A Model for Slices’ Resource Deployment in 5G Scenario
Jing WANG, Sui-xiang GAO*, Sai-nan WANG and Qian ZHU
University of Chinese Academy of Sciences, Beijing, China
*Corresponding author

Keywords: Slice, Resource Deployment, C-RAN in 5G.

Abstract. Compared with 4G, the fifth-generation mobile network dedicated to provide users with personal and flexible services. The new network architecture effectively isolate the core layer and control layer, and generate component Cloud-RAN. This paper founded in Cloud-RAN and go into factors like distance, intercellular interference, and intra slice interference. Furthermore, the model firstly take into account effect of users' heterogeneous demands on network utility optimization model, will optimize radio-resource allocation scheme among different slices. Finally, the results of the deployment in the slice-free environment compared with those in the slicing environment, indicates that the slicing idea can simplify the computational complexity in the deployment of 5G network resources, and the network utility performs better in the case of rapid increase of user requirements number and complexity.

Background

Due to the continuous development of mobile communication technology, we are able to broaden the connectivity edge of Internet-of-things (IoT). With increasing number of connected devices, a huge variety of data services are brewed with their dedicated communication quality demands. These demands including throughput, latency, packet loss ratio etc. In a common network environment, such demands conflicts due to limited radio resource in the whole network. As LTE is developed so far, previous solutions are invented like smarter scheduler, multiple radio bearers, link adaption and so on.

In 5G era, low latency, high throughput and good reliability triggers more user-customized use-cases. Thus, the complexity of communication Quality of Service (QoS) increases much more with various network Key Performance Indicators (KPIs\(^1\)). On the one hand, this contributes to marketing like network as a service (NaaS), which shapes network according to industries and use-cases. On the other hand, network subscriber puts highly customized subscription package and demands the operators to provide tailored services.

Network Slicing, as one of the major reform in 5G core, according to the introduction in IMT-2020 5G Trials in China\(^1\), contributes to provide network resource in an optimized manner. Based on shared infrastructure, logic network section created for use case specific demands in 5G evolved packet core (EPC). In Radio Access Network (RAN), according to Ericsson Network Slicing Report\(^2\), the process furtherly assisted by radio resource partitioning (RRP) to enhance the performance of QoS levels. From the bandwidth perspective, the service reliability is increased due to parts of the bandwidth is booked in advance for the service in a network slice. The trade-off happens when balancing multiple such pre-booked network resource with, if applicable, new subscriber’s QoS demand. Beyond conventional Radio Resource Management (RRM) approach, we employ optimization theory and heuristic algorithm to find the best deal among several counters.

Related Works

Introduction on Report\(^1\) focuses on 5G scenarios based on different user QoS. Each scenario has different levels of requirements for each module of the network. According to the statistics of mobile services, the amount of data requests of users has peaks and dips. It is foreseeable that 5G networks need to face the needs with different types of requirement, Distribution pattern over time.
In order to utilize the resources throughout the grid effectively, the network must have the ability to allocate resources adaptively.

Based on this idea, several organizations began to study a more flexible resource scheduling architecture. A slice is a logical network that segmented from the virtual resources of the entire network and can serve a certain type of users independently. Therefore, the sliced programmable network traverses Core Network (CN) and the access network (AN).

With the improvement of SDN and NFV (Network Function Virtualization) technology, the network gradually can adaptively allocate resources. In the context of NFV technology, a slice is regarded as a number of virtual users of a virtual base station, and each virtual user’s demand is attached to the number of similar business users in 4G. In order to meet users' service quality needs, this article used impact factors and allocation decisions to calculate SIR (Signal-to-Interference Rate) and all slices’ SIR is regard utility for the entire network, then guarantees the user's QoS by max networks utility.

Combined with the technical background and practical application, 5G resource allocation problem has several basic aspects: slices’ geographical distribution & servicing quality, network resource action set & utility function. We will sort out the above contents after this paper.

In the future 5G scenario, the mobile communication network is a complex system of heterogeneous devices because of complex users’ needs. The statistics of the existing communication network services show that the traffic generated by the users is extremely uneven in geographical and temporal distribution, which is likely to cause redundancy of the infrastructure. In order to solve these problems, many research institutions advocate 5G to adopt SDN and NFV technologies. SDN exit in CN and can flexibly support multiple types of users and coordinate resources through defended software. It separates the control plane from the forwarding plane and forms a logical network in the control plane. The logical network can deliver multiple distributed controller instances down and provide an interface to the infrastructure; some researches build a hierarchical distribution model based on this network architecture or priority. In the process of the above process, SDN realizes the control function of deploying network management resources on-demand through NFV under the premise of independently operating the wireless protocol. To express the radio resources logically, people define the concept of slicing: it exists to implement a particular service. In addition, the virtualization technology realizes the logical isolation between the slices. Therefore, each slice contains resources for the complete process across different network layers, there are also studies on resource modeling through inter-slice relationships. However, the most fundamental thing is deployment for slice have to fulfill the user's QoS.

In the network in 4G, the base stations distributed in various places provide the functions of access and data transmission of the user base. The maximum number of users that a base station can carry depends on the minimum ratio of resources to user demand. However, in 5G radio resource allocation, one slice can be cooperatively supported by several base stations. On the one hand, the varity of the user's acquisition of resources is expanded; On the other hand, the slice can customize the set of resources grasped by the slice based on the type of user request around the subordinate base station. For example, a slice corresponding to a smart home service combines a plurality of base stations distributed in various places and calls the base station basic contact function. This not only improves the stability of the user's contact network, reduces the average coverage of the base station, reduces the power of the base station, but also splits the data storage function and transmission function of the base station in order to achieve other high data transmission and low latency user requirements.

As can be seen from the above example, the biggest role of the slice is to achieve network utility under the premise of multiple types of user requirements. The model, which is built on a single slice, as mentioned in other papers on allocating resources, is meaningless out of the practical usecase. This paper first establishes the idea of using slices as virtualized users proposing for the first time, and the resource deployment scheme of multiple types of slices discussed as insights.
**Model Description**

**System Describe**

Suppose a gNodeB (gNB) set \( C = 1, 2, 3, \ldots, n \) to form up the local RAN topology, we set User Equipment (Ue) in slice \( S \) be \( U = U_S \) attached to a certain gNB \( C_i \). Every gNB has a set of parameters including total transmission bandwidth \( V_i \), Base station power \( P_i \), geographical position \( G = \{(g_i)\} \).

Then \( C^T 1_n - 1_n^T C \) described the relationship between two stations:

\[
(C^T 1_n - 1_n^T C)_{ij} = \begin{cases} \frac{1}{|a_i-a_j|} & i \neq j \\ \frac{1}{|b_i-b_j|} & i = j \\ 0 & \end{cases}
\]

(1)

The larger variable indicates that there are more susceptible between two base stations when they at the same frequency. For a certain base station \( i \), we make the following appointment:

\[
\|\left(\frac{1}{|a_i-a_j|}, \frac{1}{|b_i-b_j|}\right)\| > 100 \quad c_{i,j} = 0 \\
\|\left(\frac{1}{|a_i-a_j|}, \frac{1}{|b_i-b_j|}\right)\| \leq 100 \quad c_{i,j} = 1
\]

(2)

In addition, \( c_{i,i} = 0 \).

We regard sliced user \( U_S \) as gNodeB user. In 5G system, slices’ resource provided by multiple base stations. Therefore, sliced user can deploy on multiple yNodeB simultaneously. In the time interval, yNodeB \( i \) provide \( \xi_{s,i} \in [0, 1] \) proportion of broadband for slices \( S \). The characteristics of a slice \( s \) is defined as follow: \([|U_s|, G_s, R_s, SIR_s]\), representing the number of users of slice \( s \), geometric center of slice’s user location: \( G_s = \{(a_s, b_s)\} \), rate of use’s lowest requirements, SIR power ratio. The slices’ user requirements then translated into the bandwidth resources of yNodeB:

Firstly, according user’s rate demand and location distribution calculate \( SIR_s \) of \( s \). Suppose distance between yNodeB \( i \) and slice \( L = \{(a_s - a_i, b_s - b_i)\} \). Expressed by a matrix:

\[
L = G^s_t 1_n - 1_m^T G^s_t
\]

(3)

Set the broadband of yNodeB \( i \) is \( V_i \), Total bandwidth obtained by slice \( s \) will be:

\[
V_s = \sum_s \xi_{s,i} V_i
\]

(4)

By Shannon formula, user on slice \( s \) minimum \( SIR_s \) will be:

\[
\frac{\sum \xi_{s,i} V_i}{2^{\xi_{s,i} V_i SIR_s} - 1}
\]

(5)

Equation shows that the higher the number of slicing users and required rate of slicing is, more \( \xi_{s,i} \) of yNodeB \( i \) needed, for better \( SIR_s \). This is consistent with the actual status.

Under a condition of the actual geographical distribution of current users, interference of slice \( s \) users includes the following situations:

1. Interference between similar users;
2. Interference of base station.

Set noise function \( N(C, \xi_{s,i}) \), and \( \sum_j \xi_{s,j,t} * c_{i,j} \) is disturbance of slice \( s \) users’ when they use resource of yNodeB. Moreover, suppose thermal noise in the channel is “1”. Then, in the actual environment, the Signal-to-Interference radio received from other users expressed as follows:

\[
SIR_{s,t}(l_{s,i,t} \sum \xi_{s,j,t} * c_{i,j}) = \frac{P_i}{\sum \xi_{s,j,t} * c_{i,j+1}}
\]

(6)

**RAN Slice Model**

To investigate network quality from satisfaction degree of user requirements and resource deployment environment, the following objective function is established:
\[ f_{s,i}(\xi_{s,i}) = SIR_s(\xi_{s,i}^{3/2}, \Sigma_j \xi_{s,j}^{2}c_{i,j}) + SIR_s(\xi_{s,i}) \] (7)

Function contains three facts: When the deployment strategy meets the basic needs of users, the farther away the base station is from users, the less bandwidth. This means that the function (7) gives priority to assigning users to the base station at close range. (7) Increases first and then decreases with the number of users. The lower the rate required by the user, the better the function.

Considering the resource deployment of all slices, sum \( f_{s,i} \) over \( s \) to obtain the target function of slicing resource deployment model:

\[
\max \sum_{s,i} f_{s,i} \quad \text{(8)}
\]

\[
\sum_i \xi_{s,i} = 1 \text{ for } \forall i \in c \quad \text{(9)}
\]

\[
\xi_{s,i} \in [0,1] \text{ for } \forall i \in c \quad \text{(10)}
\]

**Model Analysis & Algorithm:**

In RAN Slice Model, only \( l_{s,i} \) is unknown variables. Derivative of \( l_{s,i} \) is

\[
\frac{\partial F}{\partial \xi_{s,i}} = \sum_k \frac{\partial}{\partial \xi_{s,i}} \left( -\frac{2P_{l_{s,k}}^{3/2}c_{i,k}\xi_{s,k}}{(l_{s,k}^{3/2} \Sigma_j \xi_{s,j}^{2} + 1)^2} c_{i,j} \right) V_i |U_s| R_s \ln 2 \quad \text{(11)}
\]

After observation, it is found that the maximum value of the above function is greater than 0 and the minimum value is less than 0, so there is a chance to reach the best advantage. Furthermore,

\[
\frac{\partial^2 F}{\partial \xi_{s,i} \partial \xi_{s,k}} = \sum_{c=0}^k \sum_{c'=k}^\infty \left( A \right) \quad \text{(12)}
\]

\[
A = \frac{8P_{l_{s,k,t}}^{3/2}c_{i,k,t}\xi_{s,k,t}c_{i,k0}c_{i,k0}^{3/2}c_{i,j}^{3/2}c_{i,j}^{3/2}c_{i,j}^{+1}}{(l_{s,k}^{3/2} \Sigma_j \xi_{s,j}^{2} c_{i,j}^{+1})} \quad \text{(13)}
\]

Therefore, the objective function is convex and the optimal value can reach the global optimal solution. Now use the idea of the steepest descent to design approximation algorithm:

**Step1:** Initial value: \( \xi_i = \{\xi_{1,i}, \xi_{2,i}, \xi_{3,i}\} = (\frac{1}{3}, ..., \frac{1}{3}), \quad \varepsilon = 0.1\)

**Step2:** Calculate \( \frac{\partial F}{\partial \xi_i} = (\frac{\partial F}{\partial \xi_{s,i}}) \); 

**Step3:** for all \( s \):

**Step3.1:** Find max \( \frac{\partial F}{\partial \xi_i} \) in \( \frac{\partial F}{\partial \xi_i}, \quad c = \xi_{s,i}, \quad \text{ caculate } \xi_{s,i} = \min\{1, \xi_{s,i} + \alpha \cdot \frac{\partial F}{\partial \xi_i} \}, \quad \alpha = 0.1. \text{ If } \xi_{s,i} = 1, \text{ keep } c = 1 - \xi_{s,i};

**Step3.2:** Find min \( \frac{\partial F}{\partial \xi_i} \) in \( \frac{\partial F}{\partial \xi_i}, \quad c = c - \xi_{s,i}, \quad \text{ if } c < 0, \quad \text{ caculate } \xi_{s,i} = -c, \quad c = 0. \text{ Or } \xi_{s,i} = 0, \text{ keep } c = c;

**Step3.3:** Repete 3.2 until \( \frac{\partial F}{\partial \xi_i} \leq 10\varepsilon \), let \( \alpha = 0.1\alpha \).

**Step4:** Repete Step3 until \( \frac{\partial F}{\partial \xi_i} \leq \varepsilon \).

**Step5:** Repete Step 2-4 until \( \xi_i \)'s changement less than \( \varepsilon \).

**Simulation and Summary**

We construct a simple example, suppose there are three types of slices and each with 5, 10, and 20 users. Each requires a bandwidth of 5, 0.5, and 0.01 MHz, and the downlink rate is 100, 10, 10 Gbps. Four gNB with bandwidth of 6, 5, 5, 15MHz and same power allocation of, 60w. The simulation results shows as Figure 1:

The resource allocation model in this paper takes full account of the elements, like, distance between users and cell, the interference between users, the interference generated by base station when they are serving the same type of users, the number of users and their diversified demands. In the previous 4G resource allocation model, only a few factors considered, while in the 5G scenario,
most researchers only study one type of slice. Model in this paper integrates all the factors mentioned above to obtain the practical scenario application.

Moreover, when a slice is set up for each user, it is equivalent to restoring the flexible allocation of resources in 4G scenario, then in the algorithm, each iteration has calculated the decreasing derivative for every user, and the computational complexity will change from $O(|S|)$ to the $O(|U_s|)$. When the number of requests is increased in vain, the complexity will also increase greatly. Therefore, this model only serves 5G scenarios and fully illustrates the value of slicing in large-scale requests.

![Figure 1. Simulation results.](image)

**Literature References**


