Multi-user Space-Time Modulation Scheme for Ultra Reliable Low Delay Communication

Gaolong Jiang*
School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Jiangsu, China

*Corresponding author e-mail: b18011725@njupt.edu.cn

Abstract. 5G (The Fifth Generation) mobile networks are expected to achieve significant improvements in several key performance metrics, where low latency is urgently needed by many emerging IoT businesses, such as remote surgery, augmented reality, haptic Internet and IoV. In this paper, a multi-user space-time modulation scheme for super-reliable low-delay communication is studied. Based on 5G technology support in the new wireless frame structure of the coexistence of a variety of business, for business and eMBB URLLC business coexist in the data transmission of conduct the thorough research to the downlink scheduling mechanism, this mechanism allows URLLC business data to grab some data are transmitted eMBB wireless resources, occupied in eMBB business to transmit data on punch to URLLC business can be scheduled in time, to meet the requirements of URLLC business ultra-low delay.

Keywords: Super Reliable Low Delay Communication, Multi-user Modulation, Scheduling Algorithm, Punching Scheduling

1. Introduction
After 2020, the global mobile communication will enter the fifth generation. Its spectrum efficiency, energy efficiency and the number of connected devices supported at the same time have been greatly improved, which can better support the rapid development of the mobile Internet and the Internet of Things. 5G includes three typical applications that enable a wide variety of enhanced wireless services. Enhanced Mobile Broadband (eMBB) can significantly improve user data rates. Massive Machine Type Communications (MMTC) will connect billions of orders of magnitude low data rate IoT devices; Ultra-reliable and Low Latency Communications (URLLC) will support short-packet communication with ultra-high reliability and extremely Low delay [1]. Specifically, in URLLC, the user-side latency for transmitting a 32-byte packet is reduced to less than 1ms, an unprecedented low latency performance that is sure to lead to a number of promising applications. Thanks to URLLC related technologies, emerging applications such as intelligent transportation and "Industry 4.0" are gradually changing people's lives. Although URLLC's vision is good, it also faces all-round challenges from theoretical research and engineering implementation. There are some defects in the current network architecture and technology, and the research on URLLC is still in its infancy. Compared to
traditional data traffic businesses, URLLC requires new architectures and technologies, including smart and flexible network slicing and new port designs, to more efficiently meet diverse latency and reliability requirements. Therefore, it is urgent to carry out related research on low delay and high reliability communication.

Emerging application scenarios and services bring new demands and challenges to wireless communication systems at the same time. In order to meet the extremely high real-time reliability requirements of emerging applications, wireless communication systems should be capable of supporting millisecond end-to-end delay and close to 99.999% reliability. Nielsen proposed a framework based on Reinforcement Learning (RL), which can select the scheduling strategy that can maximize user satisfaction in each Transmission Time Interval (TTI) according to different QoS of users [2]. According to the finite code length theory, Ryu discussed the research directions and challenges of large 5G connection, low delay and high reliability communication, and further pointed out the importance of control overhead optimization for short frame transmission [3].

From the above research status, it can be seen that the current research on low delay and high reliability communication is still immature, and further research is needed in many aspects. The cross integration of information and communication technology with other technologies has become a new trend of future development. The communication with low time delay and high reliability will eventually "serve" the application of vertical industries. However, the researches of control, perception and communication are separated from each other at the present stage, and it is urgent to integrate the researches.

2. User Scheduling Algorithm with Low Delay and High Reliability

2.1. Meaning of Delay and Reliability

2.1.1. Meaning of Delay. End-to-end (E2E) delay refers to the time required for a packet to travel from the sending node to the receiving node in the network, and is an indicator to measure the network transmission performance [4]. The generalized end-to-end delay mainly consists of the time experienced by the packet in the wireless access network and the core network, while the narrow end-to-end delay mainly describes the time experienced by the packet in the wireless access network. Generally speaking, the (narrow) end-to-end delay is mainly composed of propagation delay, transmission delay, queuing delay, computation or processing delay and other parts [5]. Specifically, their respective meanings are:

1) Transmission delay

Propagation delay, also known as propagation delay in media, refers to the time required for a packet to travel from the sending node to the receiving node in the physical medium. Since the main medium of wireless communication is electromagnetic wave, the propagation delay can be calculated as the ratio of the transmission distance to the speed of light C, namely:

$$L_{\text{Propagation}} = \frac{D}{c}$$

2) Physical layer transmission delay

Physical layer transmission delay, also known as air interface delay, refers to the time required for data packet to travel from the sending node to the receiving node in the air interface of wireless access network. A simple and common way to model this is to calculate it as the ratio of the amount of information to be transmitted to the communication rate.

3) MAC layer queuing delay

MAC queuing delay, also known as scheduling delay, refers to the time that packets experience in the queue. Specifically, when the arrival rate of a packet is greater than its departure rate, a packet queue will be generated, and the latter packet will wait for the previous packet to leave the queue before being sent. This waiting time is the so-called queuing delay.
4) Delay in calculation or processing
The computation or processing delay is the time required to perform various operations on a packet. Computing or processing latency is heavily dependent on the capabilities of the hardware platform, so there is usually no uniform rule to model them.

2.1.2. Meaning of Reliability. Reliability is another important index to measure system performance. Usually, reliability is mainly composed of equipment reliability and network (communication) reliability. For network reliability, it can be divided into physical layer reliability, MAC layer reliability and other aspects [6-7]. Specifically, their respective meanings are:

1) Equipment reliability
Equipment reliability refers to the probability that a device (also known as an entity) works normally and stably. The common modeling method is the complementary cumulative distribution function, that is, the probability that the equipment life \( X \) is greater than a certain threshold \( t \):

\[
Pr[X > t] = 1 - F_X(t)
\]  

(2)

2) Physical layer reliability
Physical layer reliability refers to the error probability of information transmission. In general, when the probability of error increases, the network communication will become unreliable. Although there are many modeling methods of physical layer reliability, such as bit error rate, block error rate, interrupt probability, etc., they are all the reflection of error probability.

3) MAC layer reliability
MAC layer reliability, usually refers to queue stability. Consider a dynamic queue \( W(t) \) whose queue stability can be defined as

\[
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[W(t)] < \infty
\]  

(3)

2.2. User Scheduling Algorithm
The algorithm proposed in this paper first selects users from the scheduling result set of non-real-time services for punching, and then selects users from the real-time business. On this basis, in order to avoid users with poor channel quality being punched all the time, a weighting factor, namely the number of punching, is introduced to improve the fairness among users [8].

At the base station side, a counter is set for each EMBB user, which is used to record the number of holes punched by URLLC data transmission, and the update period of the number of holes is set as 5ms. The decision priority formula for IPS is shown below.

\[
\begin{align*}
M_{v,p}(t) &= \frac{r_{v,p}(t)}{R_v(t)} \times \frac{1}{W_p(t)} \times \frac{1}{1 + C_v(t)} \\
M_{v,p}(t) &= \frac{r_{v,p}(t)}{R_v(t)} \times \frac{1}{W_p(t)} \times (r_v - t_v) \times \frac{1}{1 + C_v(t)}
\end{align*}
\]  

(4)

As can be seen from the type, \( t_v \) difference on behalf of the near and far from the time delay threshold, for the user the nearer the time delay threshold, if the punch and cause for the retransmission, the greater the likelihood of its packet loss, or the packet due to the time of transmission of more than delay threshold and meaningless, and also wasted a wireless bandwidth resources [9-10]. For users far away from the delay threshold, the packet retransmission caused by punching can be tolerated to a certain extent, thus improving the user's service quality and reducing the packet loss rate of the system.

The process is described as follows:
(1) Execute the scheduling algorithm at the beginning of each TTI(0.143ms) interval;
(2) Determining whether the queue to be scheduled for buffer URLLC user is empty, if so, jump to step 1, if not, execute step 3;
(3) Determine whether there are any remaining resources in the current TTI. If there are any remaining resources, skip to Step 7 to schedule URLLC users. If there are no remaining resources, it means that EMBB users occupy all the wireless resources, and then execute Step 4;
(4) Determining whether the PRB set pnRT allocated by the business containing non-real-time EMBB users is empty; if so, jump to step 8; if not, execute step 5;
(5) According to the algorithm, select the allocated PRB corresponding to EMBB user from the non-real-time PRB set PNRT for punching scheduling;
(6) Determine the scheduling. If the scheduling is over, go back to Step 1 and execute the above steps.

3. Simulation Experiment

3.1. Scenario and Parameter Settings
In this paper, MATLAB is used as the simulation platform to verify the performance of the algorithm, and the performance of the improved IPS algorithm and PS algorithm is simulated and compared. Because the mobility of users is not considered in this paper, the users in the cell obey the relatively static random uniform distribution. The cell contains users of two types of business: EMBB and URLLC. VoIP business model is used for EMBB real-time business, FTP model is used for non-real-time EMBB business, and the business of URLLC users follows the Poisson distribution arrival process. As shown in Table 1, is the simulation parameter.

| Simulation parameters                  | Value                   |
|----------------------------------------|-------------------------|
| Number of base stations                | Single cell             |
| System bandwidth                       | 20MHz                   |
| System contains RB numbers             | 100                     |
| UE distribution in the community       | Random distribution     |
| Number of users in the cell            | 30                      |
| Dduplex                                | FDD                     |
| PHY parameter set                      | 15kHz                   |
| TTI size                               | URLLC=0.143ms eMBB=1ms  |

3.2. Content of the Experiment

3.2.1. Fairness Comparison. When multiple users communicate, the channel quality may vary between users. If the highest throughput is blindly pursued, users with poor channel quality may be kept waiting for resources, resulting in unfairness among users. The fairness of users can be measured by the waiting time for scheduling and the data transmission rate of users. If the waiting time between users is basically the same, then we think it is fair. The data transfer rate is basically the same, and we think it's fair. In addition, there are many other indicators to measure the fairness between users, which will be determined according to the actual situation.

3.2.2. Throughput Comparison. The throughput of an individual user is called user throughput, and the throughput of the whole system is called system throughput. The more bits passed by the channel per unit time, the greater the throughput of the system, and the greater the throughput of the system represents the greater the data transmission capacity of the whole system.

3.2.3. Analysis of Packet Loss Rate. The rate of packet loss is often one of the indicators of a system's superiority. When the data is transmitted in the channel, if the delay exceeds the threshold value, the
packet to be transmitted will be discarded. If the channel quality suddenly deteriorates, the packet may also be lost. For the system, the lower the packet loss rate, the better the performance.

4. Simulation Results

4.1. Fairness

![Bar chart comparing algorithm fairness with different URLLC users](image)

**Figure 1.** Comparison of algorithm fairness with different URLLC users

As shown in Figure 1, the fairness among EMBB users using IPS algorithm is significantly higher than that of PS algorithm. Due to the introduction of weighting factor in IPS algorithm, for EMBB users with poor channel quality, the priority of re-punching decreases with the increase of the number of holes in the past period of time, thus improving the fairness among users.

4.2. Throughput Comparison

![Line chart comparing cell throughput](image)

**Figure 2.** Cell throughput comparison

As shown in Figure 2, the total throughput of the IPS algorithm is about 4MB /s lower than that of
the PS algorithm. This is because the weighting coefficient is introduced into the IPS algorithm, and under the user channel condition, the weighting coefficient is slightly lower than the priority of the punching data in the PS algorithm.

4.3. Comparison of Packet Loss Rates

Table 2. Comparison of packet loss rates

|       | 4   | 8   | 12  | 16  | 20  |
|-------|-----|-----|-----|-----|-----|
| IPS   | 1.3 | 1.9 | 2.6 | 4.1 | 7.0 |
| PS    | 1.7 | 2.8 | 4.5 | 7.8 | 12  |

As shown in Table 2, under the same simulation conditions, the change of packet loss rate of EMBB service under IPS algorithm and PS algorithm with the change of the number of users is given. It can be seen from the table that compared with PS algorithm, IPS algorithm has a smaller packet loss rate. This is because the IPS algorithm preference in a real-time business conducted on the resources of stiletto eMBB occupies, avoid playing in May is almost close to the time delay threshold eMBB real-time data transmission business resources, reduces its been URLLC transmission of business data and choose the priority punching, and thus to a certain extent, reduce the probability of decoding failure and retransmission again, so reduces the eMBB real-time packet loss rate of the business.

5. Conclusions

Low delay and high reliability communication application is one of the three main business scenarios of 5G mobile communication system, and it is also an essential part of the realization of emerging real-time reliable applications. Based on the real-time and reliable demand of wireless communication in the future, this paper comprehensively applies the relevant knowledge of wireless communication theory, information theory, cybernetics and convex optimization theory to study the key technologies of wireless transmission and cooperative control with low delay and high reliability. The algorithm proposed in this paper takes the number of holes punched by EMBB users by URLLC users as an important parameter in the calculation of user priority, which enables users with fewer holes to obtain high priority, thus improving the fairness among users. And the algorithm adopts different punching strategies according to the different EMBB services in the cell, which reduces the packet loss rate to a certain extent.

References

[1] Wu X, Jiang M, Zhao C, et al. Low-Rate PBRL-LDPC Codes for URLLC in 5G. IEEE Wireless Communications Letters, 2018, 7(5):800-803.
[2] Nielsen J J, Liu R, Popovski P. Ultra-Reliable Low Latency Communication (URLLC) using Interface Diversity. IEEE Transactions on Communications, 2017, 66(3):1322-1334.
[3] Ryu W J, Shin S Y. Performance evaluation of a power allocation algorithm based on dynamic blocklength estimation for URLLC in the multicarrier downlink NOMA systems. Turkish Journal of Electrical Engineering and Computer Sciences, 2021, 29(1):310-320.
[4] Haija A A A, Tellambura C. Decoding Delay and Outage Performance Analysis of Full-Duplex Decode-Forward Relaying: Backward or Sliding Window Decoding. IEEE Transactions on Communications, 2016, 64(11):4520-4533.
[5] Beckmann K, Manem H, Cady N C. Performance Enhancement of a Time-Delay PUF Design by Utilizing Integrated Nanoscale ReRAM Devices. IEEE Transactions on Emerging Topics in Computing, 2017, 5(3):304-316.
[6] Rao J, Vrzic S. Packet Duplication for URLLC in 5G: Architectural Enhancements and Performance Analysis. IEEE Network, 2018, 32(2):32-40.
[7] Haider A, Rahman M, Khan T, et al. Multi-Gigabit CO-OFDM System over SMF and MMF Links for 5G URLLC Backhaul Network. Cmc 'Tech Science Press', 2021, 67(2):1747-1758.
[8] Porretta M, Schlarmann B K, Ballereau A, et al. A Novel Uplink Scheduling Algorithm for the
Galileo System. IEEE Transactions on Aerospace and Electronic Systems, 2018, 54(2):819-833.

[9] Elkader A A. Tuples: A New Scheduling Algorithm. Journal of Computers, 2018, 13(11):1309-1315.

[10] M, Sudhakar. Wmns: Efficient Scheduling Algorithm and Classification of Routers With Multiple Gateways. Advances in electronic and electric engineering, 2016, 7(2):119-134.