Hierarchical resource allocation strategy for RAN slices in smart grid

Mingyue Liu¹²*, Yang Wang¹, Sachula Meng¹ and Xiongwen Zhao³

¹Electric Communication, China Electric Power Research Institute, Beijing, 100192, China
²Communication Engineering, North China Electric Power University, Beijing, 102206, China
³Corresponding author’s e-mail: 1182201423@ncepu.edu.cn

Abstract: With the rapid development of the new generation of smart grid, the variety of smart grid services are becoming more and more diversified, which lead to various service demands. According to different service requirements, it is necessary to allocate network resources flexibly and dynamically. In order to solve the problem of resource allocation among network slices in smart grid, we propose a two-layer dynamic resource scheduling strategy for network slicing. It can be divided into inter slice resource allocation and intra slice resource allocation. The former refers to the allocation of network resources to network slices. The latter refers to that the network slice allocates the allocated resources to the users in the slice. The simulation results show that the proposed two-layer network slicing dynamic resource scheduling strategy is superior to the current resource scheduling strategy in terms of quality of experience (QoE) of users and system throughput, which can better meet the needs of users.

1. Introduction

With the rise and development of the Internet of vehicles, the Internet of things (IoT), industrial intelligent control and various vertical industries, mobile communication will face thousands of times of data flow growth and hundreds of billions of equipment networking needs in the future. 5G, as the development direction of new generation mobile communication technology, will further meet the massive demand of future IoT applications on the basis of improving the service experience of mobile Internet users. 5G needs to implement customized services in multiple fields and scenarios and establish links between various industries. How to meet the diversified service needs of mobile Internet and IoT in a low-cost and efficient way is the focus of communication research, and network slicing technology just can solve the problem very well. Network slicing is one of the key technologies in 5G network. Network slicing abstracts all kinds of physical resources into virtual resources through network virtualization technology. Based on the specified network function and specific access network technology, network slicing can build an end-to-end logical network on demand and provide one or more network services. By customizing the network and constructing the network flexibly, the network slice can provide the optimal network resource allocation scheme.

Driven by the growth of energy and power demand, the world grid has entered a new stage of smart grid from the traditional network. Smart grid covers power generation, transmission, transformation, distribution, power consumption and other links. Due to the diversity of power business, a large number of data generated in the smart grid needs a strong network for communication. Therefore, as a typical
representative of the vertical industry, smart grid brings new challenges to the communication network. According to the different needs of power users, a network with high reliability, isolation, ultra-low delay and low cost is needed. 5G network slicing technology, a new generation of communication technology, has the possibility and feasibility of application in power services. Different network slices can be created to carry different power services in the smart grid, so as to ensure the quality of service (QoS) of the services carried by the network slices.

In order to improve the overall performance of the network system, first of all, we should reasonably allocate the limited resources in the network to achieve efficient and maximum utilization of resources. At present, most of the resource allocation strategies only consider a single level of resource allocation, which can maximize the rate and delay of users in the system, but can not meet the resource allocation in the slicing environment, such as in [4-8]. Power service is divided into elastic service and real-time service and a dynamic resource allocation strategy based on reinforcement learning algorithm is proposed in [4]. In [5], resources are iteratively assigned to the users in order to increase their level of satisfaction that is differently defined for each slice. In [6], a resource allocation method based on proportional fairness algorithm is proposed. At each iteration a slice and a user are selected according to a proportional fair (PF) approach: the selected slice picks the physical resource having the best channel response, and assigns it to the selected user. An analysis of the share-constrained proportional allocation mechanism is provided in [7]. In particular, each slice has a given amount of resources that depends on its sharing level, and customizes its own users’ allocation reacting to the allocation of other slices, so as to maximize its own utility. In order to maximize the resource utilization with a constraint on the outage probability, [8] proposes the use of an offline reinforcement learning followed by a low-complexity heuristic. At present, some papers have proposed two-tier resource allocation strategy, such as in [9-11]. In [9], a two-level priority access resource control mechanism based on heuristic algorithm is proposed, which can effectively improve the user’s experience quality and network resource utilization. In [10], a two-layer scheduler is proposed to allocate network resources in wireless access networks. The resource allocation between slices mainly considers the priority and isolation of slices, and the resource allocation within slices mainly depends on the types of slices. In [11], a dynamic resource scheduling strategy based on improved greedy algorithm is proposed. In the stage of slice scheduling and user scheduling, a two-layer model is designed. Combined with the improved greedy algorithm, the slice with the largest service weight value and the user combination with the highest priority can complete the corresponding resource allocation process. So as to make the resources of the whole system and the quality of user experience reach the best. Through the analysis of the current literature, it can be found that priority has become one of the important factors in the field of power grid.

At present, there are not many literatures about resource allocation in smart grid. Due to the large number of users in smart grid, services in different slices have their own quality of service requirements such as delay, speed, reliability, etc. Therefore, in this paper, we propose a two-layer network slicing dynamic resource scheduling strategy for smart grid scenarios. The main contributions of this paper are as follows:

(1) Unlike most of the current smart grid resource allocation algorithms, only the user's resource allocation is considered, in this paper, the resource management of the network is divided into two parts, which are used to solve the problem of resource allocation between slices and within slices.

(2) In this paper, resource management is divided into inter slice and intra slice. The proportional fair allocation algorithm is used in the resource allocation among slices, so that every slice can get the resource, which ensures the fairness of slices. The resource allocation algorithm in the slice ensures the maximization the QoE of users, and the throughput capacity of the system is also significantly improved.

The rest of the paper is as follows: After introducing the requirements of services in smart grid in Section 2, our system model are elaborated in Section 3. Then the RAN slicing resource allocation problem are presented in Section 4. The evaluation and simulation of the proposed method is given in Section 5. Finally, conclusions are drawn in Section 6.
2. Service classification in amart grid

In this paper, we mainly consider three kinds of service slices in 5G: enhance mobile broadband (eMBB) slice, ultra reliable low latency communications (uRLLC) slice and massive machine type communications (mMTC) slice. EMBB, uRLLC and mMTC are three application scenarios in 5G defined by ITU[12]. At present, the stock service of power grid can be divided into power grid control, information collection and mobile application. Corresponding to three application scenarios, as shown in Table 1.

| Scenario category | Typical service scenario | Service category |
|-------------------|-------------------------|-----------------|
| uRLLC             | Distribution automation  |
|                   | Distribution network protection |
|                   | Precise load control     | Control services |
| mMTC              | Distributed power        |
|                   | Electricity information collection |
|                   | Power quality monitoring |
|                   | Comprehensive monitoring of distribution substation |
|                   |                          | Information collection services |
| eMBB              | Mobile office            |
|                   | UAV inspection           |
|                   | Mobile video conference  |
|                   | Robot inspection of transmission and distribution |
|                   |                          | Mobile application services |

Control service in smart grid mainly includes distribution automation, precise load control, distribution network protection, etc. This kind of service involves the safe and stable operation of power grid, mainly used for fault isolation, self-healing control, precise timing, etc. Ultra low delay and high reliability are typical characteristics of the control services. In the future, there will be more distributed point-to-point connections in the service connection mode of power grid control. The main station system will gradually sink, and there will be more local and nearby control, and the demand for linkage with the master network control. Take the distribution automation service as an example, the delay is required to be 10ms, and the reliability is required to be 99.999%. In the uRLLC scenario, it can provide millisecond level end-to-end delay and nearly 100% service reliability guarantee, so it is suitable for control services in the power grid.

Information collection service mainly includes electricity information collection, electric vehicle charging pile, power quality monitoring, comprehensive monitoring of power distribution station, etc. This kind of service involves a large number of communication terminals and is widely distributed in all aspects of the power grid. Therefore, this kind of service has the typical characteristics of wide coverage and wide range connection. At present, the frequency of acquisition service is monthly, daily and hourly, and the acquisition content is mainly based on basic data and images. The connection density is generally 100 / km². In the future, the acquisition frequency tends to minute level, and the number of connections at least doubles. Under the direct reading mode, the number of connections is expected to increase by 50-100 times. The collected content also tends to be video and high-definition from the original simple data, and the connection demand will reach billions. Take the electricity information collection service as an example. When the collection object sinks to the user, the number of connections will reach thousands of levels, even thousands of levels per square kilometer. mMTC application scenario is an application scenario aiming at large-scale sensing and data acquisition. It mainly provides massive device connections and has the ability to support over 100 billion network connections. Therefore, the typical slicing scenario of information collection business is mMTC.

Mobile application business mainly includes UAV inspection, transmission and distribution robot inspection, mobile office, mobile video conference, etc. This kind of service has obvious demand for network bandwidth and mobility, and has typical characteristics of large bandwidth and flexible access.
Smart grid video services in mobile applications need to send back a lot of high-definition video. Taking the inspection service of transmission and distribution robot as an example, the required bandwidth is 4-10mbps. The application scenario of eMBB mainly includes hot areas with wide area coverage and high user density. Through 5G enabled smart grid, the high requirements of communication bandwidth are properly solved. The typical network slicing scenario is eMBB.

In the future, the requirements of different services for the network will be diversified. Different services in smart grid have different requirements for network delay, bandwidth, reliability, etc. In the 5G network, network slicing technology can be used to provide specific services for different services to meet their specific needs. Table 2 shows the requirements for bandwidth, delay, reliability and other service indicators of various services in smart grid. Bandwidth refers to the amount of data that can be transmitted online in unit time. The higher the bandwidth, the faster the transmission rate. The measurement of reliability in the network is generally expressed by the number of 9. The number of 9 represents the ratio of system normal working time to total time (generally referred to as one year of the use time). Different grid services have different delay requirements, usually the number of nines is 3 to 5[13]. Delay refers to the time required for a packet to be transmitted from one end of a network to another. Different power services have different requirements for delay.

Table 2. Key network requirements for typical service scenarios of smart grid

| Service category | Service name | Communication requirements | Number of terminal connections |
|------------------|-------------|---------------------------|-------------------------------|
| Control services | Distribution automation | $\leq 10$ms | $\geq 2$Mbps | X*10/km$^2$ |
|                  | Precise load control | $\leq 50$ms | 10kbps-2Mbps |                               |
| Information collection services | Advanced measurement | $\leq 3$s | 1-2Mbps | X*100/km$^2$ |
|                  | Electricity information collection | $\leq 5$s | 2.5Kbps | X*1000/km$^2$ |
| Mobile application services | Substation inspection robot | $\leq 200$ms | 4-10Mbps | 1-2 in the concentrated local area |
|                  | Video integrated monitoring of power distribution room | $\leq 200$ms | 20-100Mbps | 5-10 in local area |

3. System model

The system model of this paper is shown in Figure 1. We mainly consider the bandwidth allocation of the downlink in the RAN network slices. We assume that in the network there is a BBU pool and $K$ RRHs, which are expressed as $\mathbf{R} = \{1, 2, 3, \ldots, K\}$. Our model consists of three main elements: the RAN slices, users in the smart grid and the control manager. We indicate with $\mathbf{S} = \{1, 2, 3, \ldots, S\}$ the set of slices in the virtual network. Each slice $s$ has a set of users, which are expressed as $\mathbf{U} = \{1, 2, \ldots, U\}$. A set of users of a specific slice $S$ are denoted by $U_s$, and $u_s$ denotes a single user of the slice $S$. The allocation of the resources of each RRH among its associated users is denoted by $b_{u,k}$, where $b_{u,k}$ is the proportion of all the resources of RRH $k$ (denote by B) that are allocated to user $u$. The
$b_{u,k}$ is continuously chosen in the range $[0, B]$ despite the discrete nature of such resources. The transmission rate is computed as in (1),

$$R_{u,k} = b_{u,k} \log_2 \left(1 + \frac{p_k h_{u,k}}{\sum_{i \in S, i \neq k} p_{i} h_{u,i} + \sigma^2} \right)$$

(1)

Where $R_{u,k}$ is the achievable data rate of user $u$ in slice $s$ associated with RRH $k$, $b_{u,k}$ denotes the proportion of all the resources of RRH $k$ (denote by $B$) that are allocated to user $u$, the transmit power of RRH $k$ is denoted by $P_k$. $h_{u,k}$ denotes the channel gain between user $u$ and RRH $k$, which includes path loss, fast fading, shadowing and antenna gain. $\sigma^2$ denotes thermal noise power at the user.

Figure 1. Intra-slice and inter-slice resource allocation scenario

4. Problem formulation and solution

In this paper, we propose a two-layer dynamic resource scheduling strategy fully considering the service priority, fairness and user’s QoE requirements. One layer is used to determine the amount of resources for each slice (inter-slice scheduler), and the other is to assign and allocate resources to end users for each slice (intra-slice scheduler).

4.1 Inter-slice scheduler

First of all, we assign resources in the RRH to slices according to the slice requirements and slice weights. $w_s$ represents the weight factor of the slices, which depends on the type of business it serves. As discussed above, we define the priority of the power service according to the respective the QoS requirements. The stricter the QoS requirements, the higher the priority. Figure 2 shows the priority level of the power services. Accordingly, we define $w_1, w_2, w_3$ as the weight factor of uRLLC, mMTC and eMBB respectively. Slice demand refers to the number of users in slice. $L_{i,k}$ denotes the number of users belong to the slice $i$ in the RRH $k$. Resources are allocated to slice $s$ according to its weight and users number considering proportional fairness.

$$B_i = B \frac{w_i L_{i,k}}{\sum_{j=1}^{N} w_j L_{j,k}}, i \in S$$

(2)
Figure 2. Power service priority

4.2 Inter-slice scheduler

Intra slice resource allocation refers to the reallocation of resources allocated by slices to their users. Driven by the requirements of high quality power services such as distribution automation, emergency communications, precision load control, an appropriate level of quality of experience in smart grid is desired. In this paper, QoE is measured by considering the effective throughput experienced by the users, normalized according to their maximum demanded data rate. With this target, the resources allocated to a service with low priority can be reduced, if needed, down to the minimum amount capable of meeting the basic QoS requirements in order to admit new services with higher priority. $q_{u,s}$ is the QoE of user $u$ in the slice $s$, which is defined as:

$$q_{u,s} = \frac{R_{u,s,k}}{R_{s}^{\text{max}}}.$$  

$R_{u,s,k}$ is transmission rate of user $u$ in slice $s$ of RRH $k$, $R_{s}^{\text{max}}$ is the maximum data rate constraint of the user in packets per second.

In this paper we focus on maximizing the network utility while dynamic allocating resources fairly, and the optimization problem can be formulated as follows.

Slice $s$’s utility: $\alpha_{s,k} = \sum_{u \in U} q_{u,s,k} * w_s$  

Network utility $\alpha = \sum_{s \in S} \sum_{k \in K} \alpha_{s,k}$  

maximize $\alpha$  

s.t. $0 \leq b_{u,k} \leq B_{u,k}, \forall u, k$  

$\sum_{u \in U} b_{u,k} \leq B, \forall k$  

$R_{s}^{\text{min}} \leq R_{s} \leq R_{s}^{\text{max}}, \forall u \in U$  

$w_s \in \{0.2, 0.3, 0.5\}$ corresponding to eMBB, mMTC and uRLLC respectively. $q_{u,s}$ represents experience quality of user $u$ in slice $s$. Constraint (7) indicates that the bandwidth allocated by the user is limited by the bandwidth of the relevant slice. Constraint (8) means that the total bandwidth allocated to users is limited by the total bandwidth of the system. Constraint (9) indicates that the data rate received by the user is limited by the requirements of related slices.

5. Simulation results and analysis

In this section, we introduce the difference between utility maximize algorithm (UMA), priority algorithm (PA)\(^9\) and proportional fair algorithm (PFA). Then we analyze the advantages of our algorithm and give the comparative simulation results.
In order to evaluate the performance of our proposed scheme, we use MATLAB software for numerical simulation and analysis. When the power business arrives, we assume that there is a base station. Users in the base station are evenly distributed. Table 3 lists the parameters needed for the experiment.

| Parameter               | Value          |
|-------------------------|----------------|
| Number of cells         | One macrocell  |
| bandwidth               | 20MHz          |
| UEs distribution        | uniform        |
| Overall number of UEs   | 200            |
| Overall interval        | 10s            |
| TTI duration            | 0.01s          |

| Parameter               | Value          |
|-------------------------|----------------|
| Number of cells         | One macrocell  |
| bandwidth               | 20MHz          |
| UEs distribution        | uniform        |
| Overall number of UEs   | 200            |
| Overall interval        | 10s            |
| TTI duration            | 0.01s          |

Utility maximize algorithm (UMA): in this algorithm, we mainly consider the resource allocation of intra slice and inter slice. The resource allocation between slices is mainly based on the priority of slice and the number of users. The resources in slices are allocated according to the goal of maximizing network utility.

Proportional fair algorithm (PFA): the algorithm is a classic algorithm of resource allocation. The resources are evenly allocated to each user in the network.

Priority algorithm (PA): the algorithm is based on priority allocation and heuristic algorithm to get the optimal resource allocation strategy. The overall problem in resource allocation is to consider both inter slice and intra slice priorities to maximize user experience quality. In formula (10), $\mu_u$ represents the priority of the user, $\rho_s$ denotes the priority of the slice.

$$\max \sum_{s \in S} \left[ \sum_{u \in U_s} \left( \frac{r_{u,s}}{\mu_u} \right)^{\mu_u} \right] \rho_s$$

Figure 3 shows the comparison of the overall network utility of the resource allocation strategy in this paper, the resource allocation strategy based on proportional fairness algorithm and the resource allocation strategy according to priority under different number of users. It can be seen that with the increase of the number of users, the overall network utility is on the rise. In the PFA, with the increasing number of users, the growth rate of network overall utility value gradually slows down. When the number of users is less than 140, the overall network utility of UMA and PA is almost the same. When the number of users is more than 140, UMA has a certain advantage over PA in the overall network utility.

Figure 4 is a comparison of the system throughput of the three algorithms with different number of users. From the simulation results, the average throughput of UMA is higher than that of PFA and PA. With the increase of the number of users, the throughput of the system is also increasing. With the consumption of system resources, the throughput difference between the three gradually decreases. This also explains the simulation results in figure 4. The higher average throughput is the guarantee of higher user experience quality, which reflects the strong advantages of UMA.

Figure 5 shows the comparison of user satisfaction among the resource allocation algorithm, proportional allocation algorithm and priority allocation algorithm in the case of insufficient resources. As we can see from the figure, compared with proportional fair allocation algorithm, UMA and PA have greatly improved the quality of user experience in the three kinds of slices. UMA algorithm provides better fairness for users in different slices. Although the QoE value of PA algorithm is slightly higher than that of UMA in uRLLC slice, the QoE value of PA is very low in mMTC slice, which is lower than that of PFA and UMA. Because in PA, when the resource is scarce, the resource will be allocated to the slice with high priority. The pressure of insufficient resources is completely transferred to slices with low priority, which will lead to low satisfaction of slices with low priority. The algorithm in this paper is based on the two-layer allocation algorithm, which ensures that the slices with high priority are allocated with resources first, and the slices with low priority will also be allocated with certain resources.
We can observe that UMA ensures that the QoE difference of these slices is lower than that of PA. That is to say, our resource allocation algorithm ensures the fairness of users.

Figure 3. Comparison of network utility under three strategies

Figure 4. Comparison of system throughput

Figure 5. Comparison of average QoE in three kinds of slices

6. Conclusion
With the wide application of 5G network slicing in smart grid, resource allocation of slice will become one of the problems to be solved. Based on the analysis of current network slicing resource allocation strategy and considering that slices with the same service type have different weights, we propose a two-layer network slicing resource allocation strategy based on dynamic weight and utility. Different resource allocation methods within and between slices can better meet the QoE of users in different slices. Theoretical analysis and experimental results show that the proposed UMA strategy has great advantages in terms of average throughput and QoE satisfaction. In the future, we will further study the resource allocation scheme in smart grid services.

Acknowledgment
This paper is supported by “Project Evolution of Power Wireless Private Network and Application Analysis of 4G and 5G Technologies (5442XX190007)”.

References
[1] Wu, Q. Wu, C. (2019) Open ICT-PaaS platform enabling 5G network slicing. IET Communications, 13: 1242-1252.
[2] Foukas, X., Patounas, G., Elmokashfi, A. (2017) Network Slicing in 5G: Survey and Challenges. IEEE Communications Magazine, 55: 94-100.
[3] Ju, P., Zhou, X. (2018) Smart Grid Plus research overview. Electric Power Automation Equipment, 38: 1-11.

[4] Meng, S., (2019) RAN Slice Strategy Based on Deep Reinforcement Learning for Smart Grid. In: Computing, Communications and IoT Applications. Shenzhen. pp. 6-11.

[5] Kalil, M., Shami, A., (2014) Wireless resources virtualization in LTE systems. In: IEEE Conference on Computer Communications Workshops. Toront. pp. 363–368.

[6] He, J., Song, W., (2015) AppRAN: Application-oriented radio access network sharing in mobile networks. In: IEEE International Conference on Signal Processing for Communications. London. pp. 3788–3794.

[7] Caballero, P., Banchs, A., (2017) Network slicing games: Enabling customization in multi-tenant networks. In: IEEE Conference on Computer Communications. Atlanta. pp. 1–9.

[8] Albonda, H.C., Pérez-Romero, J. (2019) An efficient RAN slicing strategy for a heterogeneous network with eMBB and V2X services. IEEE Access, 7:44771–44782.

[9] Jiang, M., Condoluci, M., Mahmoodi, T., (2016) Network slicing management & prioritization in 5G mobile systems, In: European Wireless Conference. EW. pp. 197-202.

[10] Marabissi, D., Fantacci, R. (2019) Highly Flexible RAN Slicing Approach to Manage Isolation, Priority, Efficiency. IEEE Access, 7:97130-97142.

[11] Chen, Q., Liu, C. (2018) Dynamic Resource Scheduling Strategy for 5G network slices based on improved greedy algorithm. Journal of Network and Information Security, 4: 60-68.

[12] Tang, J., Shim, B. (2019) Service Multiplexing and Revenue Maximization in Sliced C-RAN Incorporated With URLLC and Multicast eMBB. IEEE Journal on Selected Areas in Communications, 37: 881-895.

[13] Afolabi, I., Taleb, T. (2018) Network Slicing and Softwareization: A Survey on Principles, Enabling Technologies, and Solutions. IEEE Communications Surveys & Tutorials, 20: 2429-2453.