

QoS-Oriented Scheduling Algorithm For Mobile Multimedia In OFDM

Liang Xiao, Shidong Zhou, Yan Yao

State Key Lab on Microwave and Digital Communications
Tsinghua University, Beijing 100084, China. Tel: 8610-6278-1398
E-mail: xiaoliang00@mails.tsinghua.edu.cn

Abstract: In this paper, we propose and evaluate a new algorithm to schedule multimedia data in a multi-user OFDM wireless system. This algorithm dynamically allocates the resource in both time and frequency domains based on feedback channel states, packet QoS requirements and the amount of data to be transmitted. Accordingly, maximum delay, BER requirement and priority of the packet are well met with improved system throughput, even with harsh channel conditions. Moreover, several neighboring subcarriers may be grouped together, while possibly incorporating with an adaptive coding & modulation, to deliver multiple data rates depending on actual channel state information. Performance assessment has been conducted per a channel model with Doppler shift, multi-path fading and path-loss. It is shown that the proposed algorithm is quite effective to guarantee QoS for high rate multimedia transmission in OFDM systems.

I. INTRODUCTION

The demand on high rate multimedia information has pushed the development of wireless communication systems in an unprecedented pace. System design of fourth-generation (4G) technology has already begun, which seeks to support a wide range of packet services with highest data rates reaching 20Mbps. With limited bandwidth resources, efficient scheduling algorithm has to be put up to provide satisfactory services to the users.

Multimedia information sources not only exhibit highly bursty traffic rates, but also have variable quality of service (QoS) requirement and traffic rate characteristics. It is especially challenging to develop multimedia packet scheduling algorithm in the wireless systems because of the limited bandwidth and time-varying link capacity. Many existing wireless scheduling algorithms, such as idealized wireless fair queueing (IWFQ), channel-condition independent packet fair queueing (CIF-Q), server-based fair approach (SBFA), are all based on the assumption of two-state Markov channel [1]. These algorithms assume the wireless link capacity has two states: the full capacity in error-free state and the zero capacity in error state. But the real wireless transmission has multiple states link and different physical capacity. The two state assumptions damage the system efficiency.

Considering this, the CDMA/HDR (High Data Rates) introduced a *dynamic data rate* transmission technique

allowing efficient data transmission with variable data rate according to the channel condition of each Mobility Terminal (MT). Accordingly, its scheduling adopts the proportional fairness algorithm (PF) that maximizes channel throughput with fairness constraint [2].

In the traditional wireless system, the QoS is guaranteed by the access admission control, the resource reservation protocols and etc. However, some QoS-based scheduling algorithms for HDR have been put forward recently, including the maximum rate rule (MAX-RATE), the modified largest weighted delay first rule (PROP-FAIR), the exponential rule (EXP) etc [3]. They improve the performance of real-time data and lessen the burden of the access admission process.

OFDM is an important candidate technique for the future wireless transmission. Its multi-carrier nature makes its scheduler different from that of CDMA/HDR, which allocates resource in only time domain. Most current OFDM scheduling research doesn't take QoS into consideration and too complex to be fitted in the real system [4][5].

In this paper, we propose a QoS-based dynamic resource allocation scheme for mobile multimedia service in OFDM. The scheduling is performed in both time domain and frequency domain. It aims at close-loop system and to reduce feedback and complexity, several neighboring subcarriers forming a subband are allocated as a units at each slot. Corresponding to each packet preparing to be sent, the scheduling algorithm generates a token containing the packet's ID, length and QoS information (Maximum delay, class, etc) at the end of the queue for that service in the token pool. Variable transmission rate service can be supported since there is no restriction on the packet character for one user. The services are scheduled with a priority order determined by their current transmit capacity, QoS requirements and average transmission rates.

This paper is organized as follows. Section II builds up the systems model. Section III presents our algorithm and its implementation. Section IV analyzes the simulation results. And Section V concludes the paper.

II. SYSTEM MODEL

Orthogonal Frequency Division Multiplexing (OFDM) is an important candidate technique for 4G. Combating the inter-symbol interference, OFDM has high spectrum efficiency and supports high-rate data transmission. In this paper, we suppose a 1024 subcarriers OFDM system. We divide each 10ms-length frame into 10 slots and suppose one slot contains eight symbols. In order to reduce the feedback and algorithm complexity, we propose the “subband” concept, which is made up of several neighboring subcarriers. Each subcarrier in the same subband is coded and modulated with the same scheme. The scheduling decisions are made once every “time-slot”.

To improve the system efficiency, many OFDM systems recently adopt the Adaptive Modulation Coding (AMC). This technique utilizes the feedback channel information and supply variable transmission rate. Table 1 gives an eight-state transmission scheme in AMC.

In this paper, we consider the single cell multiuser downlink OFDM system, which contains one Access point (AP) and several Mobility Terminals. With maximum transmission power, the scheduling mainly focuses on the data rate control.

In the close-loop system, each MT gets the signal noise rate (SNR) on each subband based on the downlink pilot information, and thus obtains the Data Rate Control information (DRC) in each subband. In each slot, all the MTs feedback the DRC information to AP.

DRC	SNR (dB)	Code Rate	Modulate	Capacity bits/slot)
0	/	/	/	0
1	-3.4	1/4	BPSK	128
2	-0.4	1/2	BPSK	256
3	2.2	1/2	QPSK	512
4	5.2	3/4	QPSK	768
5	7.6	2/3	8PSK	1024
6	10.9	3/4	16QAM	1536
7	14.5	2/3	64QAM	2048

Table 1. Adaptive Transmission Scheme for OFDM

AP sets a token pool for the real-time data shown in figure 1, which generates and maintains a token queue for each admitted MT. For each real-time packet from the higher layer, the scheduling algorithm sets a token containing its ID, length, maximum delay and other information in the corresponding queue. This method supports variable data rate, because the token queue screens the restriction on the packet character of. Similarly, if the token pool space is large enough, the packet arrival character will not affect system performance.

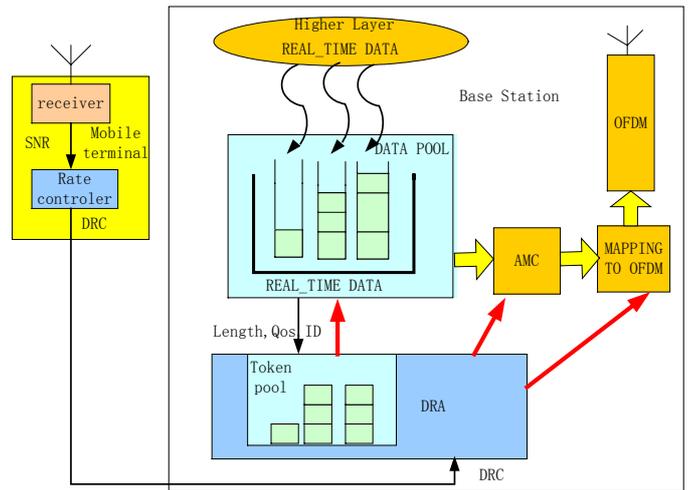


Figure 1. System architecture of the transmit side.

III. ALGORITHM DESCRIPTION

The scheduling process in figure 2 is carried out in each slot for the real-time data. We set a priority-evaluation-value for each real-time data service token queue, which can be get from the following equation

$$Prio_eval_i = \frac{\sum_j^M Capa_{i,j}}{R_i \cdot M} + \beta \frac{\sum_j^{num_token_i} \alpha_{i,j} \cdot Len_{i,j}}{Delay_{i,j}} \quad (1)$$

Here, β is a tunable parameter to give more weight to the second factor in the equation. $\alpha_{i,j}$ is the priority coefficient for the j th packet in the i th user queue. It is determined by the higher layer based on the packet's QoS requirement such as BER, jitter and packet importance. $Len_{i,j}$ and $Delay_{i,j}$ are respectively the packet length and the time before maximum delay for the j th packet in the i th user queue. Their values are firstly set by the higher layer as the packet length and the maximum delay, and then are updated and maintained by the scheduling algorithm. num_token_i is the number of the packets in the corresponding service's token queue. M is the number of subbands in OFDM. $Capa_{i,j}$ is the estimate link capacity in a slot and is derived from the DRC for the j -th subband of the i -th user. R_i is the long-time average transmission rate shown in equation (4).

In each slot, the real-time data are scheduled in a turn with descending priority-evaluation-value. One service queue is scheduled at a time. And after the packet queue with the lowest priority-evaluation-value get scheduling right, a new scheduling turn will begin from the service with the highest priority-evaluation-value, unless the overall scheduling turn in this slot exceeds the maximum turn parameter.

From equation (1) we can see the priority is determined by two factors. Firstly, the users with heavier transmission load and poorer transmit capacity obtain a higher priority.

Secondly, the algorithm also puts some priority to the user whose instant transmit capacity most reaches its peak capacity in a long time. The first factor has more weight in the scheduling turn.

When a service queue gets the scheduling right, it will be allocated in the following way. Firstly, it selects as many best available subbands as it can before its allocated resource in this scheduling turn meets the capacity goal, which can be shown as

$$Goal_i = \sum_j^{num_token_i} \frac{Len_{i,j}}{Delay_{i,j}}, i = 1, 2, \dots, N_{real} \quad (2)$$

Secondly, the scheduling right must be released if the subband number allocated in this scheduling turn exceeds

$$Max_subband_i = M \cdot Goal_i / \sum_k Goal_k, i = 1, 2, \dots, N_{real} \quad (3)$$

Thus we can get the subbands, as well as the transmission scheme and capacity of the user in this turn.

After the scheduling turn exceeds the maximum turn parameter, the final transmission scheme of each subband is decided by the feedback parameter DRC of the same user at this subband. And the parameters such as Len and Delay in the tokens should be updated. The parameter R of the real-time data is updated with

$$R_i = (1 - 1/T_c) \cdot R_i + Len_i / T_c \quad (4)$$

where, T_c is similar to that in [2] and Len_i is the overall length of the transmission resource allocated to this user at this slot.

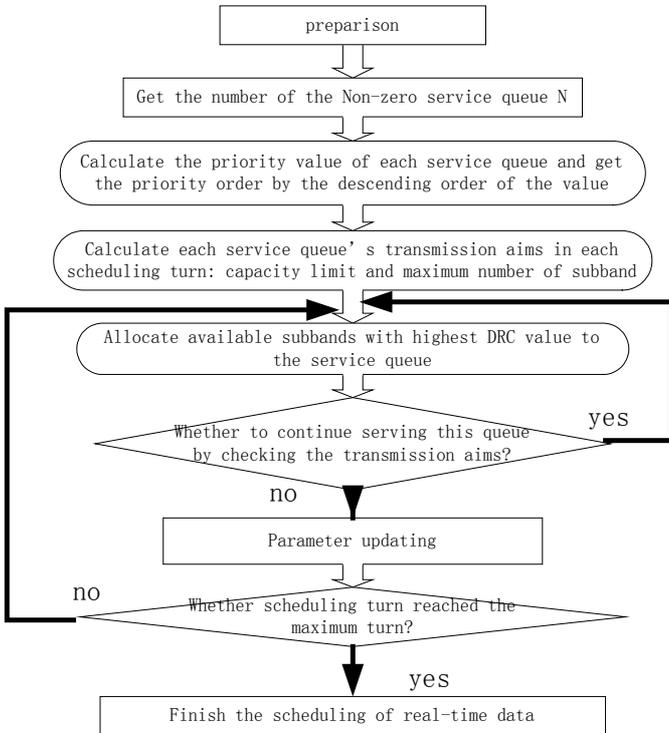


Figure 2. Scheduling flowchart for real time data.

IV. SIMULATION RESULTS

The algorithm performance with seven users is analyzed by Synopsys's CCSS. In the simulation, the 1024 OFDM sub-carriers are divided into 16 subbands. The carrier frequency is 2GHz and the bandwidth is 10MHz. The sampling rate of receive signal is 20mps.

Tap#	Delay [us]	Power [dB]
1	0	0
2	0.31	-1
3	0.71	-9
4	1.09	-10
5	1.73	-15
6	2.51	-20

Table 2. Channel Parameters

The channel model considers the large-scale fading, Doppler fading and multi-path fading, adopting the six paths model in table 2. Each MT has the same average vehicle rate 20mps and the average receiving SNR ranging from 6dB to 12dB. Each MT transmits the real-time data with QoS characters listed in the following table. $\alpha_{i,j}$ is the priority coefficient, determined by the higher layer considering the packet's BER and jitter requirement.

MT ID	Rate (bps)	Maximum Delay (10ms)	$\alpha_{i,j}$	Packet Length (Bit)
1	10M	15	30	100K
2	800K	20	15	8K
3	800K	20	15	8K
4	800K	20	15	8K
5	500K	10	30	5K
6	200K	15	10	2K
7	100K	30	10	1K

Table 2. QOS Character of real-time data

Figure 3 presents the average throughput of the real-time data in seven MTs. Figure 4 gives the ratio of the real-time throughput to the required data rate. The packet outage rate is shown in Figure 5.

From the figures, we can see the average throughput of the real-time data rises with SNR before satisfying the QoS requirement. After the SNR exceeds 11dB, the throughput requirement of the all the users can be guaranteed. Before all the data can be transmitted, the throughput of each user is related to its priority coefficient.

When the channel condition is worse, the QoS requirement of the data with higher QoS priority coefficient is more likely to meet. For example, the data for user 5 has the

strictest delay requirement, but its satisfied level is still very high due to a high priority coefficient.

Meanwhile, the users with the same QoS requirement, such as user 2,3 and 4, have the same average performance on throughput and delay. These show the fairness of the algorithm.

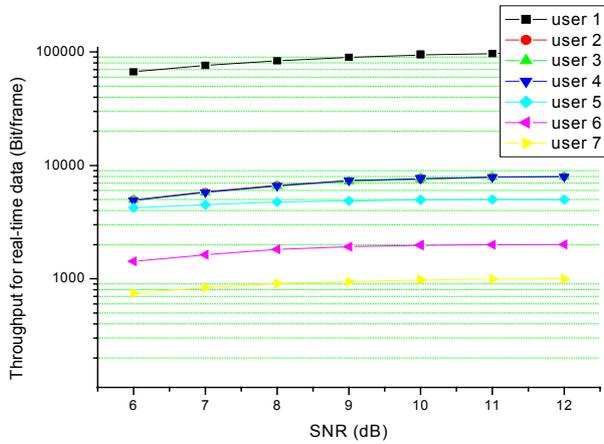


Figure 3. Average throughput of real-time data

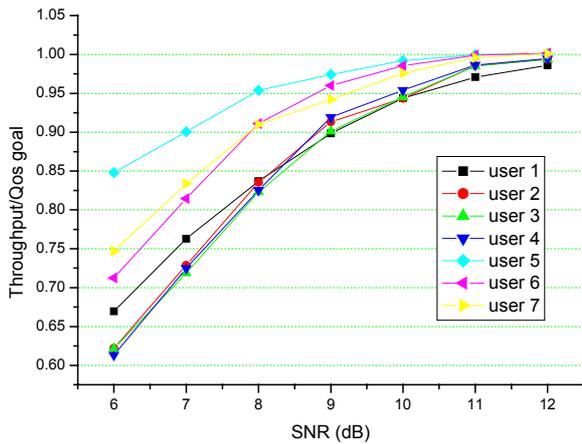


Figure 4. Ratio of the throughput to the required data rate

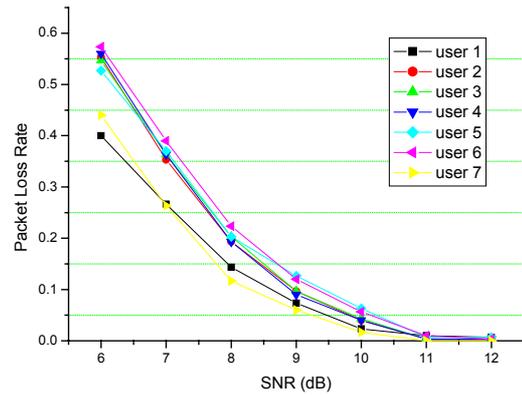


Figure 5. Packet loss rate of real-time data

V. CONCLUSION

In this paper, we proposed a QoS-orientated scheduling algorithm for the mobile multimedia service in the OFDM system. The algorithm give each service for the same MT a priority-evaluation-value and schedule the service in a turn with descending priority-evaluation-value. The number of the subbands allocated to each user in each turn by can be obtained with the two allocation rules. The final transmission scheme of each subband is decided by the feedback parameter DRC of the corresponding user. Simulation results shown that the performance of each user, such as average throughput and average packet outage rate is based on its QoS requirement.

Reference

- [1] Yaxin Cao, Victor O.K.LI, "Scheduling Algorithms in Broad-Band Wireless Networks", *IEEE Proceedings of The IEEE*, VOL.89, NO.1, 2001.
- [2] A. Jalali, R. Padovani, R. Pankaj, Data Throughput of CDMA-HDR a High Efficiency-High Data Rate, *VTC 2000-Spring*, pp.1854-51858 vol.3.
- [3] Sanjay Shakkottai, "Scheduling algorithms for a mixture of real- time and non-real-time data in HDR".
- [4] C. Y. Wong, R. S. Cheng, K. B. Letaief, R. D. Murch, 'Multiuser OFDM with Adaptive Subcarrier, Bit and Power Allocation', *IEEE J. on SAC*, vol.17, No.10, 1999.
- [5] Inhyoung Kim, H. L. Lee, et al. "On the Use of Linear Programming for Dynamic Subchannel and Bit Allocation in Multiuser OFDM", *IEEE Global Telecommunications Conference*, v 6, 2001, p 3648-3652.