

Joint Power Allocation and Network Slicing In an End to End O-RAN System

1st Mojdeh Karbalaee Motalleb
Electrical and Computer Engineering
Tehran University
Tehran, Iran
mojdeh.karbalaee@ut.ac.ir

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

Abstract—Many major telecommunication companies confirmed the unification of the xRAN Group with the C-RAN Alliance to establish a global 'carrier-led' effort to bring greater transparency of Open-RAN (O-RAN), to the next-generation of radio access network.

To increase energy efficiency and optimize allocation of resources, Network Slicing (NS) is considered as a best method in 5G in order to virtualize the common physical network into several logical end-to-end networks. Every slice consists of a part of core network resources, network functions, and radio access network resources as a functional end-to-end network. In this paper, we elaborate joint NS in RAN and Core of O-RAN system and investigate optimal power of each User Equipment (UE) to jointly maximize Energy Efficiency and minimize consumption power of physical resources in a downlink channel. The problem is formulated as a mixed integer optimization problem that can be decompose into two independent sub-problems for RAN and Core since sub-problems are independent. Heuristic algorithms is proposed to each of sub-problems to solved the first sub-problem in order to simultaneously map slices to services and optimize power and the second sub-problem in order to map slices to physical resources and minimize number of active Data Centers (DC).

Index Terms—O-RAN, Network Slicing, Energy Efficiency, Data Center

I. Introduction

Recently, O-RAN which is the integration and expansion of C-RAN and xRAN is expected to be a key technology for 5G to enhance RAN performance and solve the challenges in a best way. The Idea of O-RAN comes from two opinion. Firstly, according to real-time analytic used for artificial intelligence system, the radio access networks must be evolved to be more open and smarter than previous generations. Furthermore, O-RAN can virtualize elements of the network with appropriate interfaces [1].

The core idea driving C-RAN is to split radio remote head (RRH) and baseband unit (BBU). Several BBUs operating on a cloud server will create a BBU-pool, providing unified baseband signal processing with powerful computing capabilities. Moreover, in O-RAN technology, this separation is implemented [2]. To communicate between BBU-pool and RRHs, fronthaul interface which is fiber link is assumed with limited capacity. Compression of message which is passed through these links, is a consequence of limited

fronthaul capacity [3], [4]. In BBU-Pool, using cloud-computing, enhance performance of system by virtualizing resources into virtual machines (VMs). Each VM has a computing processor which is mapped to Virtual Network Functions (VNFs) which is processed arrival data. Also VNFs are mapped to physical resource through NS technique [3], [5], [6].

xRAN technology ,released in April 2018 as a next generation of RAN, has three fundamental features. Control plane is decoupled from User plane. In addition, a modular eNB software stack is built to operate on common-off-the-shelf (COTS) hardware. Moreover, open north-bound and south-bound interfaces is introduced [7].

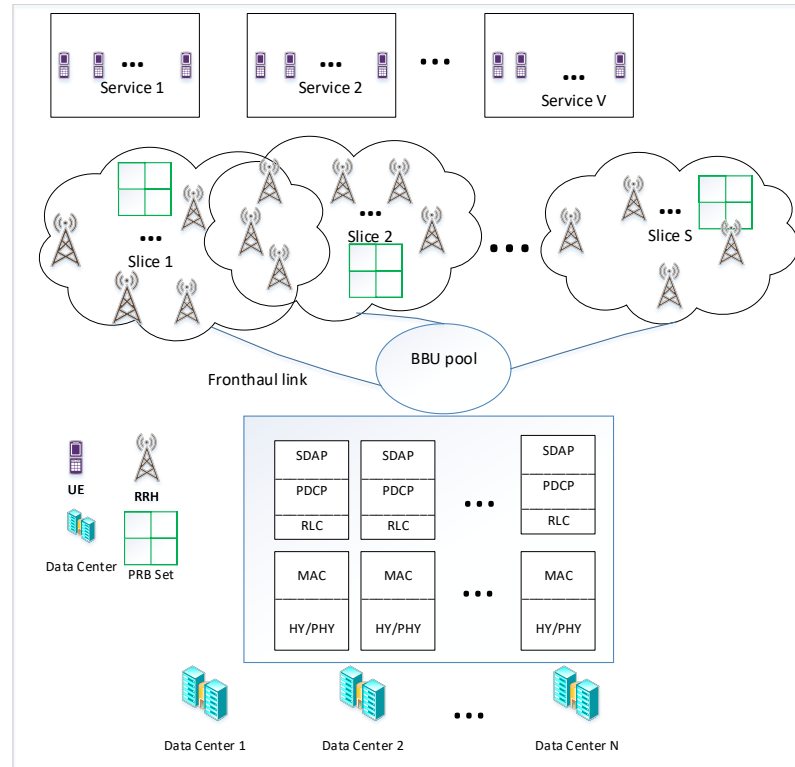


Fig. 1: Network sliced O-RAN system

Responsibility of wireless systems of the fifth generation

covers a wide range types of services. In order to provide the requirement of these services, NS is implemented to virtualize the common physical network into several logical end-to-end networks. Three different types of NS is introduced in [8] contains Core Slicing, RAN Slicing and Core-RAN Slicing. In Core-RAN Slicing, each slice of RAN is mapped to slices of Core.

In this paper, as mentioned in figure (2), the downlink of O-RAN system is assumed. UEs are divided to different groups according to their service requirements. Also RAN is decoupled to slices to provide services needs. Optimal power allocation and joint mapping slices to services is applied. In addition, mapping slices to physical resources is considered too.

II. System Model and Problem Formulation

In this section, first, we present the downlink (DL) of O-RAN System. Then we obtain achievable rate and delays. Afterward, the main problem is expressed.

A. System Model

Suppose that there are S slices Serving V services. Each Service $v \in \{1, 2, \dots, V\}$, consists of U_v single antenna users that require certain service. Each slice $s \in \{1, 2, \dots, S\}$ consists of R_s RRHs and N_s PRBs. All the RRHs in a slice that is mapped to a service, transmit signals to all the UEs in specific service. Each RRH $r \in \{1, 2, \dots, R\}$ is connected to BBU pool via an optical fiber link with limited fronthaul capacity. Also each RRH and PRB can serve more than one slice. It is considered that in BBU, the system has two processing layer consists of M_1 homogeneous VMs in first layer and M_2 homogeneous VMs in second layer.

B. Achievable Rate

In this subsection, the Achievable Rate is obtained as below. The achievable data rate for i^{th} UE in v^{th} service can be written as

$$\mathcal{R}_{u(v,i)} = B \log_2(1 + \rho_{u(v,i)}) \quad (1)$$

where B is the bandwidth of system and $\rho_{u(v,i)}$ is the SNR of i^{th} UE in v^{th} service which is obtained from

$$\rho_{u(v,i)} = \frac{p_{u(v,i)} \sum_{s=1}^{N_s} |\mathbf{h}_{R_s, u(v,i)}^H \mathbf{w}_{R_s, u(v,i)}|^2 a_{vs}}{BN_0 + I_{u(v,i)}} \quad (2)$$

Where, $p_{u(v,i)}$ represents the transmitted power allocated by RRHs to i^{th} UE in v^{th} service. Also, $\mathbf{h}_{R_s, u(v,i)} \in \mathbb{C}^{R_s}$ is the vector of channel gain of wireless link from RRHs in the s^{th} slice to the i^{th} UE in v^{th} service. In addition, $\mathbf{w}_{R_s, u(v,i)} \in \mathbb{C}^{R_s}$ depicts the the transmit beamforming vector from RRHs in the s^{th} slice to the i^{th} UE in v^{th} service. Moreover, BN_0 denotes the power of gaussian additive noise and $I_{u(v,i)}$ is the power of interfering signals. To obtain SNR as formulated in equation (2),

let $\mathbf{y}_{U_v} \in \mathbb{C}^{U_v}$ be the received signal's vector of all users in v^{th} service which is given by equation (3)

$$\mathbf{y}_{U_v} = \sum_{k=1}^{N_k} \sum_{s=1}^{N_s} \mathbf{H}_{R_s, \mathcal{U}_v}^H (\mathbf{W}_{R_s, \mathcal{U}_v} \mathbf{P}_{U_v}^{\frac{1}{2}} \mathbf{x}_{R_s} + \mathbf{q}_{R_s}) \zeta_{U_v, k, s} + \mathbf{z}_{U_v} \quad (3)$$

where $\mathbf{x}_{R_s} = [x_{r(s,1)}, \dots, x_{r(s, R_s)}]^T \in \mathbb{C}^{R_s}$ depicts the transmitted symbol vector for the s -th set of Network slice, \mathbf{z}_{U_v} is the additive Gaussian noise $\mathbf{z}_{U_v} \sim \mathcal{N}(0, N_0 \mathbf{I}_{U_v})$ and N_0 is the noise power. In addition, $\mathbf{q}_{R_s} \in \mathbb{C}^{R_s}$ indicates the quantization noise which is made from signal compression in BBU.

Furthermore, $\zeta_{U_v, k, s} \triangleq \{\zeta_{u(v,1), k, s}, \zeta_{u(v,2), k, s}, \dots, \zeta_{u(v, U_v), k, s}\}$, $\zeta_{u(v,i), k, s} \in \{0, 1\}$ is a binary parameter that map Physical Resource Blocks (PRB) to UE. Also as defined before, $\mathbf{H}_{R_s, \mathcal{U}_v} = [\mathbf{h}_{R_s, u(v,1)}, \dots, \mathbf{h}_{R_s, u(v, U_v)}]^T \in \mathbb{C}^{R_s \times U_v}$ shows the channel matrix between RRH set \mathcal{R}_s to UE set \mathcal{U}_v . The channel vector from the RRH of s^{th} slice to the i^{th} UE in the v^{th} service $\mathbf{h}_{R_s, u(v,i)} \in \mathbb{C}^{R_s}$ is modeled as below

$$\mathbf{h}_{R_s, u(v,i)} = \beta_{R_s, u(v,i)}^{\frac{1}{2}} \mathbf{g}_{R_s, u(v,i)}, \quad (4)$$

where $\mathbf{g}_{R_s, u(v,i)} \sim \mathcal{N}(0, N_0 \mathbf{I}_{U_v})$ indicates the fast fading and flat fading channel vector and $\beta_{R_s, u(v,i)} = \text{diag}(b_{r(s,1), u(v,i)}, \dots, b_{r(s, R_s), u(v,i)})$ represents the large scale fading matrix. Here, it is assumed we have perfect channel state information(CSI).

Moreover, $\mathbf{W}_{R_s, \mathcal{U}_v} = [\mathbf{w}_{R_s, u(v,1)}, \dots, \mathbf{w}_{R_s, u(v, U_v)}] \in \mathbb{C}^{R_s \times U_v}$ is the zero forcing beamforming vector to minimize the interference which is indicated as follow

$$\mathbf{W}_{R_s, \mathcal{U}_v} = \mathbf{H}_{R_s, \mathcal{U}_v} (\mathbf{H}_{R_s, \mathcal{U}_v}^H \mathbf{H}_{R_s, \mathcal{U}_v})^{-1} \quad (5)$$

Hence, the interference power of i^{th} UE in v^{th} service can be represented as follow

$$\begin{aligned} I_{u(v,i)} = & \underbrace{\sum_{s=1}^S \sum_{n=1}^{N_s} \sum_{\substack{l=1 \\ l \neq i}}^{U_v} \gamma_1 p_{u(v,l)} a_{vs} \zeta_{u(v,i), n, s} \zeta_{u(v,l), n, s}}_{\text{(intra-service interference)}} \\ & + \underbrace{\sum_{\substack{y=1 \\ y \neq v}}^V \sum_{s=1}^S \sum_{n=1}^{N_s} \sum_{l=1}^{U_y} \gamma_2 p_{u(y,l)} a_{ys} \zeta_{u(v,i), n, s} \zeta_{u(y,l), n, s}}_{\text{(inter-service interference)}} \\ & + \underbrace{\sum_{s=1}^S \sum_{j=1}^{R_s} \sigma_{q_{r(s,j)}}^2 |\mathbf{h}_{r(s,j), u(v,i)}|^2 a_{vs}}_{\text{(quantization noise interference)}} \end{aligned} \quad (6)$$

where, $\gamma_1 = |\mathbf{h}_{R_s, u(v,i)}^H \mathbf{w}_{R_s, u(v,i)}|^2$ and $\gamma_2 = |\mathbf{h}_{R_s, u(v,i)}^H \mathbf{w}_{R_s, u(y,l)}|^2$. It is assumed that interference signal for each UE, is applied from UEs which are using same PRB. If we replace $p_{u(v,l)}$ and $p_{u(y,l)}$ by P_{max} , an upper bound $\tilde{I}_{u(v,i)}$ is obtained for $I_{u(v,i)}$. Therefore, $\tilde{\mathcal{R}}_{u(v,i)} \forall v, \forall i$ is derived by using $\tilde{I}_{u(v,i)}$ instead of $I_{u(v,i)}$ in equation (1) and (2).

let $\bar{p}_{r(s,j)}$ denote the power of transmitted signal from j^{th} RRH in s^{th} slice. from equation (3) we have

$$\bar{p}_{r(s,j)} = \sum_{v=1}^V \mathbf{w}_{r(s,j), \mathcal{U}_v} \mathbf{P}_{\mathcal{U}_v}^{\frac{1}{2}} \mathbf{P}_{\mathcal{U}_v}^{H \frac{1}{2}} \mathbf{w}_{r(s,j), \mathcal{U}_v}^H a_{vs} + \sigma_{q_{r(s,j)}}^2. \quad (7)$$

As a result the user data capacity on the fronthaul link between BBU and the j^{th} RRH in s^{th} slice is formulated as below

$$C_{R(s,j)} = \log \left(1 + \sum_{v=1}^V \frac{w_{r(s,j), \mathcal{D}_s} \mathbf{P}_{\mathcal{U}_v}^{\frac{1}{2}} \mathbf{P}_{\mathcal{U}_v}^{H \frac{1}{2}} w_{r(s,j), \mathcal{U}_v}^H a_{vs}}{\sigma_{q_{r(s,j)}}^2} \right), \quad (8)$$

C. Mean Delay

Suppose that we have two processing layer in BBU of O-RAN system. The lower layer is consist of PHY and MAC and the upper layer is consist of RLC, PDCP and SDAP.

As it is mentioned before, we have M_1 VMs in the first layer and M_2 VMs in second layer. Each VM in both layers map to one or more slices. So in s^{th} slice, there are M_{s1} VMs in first layer and M_{s2} VMs in second layer. Each VM in first and second layer has computational capacity that is equal to μ_1 and μ_2 respectively.

Let the packet arrival of UEs have a Poisson Process with arrival rate $\lambda_{u(v,i)}$ for i^{th} UE in v^{th} service. Therefore, the mean arrival data rate of UEs in s^{th} slice in the first layer is $\alpha_{s1} = \sum_{v=1}^V \sum_{u=2}^{U_v} a_{vs} \lambda_{u(v,i)}$. Furthermore, the mean arrival rate of second layer is approximately equal to the mean arrival rate of first layer $\alpha_s = \alpha_{s1} \approx \alpha_{s2}$ since, by using Burke's Theorem, the arrival packets of second layer which is processed in first layer is still Poisson with rate α_s . It is assumed that there are dispatchers in each layer for each slice to divide the incoming traffic to VMs. Suppose the baseband processing of each VM is depicted as a M/M/1 processing queue. Each packet is routed by one of VMs of slices. So the mean delay of slice s which is related to incoming traffic rate routed to each VM in first layer can be written as follow

$$d_{s1} = \frac{1}{\mu_1 - \alpha_s / M_{s1}} \quad (9)$$

Also, the delay in s^{th} slice in second layer can be formulated as below

$$d_{s2} = \frac{1}{\mu_2 - \alpha_s / M_{s2}} \quad (10)$$

In addition, the arrival data rate to the queue of wireless transmission is equal to the arrival data rate of dispatcher. 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 transmission layer is

$$d_{str} = \frac{1}{R_{tot_s} - \alpha_s} \quad (11)$$

Where, $R_{tot_s} = \sum_{v=1}^V \sum_{u=2}^{U_v} a_{vs} R_{u(v,i)}$. We define a new parameter which indicates mean delay of each slice

$$D_s = d_{s1} + d_{s2} + d_{str} \forall s \quad (12)$$

D. Physical Resource

Assume each VM is mapped to one virtual network function (VNF) for simplicity. Each VNF requires physical resources which contain RAM, Memory and CPU. Let, the required resources for VNF f in slice s is represented by a three dimensional vector as follow

$$\bar{\Omega}_{(f,s)} = \{\Omega_{R_{f,s}}, \Omega_{M_{f,s}}, \Omega_{C_{f,s}}\} \quad (13)$$

Where, $\bar{\Omega}_{(f,s)} \in \mathbb{C}^3$ and $\Omega_{R_{f,s}}, \Omega_{M_{f,s}}, \Omega_{C_{f,s}}$ indicate the amount of required RAM, Memory and CPU. Moreover total amount of required RAM, Memory and CPU of all VNFs in a slice is a three dimension vector which is defined as

$$\bar{\Omega}_s^{tot} = \sum_{f=1}^{F_s} \bar{\Omega}_{(f,s)} \quad (14)$$

Where, $F_s = M_{s1} + M_{s2}$ is the total number of VNFs in slice s .

Also, in the Core Network (CN), there are D_c data centers (DC), which served VNFs. Each DC contains several servers that supply VNF's needs. The amount of RAM, Memory and CPU is denoted respectively by τ_{R_j}, τ_{M_j} and τ_{C_j} for j^{th} DC.

$$\tau_j = \{\tau_{R_j}, \tau_{M_j}, \tau_{C_j}\}$$

Also we define a weighted parameter of τ_j as follow

$$\begin{aligned} \hat{\Omega}_s^{tot} &= w_R \bar{\Omega}_{R_s}^{tot} + w_M \bar{\Omega}_{M_s}^{tot} + w_C \bar{\Omega}_{C_s}^{tot} \\ \hat{\tau}_j &= w_R \tau_{R_j} + w_M \tau_{M_j} + w_C \tau_{C_j} \end{aligned} \quad (15)$$

Where, $\mathbf{w} = \{w_R, w_M, w_C\}$ are the weight of RAM, Memory and CPU.

E. Problem Statement

One of the most important parameters to estimate the optimality of the system is energy efficiency which is represented as sum-rate to sum-power as follow

$$\eta(\mathbf{P}, \mathbf{A}) := \frac{\sum_{v=1}^V \sum_{k=1}^{U_v} \mathcal{R}_{u(v,k)}}{\sum_{s=1}^S \sum_{i=1}^{R_s} \bar{p}_{r(s,i)}} = \frac{\mathfrak{R}_{tot}(\mathbf{P}, \mathbf{A})}{P_{r_{tot}}(\mathbf{P}, \mathbf{A})}, \quad (16)$$

Assume the power consumption of baseband processing at each data center that is mapped to VNFs of a slice is depicted as ϕ . So the total power can be represented as

$$\phi_{tot} = \sum_{s=1}^S \sum_{d=1}^{D_c} y_{s,d} \phi$$

Where, $y_{s,d}$ is a binary variable which indicates whether d^{th} data-center is mapped to VNFs of s^{th} slice or not.

In this paper, the main goal is to simultaneously maximize

sum-rate and minimize sum-power with the presence of constraints which is written as follow,

$$\max_{\mathbf{P}, \mathbf{A}, \mathbf{Y}} \eta(\mathbf{P}, \mathbf{A}) + \frac{1}{\phi_{tot}(\mathbf{Y})} \quad (17a)$$

$$\text{subject to } \bar{p}_{r(s,i)} \leq P_{max} \quad \forall s, \forall i, \quad (17b)$$

$$p_{u(v,k)} \geq 0 \quad \forall v, \forall k, \quad (17c)$$

$$\mathcal{R}_{u(v,k)} \geq \mathcal{R}_{u(v,k)}^{min} \quad \forall v, \forall k, \quad (17d)$$

$$C_{r(s,i)} \leq C_{r(s,i)}^{max} \quad \forall s, \forall i, \quad (17e)$$

$$D_s \leq D_s^{max} \quad \forall s, \quad (17f)$$

$$\sum_{s=1}^S a_{vs} \geq 1 \quad \forall s, \quad (17g)$$

$$\sum_{d=1}^{D_c} \sum_{v=1}^V y_{s,d} a_{vs} \geq 1 \times \sum_{v=1}^V a_{vs} \quad \forall s, \quad (17h)$$

$$\bar{\Omega}_{\mathfrak{z}(s)}^{tot} = \sum_{f=1}^{F_s} \bar{\Omega}_{\mathfrak{z}(f,s)} \leq \sum_{d=1}^{D_c} y_{s,d} \tau_{\mathfrak{z}d} \quad \forall \mathfrak{z}, \forall s \quad (17i)$$

Where, $\mathbf{P} = [p_{u(v,k)}] \forall v, \forall k$, $\mathbf{A} = [a_{vs}] \forall v, \forall s$ and $\mathbf{Y} = [y_{s,d}] \forall s, \forall d$. Equation (17b), (17c) indicates respectively that the power of each RRH do not exceed the maximum power and power of each UE is a positive integer value. Also (17d) shows that the rate of each UE is more than a threshold. (17e) and (17f) depicts respectively that the capacity of fronthaul link is limited and the delay of receiving signal should be less than a threshold. Furthermore, (17g) ensure that each service is mapped to one or more slice. Also, (17h) guarantee that each slice (VNFs in two layers of slice) has been placed to one or more physical resources (DC). Moreover, in (17i) $\mathfrak{z} \in \{M, R, C\}$, which supports that we have enough physical resource for VNFs of each slice.

The main optimizaition problem which is formulated as (17) can be decomposed into two independent optimization problem A and B. The problem A is introduced as bellow

$$\max_{\mathbf{P}, \mathbf{A}} \eta(\mathbf{P}, \mathbf{A}) \quad (18a)$$

$$\text{subject to } \bar{p}_{r(s,i)} \leq P_{max} \quad \forall s, \forall i, \quad (18b)$$

$$p_{u(v,k)} \geq 0 \quad \forall v, \forall k, \quad (18c)$$

$$\mathcal{R}_{u(v,k)} \geq \mathcal{R}_{u(v,k)}^{min} \quad \forall v, \forall k, \quad (18d)$$

$$C_{r(s,i)} \leq C_{r(s,i)}^{max} \quad \forall s, \forall i, \quad (18e)$$

$$D_s \leq D_s^{max} \quad \forall s, \quad (18f)$$

$$\sum_{s=1}^S a_{vs} \geq 1 \quad \forall s \quad (18g)$$

and the problem B is

$$\min_{\mathbf{Y}} \phi_{tot}(\mathbf{Y}) \quad (19a)$$

$$\text{subject to } \sum_{d=1}^{D_c} \sum_{v=1}^V y_{s,d} a_{vs} \geq 1 \times \sum_{v=1}^V a_{vs} \quad \forall s, \quad (19b)$$

$$\bar{\Omega}_{\mathfrak{z}(s)}^{tot} = \sum_{f=1}^{F_s} \bar{\Omega}_{\mathfrak{z}(f,s)} \leq \sum_{d=1}^{D_c} y_{s,d} \tau_{\mathfrak{z}d} \quad \forall \mathfrak{z}, \forall s \quad (19c)$$

III. Proposed Method For Problem (18)

In this subsection, the proposed method is applied to solve the optimization problem. We want to solve (18). Since the problem is non-convex and NP-Hard iterative algorithm is applied. To solve the problem and obtain optimum \mathbf{A} and \mathbf{P} we divide problem (18) to two different part that can be solved iteratively.

A. First Part of Sub-Problem A

Firstly, we need to obtain \mathbf{A} by fixing \mathbf{P} in the problem (18) and updating this parameter at the end of each iteration. Two different method is applied to acquire \mathbf{A} . The first method is using MOSEK and second method is a heuristic algorithm. The details of heuristic algorithm are represented in Algorithm (1).

Algorithm 1 Mapping Slice to Service

- 1: Sort services according to their priority, the number of UEs in it and their requirements in descending order.
 - 2: Sort slices according to the number of PRBs, RRHs and VNFs in two layers and the Capacity of their resources in descending order.
 - 3: for $i \leftarrow 1$ to S do
 - 4: for $j \leftarrow 1$ to V do
 - 5: Set $a_{ij} = 1$
 - 6: Obtain Parameters of System
 - 7: if conditions (17b), (17c), (17d) and (17e) is not applied then
 - 8: Set $a_{ij} = 0$;
 - 9: else
 - 10: break from inner loop;
 - 11: end if
 - 12: end for
 - 13: end for
-

B. Second Part of Sub-Problem A

In this part, by assuming that \mathbf{A} is fixed, the optimal power of UEs in each service is achieved.

Theorem 1. η^* which is the optimum energy efficiency can be achieved if

$$\begin{aligned} \max_{\mathbf{P}} (\mathfrak{R}_{tot}(\mathbf{P}) - \eta^* P_{r_{tot}}(\mathbf{P})) = \\ \mathfrak{R}_{tot}(\mathbf{P}^*) - \eta^* P_{r_{tot}}(\mathbf{P}^*) = 0. \end{aligned} \quad (20)$$

Proof. See Appendix A \square

The second subproblem can be solved using Lagrangian function and iterative algorithm. Since, Interference is a function of power of UEs, for simplicity, we assume an upper bound $\bar{I}_{u(v,i)}$ for interference (the worst-case). In order to make equation (18) as a standard form of convex optimization problem, it is required to change the variable of equations (18e) and (18f). Lagrangian function is written as follow

$$\begin{aligned} \mathcal{L}(\mathbf{P}; \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\xi}, \boldsymbol{\kappa}) = & \sum_{v=1}^V \sum_{k=1}^{U_v} \bar{\mathcal{R}}_{u(v,k)} - \eta \sum_{v=1}^V \sum_{i=1}^{R_s} \bar{p}_{r(s,i)} \\ & + \sum_{s=1}^S \sum_{k=1}^{U_v} \lambda_{u(v,k)} (\bar{\mathcal{R}}_{d(s,k)} - \mathcal{R}_{u(v,k)}^{max}) \\ & - \sum_{s=1}^S \sum_{i=1}^{R_s} \mu_{r(s,i)} (\bar{p}_{r(s,i)} - P_{max}) \\ & - \sum_{s=1}^S \sum_{i=1}^{R_s} \xi_{r(s,i)} (\bar{p}_{r(s,i)} - \sigma_{q_{r(s,i)}}^2 2^{C_{r(s,i)}^{max}}). \\ & + \sum_{v=1}^V \sum_{k=1}^{U_v} \kappa_{u(v,k)} \sum_{s=1}^S (R_{u(v,k)} - \mathfrak{D}_s) a_{vs}. \end{aligned} \quad (21)$$

Where, $\mathfrak{D}_s = \frac{1}{D_s^{max} - d_{s1} - d_{s2}} + \alpha_s$. Also, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, $\boldsymbol{\xi}$ and $\boldsymbol{\kappa}$ are the matrix of Lagrangian multipliers which have non-zero positive elements. Optimal power is obtained from equation (21) as follow

$$p_{u(v,i)}^* = \frac{\eta_{u(v,i)} \mathfrak{w}_{u(v,i)} - \mathfrak{r}_{u(v,i)} \mathfrak{z}_{u(v,i)}}{\mathfrak{r}_{u(v,i)} \mathfrak{w}_{u(v,i)}} \quad (22)$$

where, $\eta_{u(v,i)} = (-\mu_{u(v,i)} + \lambda_{u(v,i)} + \kappa_{u(v,k)} + 1) \frac{B}{Ln_2}$ and $\mathfrak{w}_{u(v,i)} = \sum_{s=1}^{N_s} |\mathbf{h}_{R_s, u(v,i)}^H \mathbf{w}_{R_s, u(v,i)}|^2 a_{vs}$. Also $\mathfrak{z}_{u(v,i)} = BN_0 + \bar{I}_{u(v,i)}$ and $\mathfrak{r}_{u(v,i)} = \sum_{s=1}^S \sum_{i=1}^{R_s} (\xi_{r(s,i)} + \eta) \|w_{r(s,i), u(v,i)}\|^2$. By using sub-gradient method, the optimal power \mathbf{P} is obtained.

C. Solving two part of Sub-problem A iteratively

In (III-A) and (III-B) the details of solving each part of subproblem is depicted. Here, the algorithm of solving sub-problem A is shown in Algorithm (2)

D. Sub-Problem B

In this subsection, we want to solve equation (19) which is the placement of virtual resources to physical resource in order to minimize the number of using DCs. To achieve optimum \mathbf{Y} heuristic algorithm and Mosek is applied. The details of heuristic algorithm is written in Algorithm (3). In this algorithm, if a slice does not admit to a specific DC, it is remain for next placement. In next placement the residual slices, map to more than one DC according to their requirements.

Algorithm 2 Joint Network Slicing and Power Allocation

```

1: Set the maximum number of iterations  $I_{max}$ , convergence condition  $\epsilon_\eta$  and the initial value  $\eta^{(1)} = 0$ 
2: Set  $\mathbf{P} = \mathbf{P}_{max}$ 
3: for counter  $\leftarrow 1$  to  $I_{max}$  do
4:   Achieve  $\mathbf{A}$  by applying Algorithm (1)
5:   Obtain  $\mathbf{P}$  by using sub-gradient method which is mentioned in (III-B).
6:   if  $\Re_{tot}(\mathbf{P}^{(i)}, \mathbf{A}^{(i)}) - \eta^{(i)} P_{r_{tot}}(\mathbf{P}^{(i)}, \mathbf{A}^{(i)}) < \epsilon_\eta$  then
7:     Set  $\mathbf{P}^* = \mathbf{P}^{(i)}$ ,  $\mathbf{A}^* = \mathbf{A}^{(i)}$  and  $\eta^* = \eta^{(i)}$ ;
8:     break;
9:   else
10:     $i = i + 1$ , Setting  $\mathbf{P} = \mathbf{P}^{(i)}$  ;
11:  end if
12: end for

```

Algorithm 3 Placement of Physical resources into Virtual resources

```

1: Sort Slices according to  $\hat{\Omega}_s^{tot}, \forall s$  in descending order.
2: Sort DCs according to  $\hat{\tau}_j, \forall j$  in descending order.
3:  $\mathbf{Y} = \mathbf{0}$ 
4: for  $d \leftarrow 1$  to  $D_c$  do
5:   for  $s \leftarrow 1$  to  $S$  do
6:     if  $\sum_{d=1}^{D_c} y_{s,d} == 0$  and  $\bar{\Omega}_{s(s)}^{tot} \leq \tau_{3j} \forall j, \forall s$  then
7:       Set  $y_{s,d} = 1$ ;
8:        $\tau_j \leftarrow \tau_j - \bar{\Omega}_s^{tot}$ 
9:     end if
10:  end for
11: end for
12:  $\{ind_{residual} = s | (\sum_{d=1}^{D_c} y_{s,d} == 0)\}$ 
13: Sort residual amount of DCs same as before in descending order.
14: Sort residual slices same as before in descending order.
15: for  $r \leftarrow 1$  to  $S_{residual}$  do
16:   for  $n \leftarrow 1$  to  $D_c$  do
17:     Set  $y_{s,d} = 1$ ;
18:      $\bar{\Omega}_s^{tot} \leftarrow \bar{\Omega}_s^{tot} - \tau_j$ 
19:     if  $\bar{\Omega}_s^{tot} == 0$  then
20:       Set  $y_{s,d} = 1$ ;
21:        $\tau_j \leftarrow \tau_j - \bar{\Omega}_s^{tot}$ 
22:       break inner loop
23:     end if
24:   end for
25: end for

```

IV. Simulation

In this section, Simulation and numerical results for the main problem are depicted.

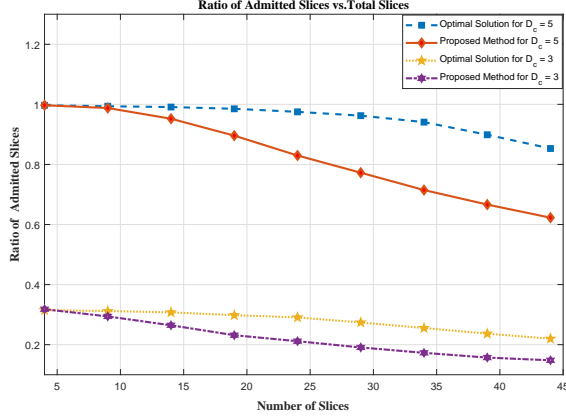


Fig. 2: algorithm

References

- [1] I. Scales, "The open ran (oran) alliance formed to lever open 5g for 'other' technologies? and much more...", 2018.
- [2] J. Wu, Z. Zhang, Y. Hong, and Y. Wen, "Cloud radio access network (c-ran): a primer," *IEEE Network*, vol. 29, no. 1, pp. 35–41, 2015.
- [3] J. Tang, W. P. Tay, T. Q. Quek, and B. Liang, "System cost minimization in cloud ran with limited fronthaul capacity," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3371–3384, 2017.
- [4] O. Simeone, J. Kang, J. Kang, and S. Shamai, "Cloud radio access networks: Uplink channel estimation and downlink precoding," *arXiv preprint arXiv:1608.07358*, 2016.
- [5] P. Luong, C. Despins, F. Gagnon, and L.-N. Tran, "A novel energy-efficient resource allocation approach in limited fronthaul virtualized c-rans," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, pp. 1–6, IEEE, 2018.
- [6] P. Luong, F. Gagnon, C. Despins, and L.-N. Tran, "Joint virtual computing and radio resource allocation in limited fronthaul green c-rans," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2602–2617, 2018.
- [7] "xran forum merges with c-ran alliance to form oran alliance," 2018.
- [8] X. Zhou, R. Li, T. Chen, and H. Zhang, "Network slicing as a service: enabling enterprises' own software-defined cellular networks," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 146–153, 2016.