Intelligent and Converged Networks
2021, 2(2): 101–107

ISSN 2708-6240
DOI: 10.23919/ICN.2020.0017

A flexible scheduling algorithm for the 5th-generation networks

Lanlan Li*, **Wentao Shao, and Xin Zhou**

Abstract: At present, the 5th-Generation (5G) wireless mobile communication standard has been released. 5G networks efficiently support enhanced mobile broadband traffic, ultra-reliable low-latency communication traffic, and massive machine-type communication. However, a major challenge for 5G networks is to achieve effective Radio Resource Management (RRM) strategies and scheduling algorithms to meet quality of service requirements. The Proportional Fair (PF) algorithm is widely used in the existing 5G scheduling technology. In the PF algorithm, RRM assigns a priority to each user which is served by gNodeB. The existing metrics of priority mainly focus on the flow rate. The purpose of this study is to explore how to improve the throughput of 5G networks and propose new scheduling schemes. In this study, the package delay of the data flow is included in the metrics of priority. The Vienna 5G System-Level (SL) simulator is a MATLAB-based SL simulation platform which is used to facilitate the research and development of 5G and beyond mobile communications. This paper presents a new scheduling algorithm based on the analysis of different scheduling schemes for radio resources using the Vienna 5G SL simulator.

Key words: 5th-Generation (5G); radio resource management; channel status information reporting; scheduling schemes; Vienna simulator

1 Introduction

According to the 3rd Generation Partnership Project timelines, New Radio Release 15 (NR Rel-15) has been frozen since 2019. With the evolution of the 4th-Generation (4G) wireless technologies, the 5th-Generation (5G) mobile communication system follows 4G's technical route and makes new breakthroughs in some key technical points. 5G networks still adopt orthogonal frequency division multiplexing transmission technologies and higher-order modulation technologies. Moreover, 5G systems can be easily integrated with advanced technologies, such as massive Multiple-Input Multiple-Output (MIMO), code block group based Hybrid Automatic Repeat reQuest (HARQ) with more HARQ processes and pre-emptive mechanisms, beam management technology for massive MIMO, and Adaptive Modulation and Coding (AMC), to further improve the performance of the system. Beam management technology enables signal transmission in narrow beams with high-frequency bands. These new features enable the 5G system to support a variety of services and achieve a relatively high system throughput. Accordingly, the introduction of these new technical features has put forward high requirements for the Media Access Control (MAC) layer scheduling algorithm. The scheduling algorithm is designed in more flexible time scales. Moreover, to make full use of these wireless technologies and radio time-frequency resources, the MAC layer scheduling algorithm must consider wireless factors to improve the networks' Key Performance Indicators (KPIs), including the Quality of Service (QoS) requirements, channel quality, and service queue

• Lanlan Li is with Purple Mountain Lab, Nanjing 210000, China. E-mail: lilanlan@pmlabs.com.cn.
• Wentao Shao and Xin Zhou are with the School of Information Science and Engineering, Southeast University, Nanjing 210000, China. E-mail: wentao1996@aliyun.com; xzhou1105@163.com.
* To whom correspondence should be addressed.
Manuscript received: 2020-07-08; revised: 2020-09-18; accepted: 2020-11-25

status of users, and consider the resource utilization of User Equipment (UE) from the point of view of the physical layer and MAC layer to optimize resource allocation. This step requires the scheduling algorithm to not only have a reliable and robust performance but also meet the requirements of computational complexity and realizability. Based on the above statements, to solve the lack of wireless resources in the wireless network and guarantee the application service quality under the dynamic changes in the bandwidth, large delay, high-speed mobility, and other factors in the wireless network architecture, an advanced scheduling algorithm should be proposed and applied.

5G technology has two network architectures: Standalone (SA) and Non-Standalone (NSA). In the two network architectures, the scheduler is located on eNodeB (eNB) (or gNodeB (gNB)) independently, so the scheduling algorithm in this paper is applicable to both architectures. In Section 2, these two types of network architectures are briefly introduced.

2 5G system architecture

2.1 NSA architecture

According to Technical Specification (TS) 38.300[1], 5G networks support Multi-Radio Dual Connectivity (MR-DC) operations. By doing so, network deployment can smoothly transform from 4G to 5G. Several network architectures have been designed for MR-DC, in which two Radio Access Network (RAN) elements can work separately in the user plane with their own independent schedulers for allocating radio resources[2]. UE in the Radio Resource Control (RRC) connected status is configured to utilize radio resources which is provided by the two distinct schedulers mentioned above, which are located in two different Next Generation (NG)-RAN elements and the two elements are connected via a non-ideal backhaul: one provides 5G access and the other one provides either 4G or 5G access. In this architecture, UE is connected to both 4G eNB and 5G gNB using Uu interfaces between RAN and user equipment in the radio side, but all communications (signaling and data) through S1-Mobility Management Entity (MME) or S1-U interface are transmitted to the Long Term Evolution (LTE) core network, where S1 represents the interface between an eNB and an EPC, providing an interconnection point between the EUTRAN and EPC and it is also considered as a reference point (see Fig. 1).

2.2 SA architecture

In the 5G SA network architecture, an RAN that supports SA 5G has a common characteristic that it is connected with 5G Core (5GC). With the evolution of 5G technology, 4G will be gradually replaced by 5G in the network deployment of operators. The 5G network element gNB provides the user plane and control plane protocol terminations toward the UE in the RAN side. gNBs are connected to the network element Access and Mobility management Function (AMF) in 5GC by means of the NG interfaces (see TS 23.501[3] and Fig. 2).

3 5G protocol stack architecture

3.1 5G protocol stack

The 5G radio protocol stack architecture is almost the same as the 4G radio protocol stack architecture. As shown in 4G technology, the 5G wireless protocol stack is divided into two planes: user plane and control plane. The user interface protocol stack is the protocol layer set used for user data transmission, and the control plane protocol stack is the protocol layer set for system control signaling transmission. In other words, different types of data are processed by different stacks. If the data are signaling messages, then they go through the

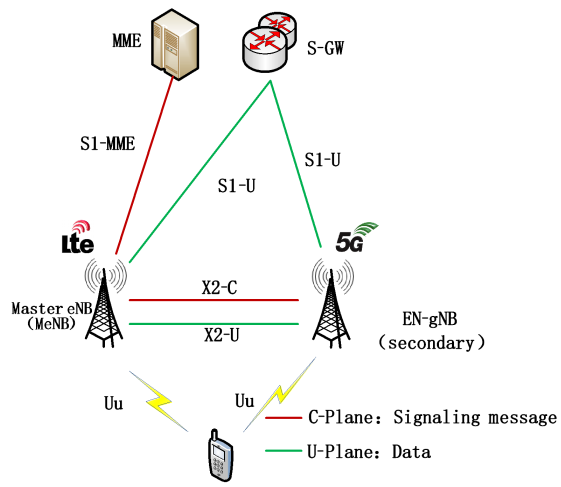


Fig. 1 NSA architecture. Here S-GW represents Serving Gate W, EN represents EUTRA-NR, and X2 is the interface between two eNBs.

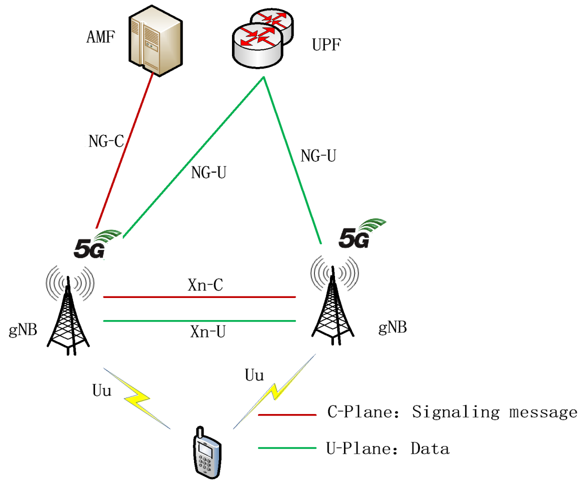


Fig. 2 SA architecture. Here, UPF represents user plane function and Xn is the interface between two gNBs.

control plane stack, and if they are user data, then they go through the user plane stack. The user plane and control plane are made up of a common stack structure, namely, Physical layer (PHY)/MAC/radio link control/packet data convergence protocol, but there is a little difference between them. The top layer of the protocol stack is different between the control plane and user plane. In the protocol stack of the user plane in 5G gNB, a new layer called service data adaptation protocol is added on the top of the radio user plane stack, and it is responsible for mapping the QoS flow to the corresponding data radio bearer. Because gNB is connected to the network element UPF via the NG-U interface in the user plane, the protocol stack in UPF is the same as the user plane in gNB. In the protocol stack of the control plane in 5G gNB, similar to 4G, the two protocol layers RRC and Non-Access Stratum (NAS) sit at the top of the protocol stack. NAS layer signaling is transparently transmitted to the 5GC network element AMF through gNB. Figure 3

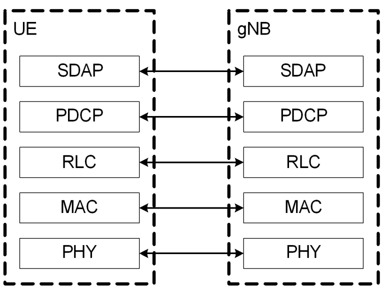


Fig. 3 User plane protocol stack. Here SDAP represents Service Data Adaptation Protocol, PDCP represents Packet Data Convergence Protocol, and RLC represents Radio Link Control.

shows the protocol stack for the user plane, and Fig. 4 shows the protocol stack for the control plane.

3.2 Radio resource management

Radio Resource Management (RRM) refers to a multitude of processes, strategies, and algorithms for monitoring parameters, such as transmission energy, user allocation, data rates, handover criteria, modulation scheme, and error coding scheme. Under the condition of a limited bandwidth, RRM provides service quality guarantee for wireless UE in the network[4]. Its aim is to flexibly allocate and dynamically adjust the available resources of the wireless transmission part and network with an uneven distribution of network traffic and fluctuation of channel characteristics due to channel fading and interference. This process aims to maximize wireless spectrum utilization, prevent network congestion, and keep the signaling transmission load as small as possible. The RRM module mainly includes the algorithm concerning the following parts: power control, channel allocation, scheduling, handoff, access control, load control, and end-to-end QoS.

3.3 Channel state information reporting

Massive MIMO is one of the key technologies of 5G. The key to improve the performance of the massive MIMO system is the acquisition of Channel State Information (CSI). During downlink MIMO system data transmission, the user estimates and quantizes the channel through the Reference Signal (RS). On one hand, it is used to demodulate the transmission information, and on the other hand, the quantization results are fed back to gNB, which selects the appropriate precoding and modulation coding scheme according to the CSI. CSI is the general name of a set of parameters used in 5G

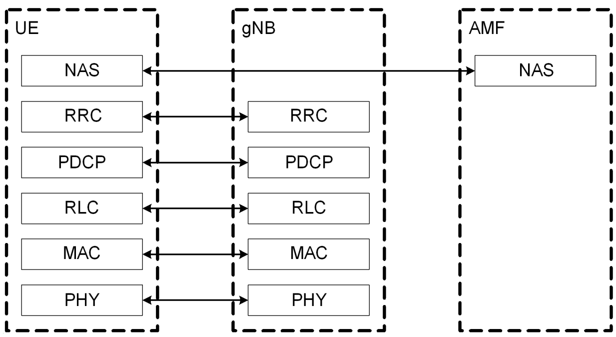


Fig. 4 Control plane protocol stack.

networks. Compared with 4G networks, there are more components on CSI in 5G networks. The following are the components of CSI in NR (based on TS 38.214[5]):

- Channel Quality Information (CQI);
- Precoding matrix indicator;
- CSI-RS resource indicator;
- Layer indicator;
- Rank indicator and/or Layer one RS Received Power (L1-RSRP).

CQI reporting is very important because it enables gNB to estimate the quality of the downlink channel. Each CQI is calculated as a quantized and scaled measure of the experienced Signal to Noise plus interface Ratio (SINR). The AMC module in RRM selects the right modulation and coding scheme according to the CQI which was reported to maximize the throughput[6]. In some scheduling algorithms, including the algorithm in this paper, CSI information is considered to design strategies for RRM to allocate the radio resource. The above description is summarized in Fig. 5.

3.4 Conventional scheduling algorithm

Scheduling algorithms are not new topics in the evolution of communication technology. Three conventional scheduling algorithms are used in 5G technology standardizations[7].

(1) Round robin

The Round Robin (RR) algorithm is a comparatively simple algorithm without complex logic deduction. In the RR scheduling scheme, RRM allocates radio resources to the users in the order of first in first served, and UE can get radio resources with the same probability. As a result, the users served by the same gNB are equally scheduled without considering the channel status. The RR scheme is easy to implement without considering the abstract theory. It provides high performances and fairness for all users in a short term. The drawback of this algorithm is that it cannot make full use of the time-varying characteristics of the wireless channel to bring multi-user diversity gain to the system.

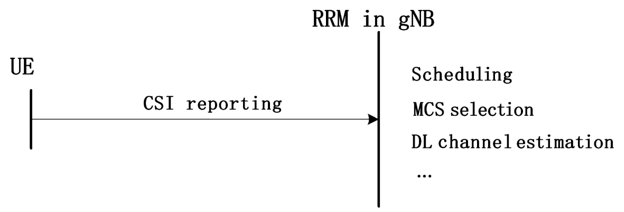


Fig. 5 CSI reporting. Here MCS represents Modulation and Coding Scheme and DL represents Downlink.

(2) Best CQI

The best CQI is another scheduling scheme used in gNB. In this algorithm, when selecting the schedulable one from all UEs attached in one cell in gNB, only the UE with the best CQI is selected. That is, the user with the best service channel condition is always given the scheduling opportunity. In this way, the user with a good service channel condition can transmit data until there are users with better service channel quality. Therefore, the system adjusts the number of multi-users continuously according to the transmission scheme. This algorithm is adapted to the time-varying and fading characteristics of the wireless channel to ensure the maximum throughput of the system. In this scheduling algorithm, users with poor channel conditions (i.e., cell-edge users) are less likely to be scheduled[7]. This scheduling approach can ensure the maximum multi-user diversity gain. From the perspective of fairness, this algorithm can be regarded as the most unfair algorithm.

(3) Proportional fair

The Proportional Fair (PF) algorithm is widely used in the industry. In the PF algorithm, RRM assigns a priority to each user served by gNB. The metrics of priority depend on the design of vendors. In each scheduling time interval, users with high priority will be scheduled first. Proper metrics are the key of the PF algorithm. It can effectively improve the network performance. On one hand, the PF scheduling algorithm makes full use of the time-varying characteristics of the user channel, and on the other hand, it ensures the balance between multi-user diversity and fairness, so it is a common dynamic resource scheduling algorithm in wireless communication systems. The PF scheduling algorithm provides a good trade-off between the system throughput and fairness to maximize the entire network throughput and to guarantee fairness among flows. It allocates radio resources by considering the experienced channel quality and the past user throughput. The PF algorithm design has many achievements in this field. One metric of priority in the PF algorithm is the ratio between the instantaneous flow available for the i-th flow and the medium flow elaborated at the moment $(k-1)$[8].

$$w_{i,j} = \frac{r_{i,j}}{\overline{R}_i},$$

where $r_{i,j}$ is the flow rate assigned to the i-th flow during the k-th Transmission Time Interval (TTI) and $\overline{R}_i$ denotes the average transmission data rate estimating. The PF scheduling algorithm uses the average data transmission rate to represent the fairness between users, so it has good long-term fairness. However, the PF scheduling algorithm cannot guarantee the short-term fairness between users.

3.5 New scheduling algorithm

In the above-mentioned PF algorithm, the flow rate and average rate are considered. In fact, more factors should be involved in the model to define priority. CQI is the quality of the wireless channel fed back by the UE to the base station, which is usually obtained according to the quality of the received RS. According to the CQI feedback from the terminal and other information, the base station implements service scheduling. User delay refers to the time delay when the UE sends data packet. The 1 ms delay of 5G networks was originally mentioned by the International Telecommunication Union International Mobile Telecommunications (IMT)-2020 M.2410-0 (4.7.1) on the minimum design requirements of the IMT-2020 system for ultra-reliable low-latency communication. The design of the scheduling algorithm should be ensured to make the packet transmission delay as small as possible. In 5G networks, a scheduling algorithm is designed based on the QoS model. The QoS flow has the finest granularity for scheduling. The PF scheduling algorithm in this paper provides a trade-off among channel quality, average transmission rate, and packet transfer delay to maximize the entire network throughput and to guarantee fairness among scheduled UEs in a short term. It allocates radio resources by considering the experienced channel quality, average transmission rate, and packet transfer delay. For this algorithm, the metric is defined as the ratio of the instantaneous flow available for the i-th flow and the medium flow which is elaborated at the moment $(k-1)$ defined below:

$$\frac{f_1(\text{SINR}_i)}{\text{avg\_rate}} \cdot \frac{f_2(\text{package\_delay}_{k-1})}{\text{DelayTh} - f_3(\text{package\_delay}_{k-1})},$$

where $f_1(\cdot)$, $f_2(\cdot)$, and $f_3(\cdot)$ are the adjustable functions adapting to the network performance requirements, and they can be linear or nonlinear functions. In the simulation of this study, let $f_1(\cdot)$, $f_2(\cdot)$, and $f_3(\cdot)$ be linear functions, i.e., $f(x) = x$. avg_rate is the average transmission rate of packet. Factor SINR_i is the quantized CQI of the wireless channel transmitting the i-th flow. DelayTh is the threshold for the transmission delay, which is set according to the network performance and can be configured in gNB. Assuming that the current scheduling time interval is the k-th TTI, the factor $\text{package\_delay}_{k-1}$ denotes the packet transmission delay in the last TTI. Thus, UE with a high SINR and long packet transmission delay is likely to be scheduled.

4 Simulation parameter and result

4.1 Vienna 5G system-level simulator

The 5G System-Level (SL) simulator is one of the Vienna Cellular Communications Simulators (VCCS), which is a simulator suite for academic usage. It was programmed using MATLAB and C++ languages. The simulator suite includes a link and SL simulators for research purpose in 4G and 5G wireless mobile communication fields. The simulations based on VCCS and the new scheduling algorithm are added into the simulator.

4.2 Simulation parameter

We analyze the performance in terms of the throughput using the fore schedulers described above. Four scheduling algorithms are simulated individually in the same simulator with the same basic parameters. The basic simulation parameters are listed in Table 1.

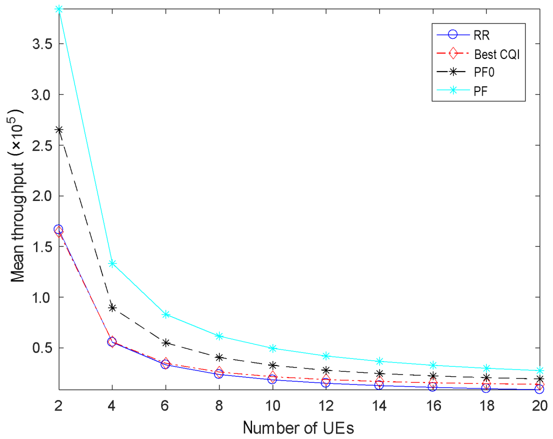
Table 1 Simulator parameters.

Simulation parameter	Value
Number of chunks	1
Slot per chunk	100
Time between chunks in slots (ms)	0
Antenna	Omni directional
Number of receiving antennas	1
Number of transmitting antennas	4
Transmit power (W)	40

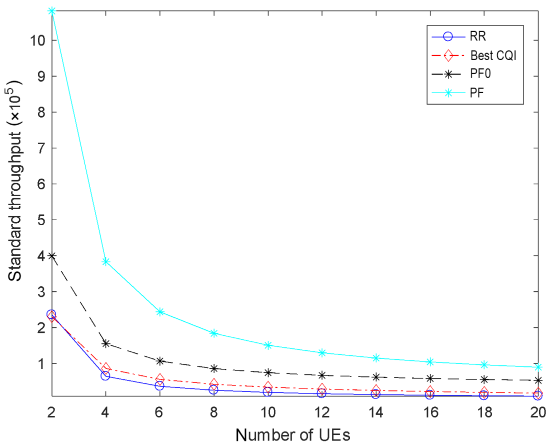
4.3 Simulation result and analysis

In this section, we present the simulation results for the SL simulations. The SL simulation results show that the newly proposed scheduling algorithm provides higher throughput than the best CQI, RR, and conventional algorithms.

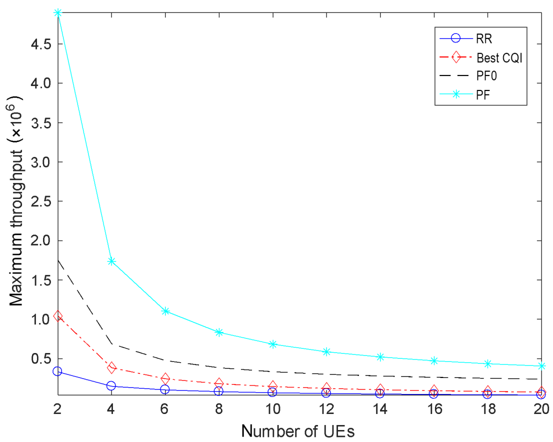
Figure 6 shows that the throughput of the PF algorithm in this paper is larger than that of the existing algorithms. This is mainly because the new PF algorithm takes into account the user packet delay. After gNB is configured with a given delay threshold value, when the packet transmission delay is too large, a larger scheduling priority will be obtained, and the time-frequency resources will be allocated preferentially to obtain a larger throughput.



(a) Mean throughput



(b) Standard throughput



(c) Maximum throughput

Fig. 6 Comparison of simulation throughput results of four scheduling algorithms. Here PF0 represents the existing Proportional Fair algorithm.

The simulation result shows that the throughput has been improved by scheduling algorithms proposed in this paper. This is because compared with the other three scheduling algorithms, the PF scheduling algorithm in this paper takes into account the delay of transmitted data packets. For a given predetermined reasonable packet delay threshold, the larger the packet delay is, the higher the proportional fairness factor is. So the time-frequency resource will be scheduled firstly for corresponding packet transmission. The network throughput can be improved effectively.

5 Conclusion

In this comparative study, we have evaluated the performance of scheduling algorithms RR, PF, PF0, and best CQI in a 5G environment for enhanced mobile broadband user experience. The SL simulation results show that the new proposed scheduling algorithm can provide higher throughput than the best CQI, RR, and conventional PF algorithms. In the future research work, algorithms will be considered in more details. Another research direction is the form of function $f(\cdot)$, which needs to be extended, and more simulations are needed.

References

[1] *Overall Description*, *Stage-2*, *Version 15.9.0*, 3GPP TS 38.300, NR, 2020.

[2] *Multi-Connectivity*, *Overall Description*, *Stage-2*, *Version 15.8.0*, 3GPP TS 37.340, NR, 2020.

[3] *System Architecture for the 5G System* (*5GS*), *Version 15.9.0*, 3GPP TS 23.501, 2020.

[4] R. Agustí, Radio resource management in beyond 3G systems, in *Proc. 2006 IEEE Mediterranean Electrotechnical Conf.*, Malaga, Spain, 2006, pp. 569–574.

[5] *Physical Layer Procedures for Data*, *Version 15.9.0*, 3GPP TS 38.214, NR, 2020.

[6] R. Bruno, A. Masaracchia, and A. Passarella, Robust adaptive modulation and coding (AMC) selection in LTE

systems using reinforcement learning, in *Proc. 2014 IEEE 80th Vehicular Technology Conf.* (*VTC2014-Fall*), Vancouver, Canada, 2014, pp. 1–6.

[7] M. H. Habaebi, J. Chebil, A. G. Al-Sakkaf, and T. H. Dahawi, Comparison between scheduling techniques in long term evolution, *IIUM Engineering Journal*, vol. 14, no. 1. pp. 67–76, 2013.

[8] G. Piro, L. A. Grieco, G. Boggia, F. Capozzi, and P. Camarda, Simulating LTE cellular systems: An open-source framework, *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 498–513, 2011.

Lanlan Li received the master degree in applied mathematics from Xinjiang University in 2008 and the PhD degree in electrical engineering from Southeast University in 2005. She is currently a senior engineering at Purple Mountain Lab. She has been involved in several national and entreprise's projects. Her research interests lie in radio resource management, signal processing for digital communications, and AI application in communication field.

Wentao Shao received the BEng degree in electronic engineering from Chongqing University in 2018. Since September 2018, he has been pursuing the MS degree at School of Information Science and Engineering, Southeast University. His current research interests include stochastic modelling, wireless network optimization, firmware development, and network virtualization. His presentation of the work: An optimal estimation of base station density based on a new 5G transmission model in ICTC 2020 was selected as the best of the session.

Xin Zhou recieved the bachelor degree in information engineering from Nanjing University of Aeronautics and Astronautics, China in 2018. He is currently pursuing the master degree at Southeast University majoring in electronic and communication engineering. His research interests include communication channel modeling, radio resource scheduling, 5G system simulation, and application of reinforcement learning in communication field.