
T_{propagation} &= T_{front} + T_{mid} + T_{back} + T_{trans2net}
\end{aligned} \tag{5}$$

Total delay is sum of processing delay, transmission delay and propagation delay. The propagation delay is the time takes for a signal to reach to its destination. So it has a constant value based on the length of fiber link ($T = L/C_r$, where L is the length of link and C_r is the capacity of the link). Here we assume the value of propagation delay is negligible compared to the rest.

2.3.1 Processing Delay

Assume the packet arrival of UEs follows a Poisson process with arrival rate $\lambda_{u(s,i)}$ for the i^{th} UE of the s^{th} slice. Therefore, the mean arrival data rate of the s^{th} slice in the UPF layer is $\alpha^1_s = \sum_{u=1}^{U_s} a_{u(s,i)}\lambda_{u(s,i)}$, where $a_{u(s,i)}$ is a binary variable which indicates whether the i^{th} UE requested s^{th} service is admitted or not.

Assume the mean arrival data rate of the UPF layer for slice s (α^U_s) is approximately equal to the mean arrival data rate of the O-CU-UP layer (α^C_s) and O-DU (α^D_s). so $\alpha_s = \alpha^U_s \approx \alpha^C_s \approx \alpha^D_s$. since, by using Burkes Theorem, the mean arrival data rate of the second and third layer which are processed in the first layer is still Poisson with rate α_s. It is assumed that there are load balancers in each layer for each service to divide the incoming traffic to VNFs equally. Suppose the baseband processing of each VNF is depicted as M/M/1 processing queue. Each packet is processed by one of the VNFs of a slice. So, the mean delay for the s^{th} slice in the first and the second layer, modeled as M/M/1 queue, is formulated as follow, respectively

$$\begin{aligned}
T^s_{DU} &= \frac{1}{\mu_d - \alpha_s/M^d_s}, \\
T^s_{CU} &= \frac{1}{\mu_c - \alpha_s/M^c_s} \\
T^s_{UPF} &= \frac{1}{\mu_u - \alpha_s/M^u_s}
\end{aligned} \tag{6}$$

Where M^d_s, M^c_s and $M_s u$ are the variables that depict the sum of VNFs in O-DU, O-CU-UP and UPF, respectively. Moreover, $1/\mu_d$, $1/\mu_c$ and $1/\mu_u$ are the mean service time of the O-DU, O-CU and the UPF layers respectively. Besides, α_s is the arrival rate which is divided by load balancer before arriving to the VNFs. The arrival rate of each VNF in each layer for each slice s is α_s/M^i_s $i \in \{d, c, u\}$.

In addition, T^s_{RU} is the mean transmission delay of s^{th} slice on the wireless link. The arrival data rate of wireless link is equal to the arrival data rate of load balancers for each service. Moreover, it is assumed that the service time of transmission queue for each slice s has an exponential distribution with mean $1/(R_{tot_s})$ and can be modeled as a M/M/1 queue. Therefore, the mean delay of the transmission layer is

$$T^s_{RU} = \frac{1}{R_{tot_s} - \alpha_s}; \tag{7}$$

where, $R_{tot_s} = \sum_{u=1}^{U_s} a_{u(s,i)} R_{u(s,i)}$ is the total achievable rate of each service. So the mean processing delay for each UE in slice s is

$$T^s_{process} = T^s_{RU} + T^s_{DU} + T^s_{CU} + T^s_{UPF} \tag{8}$$

2.3.2 Transmission Delay

The transmission delay is the amount of time required to push all the packets into the fiber link. Here, we have transmission delay in fronthaul, midhaul, backhaul and the link to transmit data to internet.

$$\begin{aligned} T_{front} =& \frac{\alpha^f_s}{R_f} \\ T_{mid} =& \frac{\alpha^m_s}{R_m} \\ T_{back} =& \frac{\alpha^b_s}{R_b} \\ T_{trans2net} =& \frac{\alpha^t_s}{R_t} \end{aligned} \tag{9}$$

Where, R_f, R_m, R_b and R_t are the rate of transmission in fronthaul, midhaul, backhaul and the link to transmit data to internet, respectively. Furthermore, the mean arrival data rate of the each link (α^i_s , $i \in \{f, m, b, t\}$) is approximately equal to others ($\alpha_s \approx \alpha^i_s$, $i \in \{f, m, b, t\}$).

2.4 Reliability of URLLC

As we know, UEs request URLLC services, require services with low latency. For the M/M/1 system, the probability of the delay for each application s in the UPF, CU, DU and RU is as follow, respectively

$$\begin{aligned} P_r\{T^s_{UPF} \geq T^{max}_{UPF}\} &= e^{-(\mu_u - \alpha_s/M^u_s)T^{max}_{UPF}} \\ P_r\{T^s_{CU} \geq T^{max}_{CU}\} &= e^{-(\mu_c - \alpha_s/M^c_s)T^{max}_{CU}} \\ P_r\{T^s_{DU} \geq T^{max}_{DU}\} &= e^{-(\mu_d - \alpha_s/M^d_s)T^{max}_{DU}} \\ P_r\{T^s_{RU} \geq T^{max}_{RU}\} &= e^{-(R_{tot_s} - \alpha_s)T^{max}_{RU}} \end{aligned} \tag{10}$$

So the probability of coincidence of these events is as follow

$$P_r\{E_1, E_2, E_3, E_4\} = A_1 A_2 A_3 A_4, \tag{11}$$

Where $E_1 = T^s_{UPF} \geq T^{max}_{UPF}$, $E_2 = T^s_{CU} \geq T^{max}_{CU}$, $E_3 = T^s_{DU} \geq T^{max}_{DU}$ and $E_4 = T^s_{RU} \geq T^{max}_{RU}$. Also $A_1 = e^{-(\mu_u - \alpha_s/M^u_s)T^{max}_{UPF}}$, $A_2 = e^{-(\mu_c - \alpha_s/M^c_s)T^{max}_{CU}}$, $A_3 = e^{-(\mu_d - \alpha_s/M^d_s)T^{max}_{DU}}$ and $A_4 = e^{-(R_{tot_s} - \alpha_s)T^{max}_{RU}}$.

2.5 Physical Data Center Resource

Each VNF requires physical resources that contain memory, storage, CPU and Network Bandwidth. Let the required resources for VNF f in slice s is represented by a tuple as

$$\bar{\Omega}^f_s = \{\Omega^f_{M,s}, \Omega^f_{S,s}, \Omega^f_{C,s}, \Omega^f_{N,s}\}, \tag{12}$$

where $\bar{\Omega}^f_s \in \mathbb{C}^4$ and $\Omega^f_{M,s}, \Omega^f_{S,s}, \Omega^f_{C,s}, \Omega^f_{N,s}$ indicate the amount of required memory, storage, CPU and and Network Bandwidth, respectively. Moreover, the total amount of required memory, storage, CPU and Network Bandwidth of all VNFs of a slice in DU, CU and UPF is defined as below, respectively

$$\bar{\Omega}_{\mathfrak{z},s}^{tot,d} = \sum_{f=1}^{M_s^d} \bar{\Omega}_{\mathfrak{z},s}^{f,d} \quad \mathfrak{z} \in \{M, S, C, N\}.$$
$$\bar{\Omega}_{\mathfrak{z},s}^{tot,c} = \sum_{f=1}^{M_s^c} \bar{\Omega}_{\mathfrak{z},s}^{f,c} \quad \mathfrak{z} \in \{M, S, C, N\}. \tag{13}$$
$$\bar{\Omega}_{\mathfrak{z},s}^{tot,u} = \sum_{f=1}^{M_s^u} \bar{\Omega}_{\mathfrak{z},s}^{f,u} \quad \mathfrak{z} \in \{M, S, C, N\}.$$

Where, $\bar{\Omega}_{\mathfrak{z},s}^{f,d}$, $\bar{\Omega}_{\mathfrak{z},s}^{f,c}$ and $\bar{\Omega}_{\mathfrak{z},s}^{f,u}$ are the amount of resource that a VNF required in DU, CU and UPF, respectively.

$$\bar{\Omega}_{\mathfrak{z},s}^{tot} = \bar{\Omega}_{\mathfrak{z},s}^{tot,d} + \bar{\Omega}_{\mathfrak{z},s}^{tot,c} + \bar{\Omega}_{\mathfrak{z},s}^{tot,u} \tag{14}$$

Also, there are D_c data centers (DC), serving the VNFs. Each DC contains several servers that supply VNF requirements. The amount of memory, storage, CPU and and Network Bandwidth is denoted by τ_{M_j}, τ_{S_j}, τ_{C_j} and τ_{N_j} for the j^{th} DC, respectively

$$\tau_j = \{\tau_{M_j}, \tau_{S_j}, \tau_{C_j}, \tau_{N_j}\},$$

In this system model, the assignment of physical DC resources to VNFs is considered. Let $y_{s,d}$ be a binary variable indicating whether the d^{th} DC is allocated the resources to the VNFs of s^{th} slice or not.

2.6 Power of the O-RU and the Fronthaul Capacity

Let P_r denote the power of the transmitted signal from the r^{th} O-RU to UEs served by it. From (1), we have,

$$P_r = \sum_{s=1}^{S} \sum_{k=1}^{K_s} \sum_{i=1}^{U_s} |\mathbf{w}_{r,u(s,i)}^k|^2 p_{r,u(s,i)}^k g_{u(s,i)}^r e_{r,u(s,i)}^k + \sigma_q^2. \tag{15}$$

Since we have a fiber link between O-RU and O-DU, the rate of users on the fronthual link between O-DU and the r^{th} O-RU is formulated as

$$C_r = \log \left(1 + \frac{\sum_{s=1}^{S} \sum_{k=1}^{K_s} \sum_{i=1}^{U_s} |\mathbf{w}_{r,u(s,i)}^k|^2 \alpha_{r,u(s,i)}^k}{\sigma_q^2}\right), \tag{16}$$

where $\alpha_{r,u(s,i)}^k = p_{r,u(s,i)}^k g_{u(s,i)}^r e_{r,u(s,i)}^k$ and σ_q^2 is the power of quantization noise.

2.7 Problem Statement

Assume the power consumption of baseband processing at each DC d that is connected to VNFs of a slice s is depicted as ϕ_s. So the total power of the system for all active DCs that are connected to slices can be represented as

$$\phi_{tot} = \textstyle\sum_{s=1}^{S} \phi_s + \sum_{d=1}^{D_c} z_d \psi_d.$$

Where,z_d is shown that whether the d^{th} DC is turned on or not and ψ_d is a static cost when a DC is active.

$$z_d = \begin{cases} 1 & \sum_{s=1}^{S} y_{s,d} \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

In addition, ϕ_s is obtained from below

$$\phi_s = M_s^u \phi_s^u + M_s^c \phi_s^c + M_s^d \phi_s^d \tag{18}$$

Where, ϕ_s^u, ϕ_s^c and ϕ_s^d are the static cost of energy in UPF, CU and DU, respectively. So the optimization problem is formulated as follow. In the large-time scale the aim of the paper is to minimize the power of the system with the presence constraints

$$\min_{\boldsymbol{M},\boldsymbol{Y},\boldsymbol{E}} \phi_{tot} \tag{19a}$$
$$\text{subject to} \quad P_r \leq P_{max} \quad \forall r \tag{19b}$$

$$\mathcal{R}_{u_{(s_1,i)}} \geq \mathcal{R}^s_{min} \quad \forall s, \tag{19c}$$

$$\phi_{tot} \leq \phi_{max}, \tag{19d}$$

$$\sum_{s=1}^{S}\sum_{i=1}^{U_s} R^r_{u_{(s,k)}} \leq C^r_{max} \quad \forall r, \tag{19e}$$

$$\textstyle\sum_{s=1}^{S} y_{s,d}\bar{\Omega}^{tot}_{\mathfrak{z},s} \leq \tau_{\mathfrak{z}d} \quad \forall d, \forall \mathfrak{z} \in \mathcal{E}; \tag{19f}$$

In the small-time scale the aim of this paper is to maximize the sum rate of UEs of eMBB services, minimum the delay of UEs of URLLC services, and minimum the power mMTC services with the presence of constraints.

$$\max_{\boldsymbol{P},\boldsymbol{E},\boldsymbol{G}} \quad \sum_{s=1}^{S_1}\sum_{i=1}^{U_s} R^e_{u_{(s,k)}} - \sum_{s=1}^{S_2} T^s_{tot} - \sum_{s=1}^{S_3}\sum_{i=1}^{U_s} p_{u_{(s,k)}} \tag{20a}$$

$$\text{subject to} \quad P_r \leq P_{max} \quad \forall r \tag{20b}$$

$$p^k_{r,u(s,i)} \geq 0 \quad \forall i, \forall r, \forall s, \forall k, \tag{20c}$$

$$\mathcal{R}_{u_{(s_1,i)}} \geq \mathcal{R}^s_{min} \quad \forall s, \tag{20d}$$

$$\sum_{s=1}^{S}\sum_{i=1}^{U_s} R^r_{u_{(s,k)}} \leq C^r_{max} \quad \forall r, \tag{20e}$$

$$T^s_{tot} \leq T^{max,s}_{tot} \quad \forall s, \tag{20f}$$

$$P_r\{E_1, E_2, E_3, E_4\} \leq \epsilon_s \quad \forall s_2, \tag{20g}$$

$$g^r_{u(s,i)} \leq g^r_{u(s,i)} \sum_{k=1}^{K_s} e^k_{r,u(s,i)} \quad \forall s, i, \tag{20h}$$

(20i)

where $\boldsymbol{P} = [p_{u(s,i)}]$ $\forall s, \forall i$, is the matrix of power for UEs, $\boldsymbol{E} = [e^k_{r,u(s,i)}]$ $\forall s, \forall i \forall r, \forall k$ indicate the binary variable for PRB association. Moreover, $\boldsymbol{G} = [g^r_{u(s,i)}]$ $\forall s, \forall i \forall r$ is a binary variable for O-RU association. Furthermore, $M = [M^d_s, M^c_s, M^u_s]$ $\forall s$ is the matrix that shown the number of VNFs in each layer of slice and $\boldsymbol{Y} = [y_{s,d}]$ $\forall s, \forall d$ is a binary variable shown whether the physical DC is mapped to a VNFs of a slice or not. Also, η is weighted variable to value between the benefit and the cost term of objective function. (20b), and (20c), indicate that the power of each RU do not exceed the maximum power, and the power of each UE is a positive integer value, respectively. Also (20d) shows that the rate of each UE requesting eMBB, URLLC and mMTC is more than a threshold, respectively. (20e) and (20f) expressed the limited capacity of the fronthaul link, and the limited delay of receiving signal, respectively. (20f) is a reliability condition that the delay in each layer should be less than threshold. (20h) guarantee that if a UE in admitted by the system, O-RU and PRB is associated to it, respectively. In addition, (19d) indicate that the static cost of energy of VNFs in each slice do not exceed from the threshold. Moreover, in (19f) $\mathcal{E} = \{M, S, C, N\}$ and the constraint supports that we have enough physical resources for VNFs of each slice.

3 Proposed Algorithm and Numerical Results

In this research, we aim to use the machine learning methods such as deep reinforcement learning and deep learning to train the O-RAN system and have an intelligent system. The deep Q-learning is implemented for the large-time scale and the multi-agent deep reinforcement learning contains DDPG (actor-critic algorithm), correlated q-learning, and the priority proportional fairness algorithm will be implemented for the small-time scale. The deep learning algorithms contain LSTM and recurrent neural networks. Also, transfer learning is an exciting way to enhance the performance and convergence of the system. In this following, the proposed algorithm for the large-time scale is implemented, and part of the numerical results is depicted.

3.1 Proposed Algorithm

Here we use the reinforcement learning method to solve the above two problems. In the Q-learning method for the large-time scale, an agent tries to find the optimal value in a specific environment. This interactive process is modeled as a Markov decision-making process that includes (S, A, R, P, γ). S represents the state space matrix, and A represents the action vector. R is also the reward of action. $P(.|S, a)$ is the probability of transfer and $\gamma\ in(0, 1]$ is the discount factor. The $\Pi(.|S)$ policy is a mapping of the state to the distribution of actions. The value-state function for state s under policy $\Pi(.|S)$ with $V^{\Pi}(s)$ indicates that the expected return value in state s under policy $\Pi(.|S)$). The value of performing operation a in state s under the $\Pi(.|S)$ policy is represented by $Q^{\Pi}(s,a)$. We have the following relations on this basis.

$$V^{\Pi}(s) = \mathbb{E}_{\Pi,P}[\sum_{t=0}^{\infty} \gamma^t R_t | S_0 = s] \tag{21}$$

The Q-value is as below.

$$Q^{\Pi}(s,a) = \mathbb{E}_{\Pi,P}[\sum_{t=0}^{\infty} \gamma^t R_t | S_0 = s, A_0 = a]. \tag{22}$$

$\mathbb{E}$ represented the statistical average. Based on the Bellman equation we have

$$V^{\Pi}(s) = \mathbb{E}_{\Pi,P}[R + \gamma V^{\Pi}(s')] \tag{23}$$

and also,

$$Q^{\Pi}(s,a) = \mathbb{E}_{\Pi,P}[R + \gamma Q^{\Pi}(s',a')] \tag{24}$$

Where, s' and a' can be obtained from $\Pi(.|s')$ and $P(.|s,a)$. The goal of reinforcement learnin is to obtain the optimal policy to maximize the $Q^{\Pi}(s,a)$. So, using Bellman equation, we have

$$Q^{*}(s,a) = \mathbb{E}_{\Pi^{*},P}[R + \gamma Q^{*}(s',a')]. \tag{25}$$

Also, T^{*} is the Bellman operator

$$T^{*}Q(s,a) = \mathbb{E}_{\Pi^{*},P}[R + \gamma Q(s',a')] \tag{26}$$

By using this operator iteratively, $Q_{t+1}(s,a) \leftarrow T^{*}Q(s,a)$ the algorithm can converge $Q_t(s,a) \rightarrow Q^{*}(s,a)$. $t \rightarrow \infty$ [16, 17]. In the Q-learning in each episode we have

$$Q(s_{t+1},a_{t+1}) = Q(s_t,a_t) + \alpha[r_{t+1} + \gamma \max_{a \in A} Q(s_{t+1},a) - Q(s_t,a_t)] \tag{27}$$

where, α is the learning rate.

3.2 Numerical Results

Here, we simplify the large-time scale problem and indicate the numerical results. The numerical results for this problem (which is the VNF placement) can be shown as follow.

$$\min_{\boldsymbol{Y}} \quad \psi_{tot}(\boldsymbol{Y})(t) \tag{28a}$$

$$\text{s. t.} \quad \textstyle\sum_{d=1}^{D_c} y_{s,d}(t) \geq 1 \quad \forall s, \tag{28b}$$

$$\textstyle\sum_{s=1}^{S} y_{s,d}(t)\bar{\Omega}_s^{tot} \leq \tau_d \quad \forall d, \forall; \tag{28c}$$

Suppose we have just one limited resource (CPU in each server). The ratio of the number of the used-servers with the optimal method to the used-servers with the Q-learning method based on time spent can be shown as figure 2. In this figure, after about 900 epochs, the algorithm converges to the 0.9 of optimal solution.

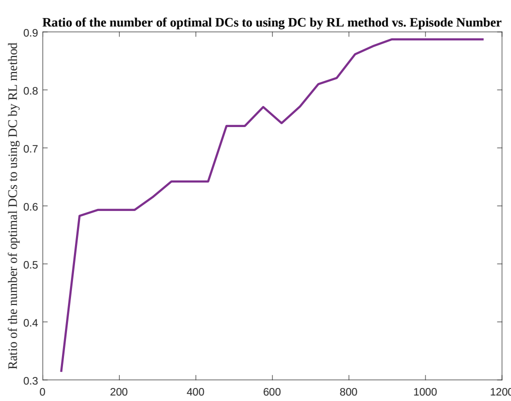

Figure 2: The ratio of the number of servers used with the optimal method to the servers used using the reinforcement learning method based on time spent

Table 1: Research Plan

Month	Tasks
First Month	Complete the System model for the Intelligent RAN slicing
Second Month	Find the best method for this system model
Third Month	Generalized this method for other system models
Fourth Month	Obtain Numerical Results
Fifth Month	Write the paper
Sixth Month	Review the paper

The figure 3, shows the ratio of the normalized cost to the average number of slices required in the optimal state and the reinforcement learning algorithm.

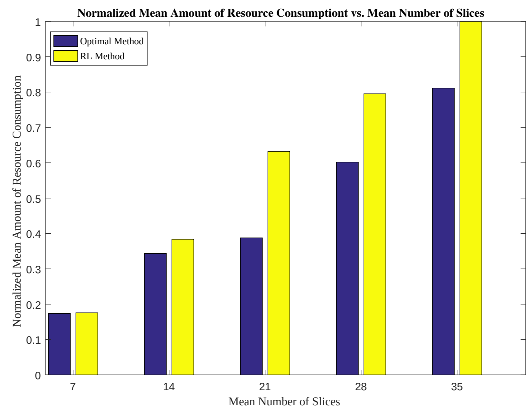

Figure 3: The ratio of the normalized cost to the average number of cuts required in the optimal state and the reinforcement learning algorithm

4 Research Plan

The table below outlines the goals and six-month planning overall.

References

[1] L. Gavrilovska, V. Rakovic, and D. Denkovski, "From cloud ran to open ran." *Wirel. Pers. Commun.*, vol. 113, no. 3, pp. 1523–1539, 2020.

[2] S. Niknam, A. Roy, H. S. Dhillon, S. Singh, R. Banerji, J. H. Reed, N. Saxena, and S. Yoon, "Intelligent o-ran for beyond 5g and 6g wireless networks," *arXiv preprint arXiv:2005.08374*, 2020.

[3] N. Kazemifard and V. Shah-Mansouri, "Minimum delay function placement and resource allocation for open ran (o-ran) 5g networks," *Computer Networks*, vol. 188, p. 107809, 2021.

[4] C. B. Both, J. Borges, L. Gonçalves, C. Nahum, C. Macedo, A. Klautau, and K. Cardoso, "System intelligence for uav-based mission critical with challenging 5g/b5g connectivity," *arXiv preprint arXiv:2102.02318*, 2021.

[5] "O-ran architecture description," O-RAN Alliance, Tech. Rep., 2020.

[6] O.-R. W. G. 2, "Ai/ml workflow description and requirements," O-RAN Alliance, Tech. Rep., 2020.

[7] B.-S. Lin, "Toward an ai-enabled o-ran-based and sdn/nfv-driven 5g& iot network era," *Network and Communication Technologies*, vol. 6, no. 1, pp. 6–15, 2021.

[8] M. Alsenwi, N. H. Tran, M. Bennis, S. R. Pandey, A. K. Bairagi, and C. S. Hong, "Intelligent resource slicing for embb and urllc coexistence in 5g and beyond: A deep reinforcement learning based approach," *IEEE Transactions on Wireless Communications*, 2021.

[9] M. Yan, G. Feng, J. Zhou, Y. Sun, and Y.-C. Liang, "Intelligent resource scheduling for 5g radio access network slicing," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7691–7703, 2019.

[10] J. Mei, X. Wang, K. Zheng, G. Boudreau, A. B. Sediq, and H. Abou-zeid, "Intelligent radio access network slicing for service provisioning in 6g: A hierarchical deep reinforcement learning approach," *IEEE Transactions on Communications*, 2021.

[11] F. Rezazadeh, H. Chergui, L. Christofi, and C. Verikoukis, "Actor-critic-based learning for zero-touch joint resource and energy control in network slicing," in *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021, pp. 1–6.

[12] M. Setayesh, S. Bahrami, and V. W. Wong, "Joint prb and power allocation for slicing embb and urllc services in 5g c-ran," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 2020, pp. 1–6.

[13] P. Yang, X. Xi, T. Q. Quek, J. Chen, X. Cao, and D. Wu, "How should i orchestrate resources of my slices for bursty urllc service provision?" *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 1134–1146, 2020.

[14] F. Saggese, M. Moretti, and P. Popovski, "Power minimization of downlink spectrum slicing for embb and urllc users," *arXiv preprint arXiv:2106.08847*, 2021.

[15] P. Korrai, E. Lagunas, S. K. Sharma, S. Chatzinotas, A. Bandi, and B. Ottersten, "A ran resource slicing mechanism for multiplexing of embb and urllc services in ofdma based 5g wireless networks," *IEEE Access*, vol. 8, pp. 45 674–45 688, 2020.

[16] P. R. Montague, "Reinforcement learning: an introduction, by sutton, rs and barto, ag," *Trends in cognitive sciences*, vol. 3, no. 9, p. 360, 1999.

[17] Y. Hua, R. Li, Z. Zhao, X. Chen, and H. Zhang, "Gan-powered deep distributional reinforcement learning for resource management in network slicing," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 2, pp. 334–349, 2019.