A novel cross-layer scheduling algorithm for IEEE 802.16 WMAN

Chong Tian , Dongfeng Yuan
School of Information Science and Engineering, Shandong University, Jinan, 250100, P.R.China
e-mail: *{tianchong,dfyuan}@sdu.edu.cn*

***Abstract*—In this paper, we propose a novel cross layer scheduling strategy for IEEE 802.16 WMAN. The proposed scheduler will first guarantee all the prescribed QoS requirements, then take into account different wireless channel conditions obtained by subscriber stations to enhance the system throughput while it ensures the property of the fairness per connection. The performance of our scheduler is evaluated via simulations. The simulation results show that our proposed scheduler can provide diverse QoS guarantees, use the wireless bandwidth efficiently and achieve an optimum tradeoff between throughput and fairness.**

***Keywords*—cross-layer, scheduling algorithm, WMAN, QoS, fairness.**

I. INTRODUCTION

In recent years, wireless technologies have significantly improved. Wireless infrastructure has been deployed to be the complement and expansion of the wired infrastructure. IEEE 802.16 WMAN refers to a broadband wireless access (BWA) technology that promises to deliver up to 70 Mbit/s for each user at a non-line-of-sight range of up to 50 km [1]. One of the main advantages of WMAN is its ability to support a wide range of data transmission, with different traffic characteristics and quality of service (QoS) requirements.

In QoS Mechanisms, Scheduling plays an important role in QoS provision because it (1) Selects a packet for transmission from the packets waiting in the transmission queue.(2)Decides which packet from what queue and station are scheduled for transmission in a certain period of time.(3)Controls bandwidth allocation to stations, calsses and applications. It is crucial to the development of future broadband wireless networks because it can support QoS differentiation and guarantees for them. Many BWA systems define only QoS architecture and signaling but do not specify the scheduling algorithm which will ultimately provide QoS support.

In this work, we apply a cross-layer design approach to design a scheduling algorithm which considers multiuser diversity. While ensuring all the QoS requirements, the scheduler will try to provide fairness for all connections.

II. SYSTEM MODEL

The IEEE 802.16 standards prescribe the PHY layer and MAC layer profile of WMAN in detail. In this paper, we will develop a cross layer model of adaptive wireless links for QoS support in WMAN.

A. Physical Layer

Each user's packet data are finally transmitted through the time varying fading channel. At the physical layer, we consider the group of transmission modes in the IEEE 802.16 standard [2] as in Table 1. Using AMC schemes that adjust transmission parameters to the wireless channel variations adaptively can realize efficient bandwidth utilization for a prescribed error performance.

As in [3], let N denote the total number of transmission modes available. The design objective of AMC is to determine the SNR boundaries $\{\gamma_n\}_{n=0}^{N+1}$ to partition the entire SNR in N+1 non-overlapping consecutive intervals. The related expressions are as follow:

$$\begin{aligned} &\gamma_0 = 0, \\ &\gamma_n = \frac{1}{g_n}\ln\left(\frac{a_n}{P_0}\right), n = 1,2,\dots, N, \\ &\gamma_{N+1} = +\infty \end{aligned} \tag{1}$$

and,

Mode n is chosen, when $\gamma \in [\gamma_n, \gamma_{n+1})$ for n=1, …, N. (2)

Note that no data will be sent when $\gamma \in [\gamma_0, \gamma_1)$ in order to avoid deep channel fades.

By employing AMC design, we can maximize the data rate by matching transmission parameters to channel conditions, while maintaining a prescribed packet error rate P_0.

1.Supported by National Scientific Foundation of China (No. 60672036)
2.Supported by Key Project of Provincial Scientific Foundation of Shandong (Z2006G04)

Mode n	1	2	3	4	5	6
Modulation	QPSK	QPSK	16QAM	16QAM	64QAM	64QAM
RS Code	(32,24,4)	(40,36,2)	(64,48,8)	(80,72,4)	(108,96,6)	(102,108,6)
CC Code Rate	2/3	5/6	2/3	5/6	3/4	5/6
Coding Rate R_c	1/2	3/4	1/2	3/4	2/3	3/4
R_n (bits/symbol)	1.0	1.5	2.0	3.0	4.0	4.5
a_n (dB)	232.9242	140.7922	264.0330	208.5741	216.8218	220.7515
g_n	22.7925	8.2425	6.5750	2.7885	1.0675	0.8125
γ_{pn}(dB)	3.7164	5.9474	9.6598	12.3610	16.6996	17.9629

Table 1. Transmission modes in IEEE 802.16 standard

B. Medium Access Control(MAC) Layer

At the MAC, each connection is mapped to a scheduling service and is associated with a set of rules that quantify the aspects of its behavior. Four scheduling services are provided by the MAC as in the IEEE 802.16 standard:

1. Unsolicited grant service (UGS) is tailored for carrying services that generate fixed units of data periodically such as E1/T1 and VoIP without silence suppression. This service provides guarantees on throughput, latency, and jitter to the necessary levels as TDM services.

2. Real-time polling service (rtPS) is well suited for connections carrying services such as MPEG or streaming video or audio. It provides guarantees on throughput and latency, and the QoS metrics are the PER and the maximum delay.

3. Non-real-time polling service (nrtPS) that provides guarantees on throughput only can tolerate longer delays and are rather insensitive to delay jitter. So it is suitable for FTP applications. The PER and the minimum reserved rate will be the QoS metrics.

4. Best Effort (BE) service provides neither throughput nor delay guarantees (e.g., HTTP, Email). The BE applications receive the residual bandwidth after the requirements of the previous three service classes are satisfied. A prescribed PER should be guaranteed for BE connections over wireless channels although no QoS parameter is specified.

The QoS architecture and signaling of each connection are defined as in the IEEE 802.16 standard [2]. However, the standard left the QoS based packet scheduling algorithms, that determine the uplink and downlink bandwidth allocation, undefined.

III. SCHEDULER DESIGN

We here propose a cross layer scheduler for multiple connections with diverse QoS requirements.

In our scheduler, the UGS connections will be given higher priority than the other three QoS classes (rtPS, nrtPS, BE) and our algorithm will only consider scheduling for rtPS, nrtPS and BE connections. Each connection under consideration adopts AMC at the PHY. Given the prescribed PER as ξ_i, we can determine the SNR thresholds for connection i by setting $P_0 = \xi$, where i denotes the connection identification. As in [4], we assume that a PHY frame has N_d time slots to convey data, and a time slot can transmit $R_i(t) \in \{2R_n\}_{n=0}^{N}$ packets with transmission mode n which is determined by the channel quality of connection i via AMC as in (2). The time slots allocated to UGS connections denoted as N_{UGS} are fixed per frame, so the time slots that can be scheduled for the other three QoS classes can be expressed as N_r=N_d-N_{UGS}.

Our scheduler is expected to achieve two goals: 1 satisfy all the QoS requirements. 2 make the optimum tradeoff between throughput and fairness. We will realize our scheduler also by two steps.

Step 1: Define $S_i(t+1)$ as the QoS guarantee function which represents the QoS guarantee level at time t+1 if the packets of connection i will not be scheduled at time t.

For a rtPS connection, it is defined as:

$$S_i(t+1) = T_i - W_i(t+1) + 1, \tag{3}$$

where Ti is the maximum latency (deadline), and Wi(t) is the packet waiting time:

$$W_i(t+1) = W_i(t) + T_f. \tag{4}$$

For each nrtPS connection, if data of connection i are always available in queue, the average transmission rate at time t usually estimated over a window size t_c as:

$$\hat{T}_i(t+1) = \begin{cases} \hat{T}_i(t)(1-1/t_c), & \text{if } i \neq i^*. \\ \hat{T}_i(t)(1-1/t_c) + N_r * R_i(t)/t_c, & \text{if } i = i^*. \end{cases} \tag{5}$$

We would like to guarantee $\hat{T}_i(t+1) \geq T_i$ during the whole service period. We express its QoS guarantee function as:

$$S_i(t+1) = \hat{T}_i(t)(1-1/t_c)/T_i, \tag{6}$$

If $S_i(t+1) \geq 1$, the QoS requirement is satisfied. If $S_i(t+1) < 1$, the packets of connection i should be sent immediately, except for $R_i(t)$=0, in order to avoid QoS dissatisfaction at time t+1, so that the highest value of priority is set. If multiple connections have their $S_i(t+1) < 1$ at the same time, the rtPS connection will be satisfied prior to nrtPS connection due to the higher QoS-class priority.

The BE connections are not involved in this step because there are no QoS guarantees for BE service.

Step 2: If all $S_i(t+1)$ are no less than 1, we will make the tradeoff between channel conditions and fairness for all the connections. The scheduler will allocate the time slots per frame to the connection $i^* = \arg\max_i \phi_i(t)$, where $\phi_i(t)$ is the priority function for connection i at time t, which can be specified as:

$$\phi_i(t) = \frac{R_i(t)}{f_i(t) + \varepsilon}, \tag{7}$$

where $f_i(t)$ is the fairness indicator, which is defined as:

$$f_i(t+1) = \frac{1}{t}\sum_{l=0}^{t} R_i(l) I_i(l). \tag{8}$$

with $I_i(t)=1$ if queue i is picked for service in time t; otherwise, $I_i(t)=0$.

In summary, this design can result in good QoS guarantees and fairness. For example, when QoS will be not satisfied for a connection at next time, our design is to assign the highest priority to such a connection to avoid QoS decreasing. So, it can provide delay bound for rtPS connections and guarantee throughput for nrtPS connections. Besides, our design can also provide comparable priorities among connections with different kinds of services, which enables exploiting multiuser diversity, to make optimum trade-off between fairness and throughput maximization. The equity function embedded into the scheduler can be formulated to prevent queue starvation and thus preserve fairness among competing queues.

IV. SIMULATIONS

A. *Channel model*

In our model, the wireless channel quality is characterized by the instantaneous SNR γ, which remains invariant during a frame. We adopt the general Nakagami-m mode to describe γ statistically. The probability density function of the received SNR per frame is:

$$p_\gamma(\gamma) = \frac{m^m \gamma^{m-1}}{\bar{\gamma}^m \Gamma(m)} \exp(-\frac{m\gamma}{\bar{\gamma}}), \tag{9}$$

Where $\bar{\gamma}$ is the average received SNR, $\Gamma(m)$ is the Gamma function and m is the Nakagami fading parameter ($m \geq 1/2$).

In order to describe small-scale channel variations, we adapt the two-state discrete time Markov chain channel variation mode presented in [5]. Assuming slow fading conditions so that transitions happen only between adjacent transmission modes, the probability of transition exceeding two consecutive modes is zero.

B. *Parameter setting*

The frame length is T_f=1ms. The packet length at the MAC is fixed as N_b=128bytes.

We assume that the arrival process to the queue for each rtPS connection is Bernoulli distributed with a given average rate T_i and parameter p_i [6]. As a result, the instantaneous arriving rate at time t can be expressed as:

$$A_i(t) = \begin{cases} 0, & \text{with probability } \mathrm{p_i}, \\ T_i/(1-p_i), & \text{with probability } 1 - \mathrm{p_i}. \end{cases} \tag{10}$$

For each nrtPS and BE connection i, we assume that the data are always available, which is reasonable for, e.g., FTP HTTP or Email applications.

We consider two rtPS connections, two nrtPS connections and two BE connections are admitted in the system with i=1,2,3,4,5,6 respectively. Their channel, QoS and traffic parameters are:

$\bar{\gamma}_1$ =15 (dB), m_1=1.2, ξ_1 =10^{-2}, D_1=30(ms), T_1=2 (Mbps), p_1=0.4;

$\bar{\gamma}_2$ =20 (dB), m_2=1, ξ_1 =10^{-2}, D_2=50(ms), T_2=1 (Mbps), p_2=0.5;

$\bar{\gamma}_3$ =15 (dB), m_3=1, ξ_3 =10^{-3}, T_3=6 (Mbps);

$\bar{\gamma}_4$ =20 (dB), m_4=1, ξ_4 =10^{-3}, T_4=3 (Mbps);

$\bar{\gamma}_5$ =16 (dB), m_5=1, ξ_5 =10^{-3};

$\bar{\gamma}_6$ =18 (dB), m_6=1, ξ_6 =10^{-3};

C. *Performance Evaluation*

1) Prescribed QoS Guarantees: The delay performance of rtPS connections is evaluated by the delay outage probability $D_i(t)$ over a window size t_c=1000ms as:

$$D_i(t+1) = \begin{cases} D_i(t)(1-1/t_c), & \text{if } W_i(t) < D_i, \\ D_i(t)(1-1/t_c) + 1/t_c, & \text{if } W_i(t) = T_i \, and \, \mathrm{i} \neq \mathrm{i}^*. \end{cases} \tag{11}$$

The rate performance of nrtPS connections is evaluated by the average service rate $T_i(t)$ over a window size t_c=1000ms based on (5).

2) Optimum tradeoff between fairness and throughput: The performance metrics here are per-connection fairness. We define the fairness index according to [7]:

$$F(t) = \frac{(\sum_{i=1}^{6} f_i(t))^2}{6 * \sum_{i=1}^{6} f_i^2(t)}. \tag{12}$$

The fairness index takes values between 0 and 1. The closer the index is to 1, the fairer the scheduling is.

The system was simulated over 60,000ms with N_r=4 and 3, respectively.

The delay outage probability $D_i(t)$ of rtPS connections i=1,2, the average transmission rate $\hat{T}_i(t)$ of nrtPS connections i=3,4 and the fairness evaluation for all the connections are plotted in Figs. 1,2 for N_r=4 and 3, respectively.

Fig. 1 depicts the performance for N_r=4, and shows that the delay outage probability $D_1(t)$ and $D_2(t)$ are always below 1%, which illustrate good delay performance for rtPS connections. Notice that $\hat{T}_3(t)$ and $\hat{T}_4(t)$ are greater than the minimum reserved rates T_3=6 and T_4=3 almost in the whole period, which indicate good rate guarantees for nrtPS connections. Furthermore, the fairness index is close to the value 1, which means good fairness per connections.

Fig. 2 illustrates the performance for N_r=3. Here one time slot is reduced from N_r=4, which may be interpreted as new UGS connections admitted. We find that the delay performance for rtPS connections is still good. Notice that $\hat{T}_3(t)$ and $\hat{T}_4(t)$ for nrtPS connections vary around the minimum reserved rates, but the variations are very small, so the rate performance is also guaranteed; while the fairness index is much smaller than that for N_r=4, due to insufficient available bandwidth that can be used in step 2 to achieve per-connection fairness, which induce that the rate performance of BE connections is much worse for N_r=3 than that for N_r=4.

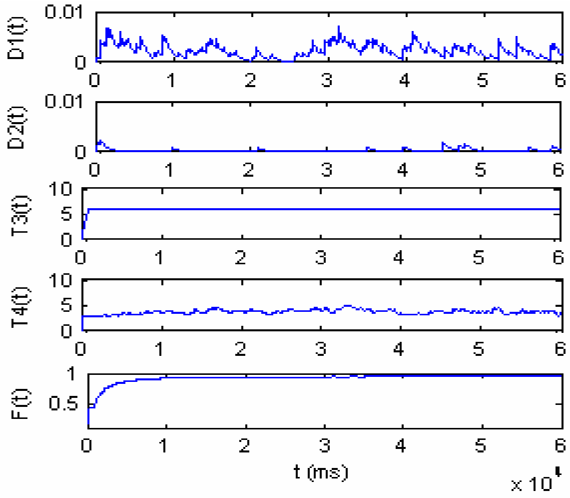

Figure 1. $D_1(t)$, $D_2(t)$, $\hat{T}_3(t)$, $\hat{T}_4(t)$,F(t) VS. t for N_r=4

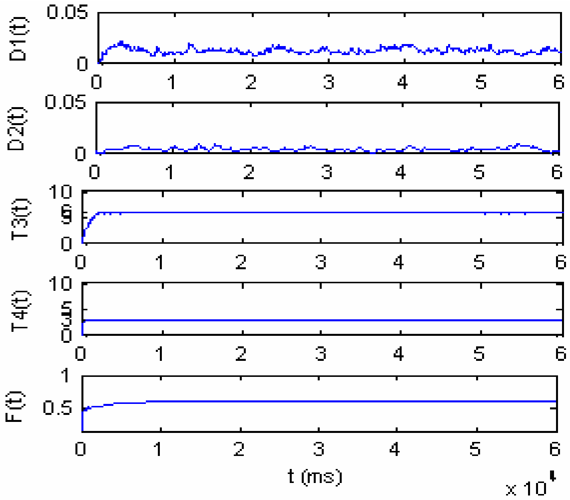

Figure 2. $D_1(t)$, $D_2(t)$, $\hat{T}_3(t)$, $\hat{T}_4(t)$,F(t) VS. t for N_r=3

V. CONCLUSIONS

In this paper, we propose a novel cross-layer scheduling algorithm at the IEEE 802.16 WMAN MAC layer for multiple connections with diverse QoS requirements, where each connection employs AMC scheme at the PHY layer. Based on the specified QoS parameters, our scheduler will first satisfy all the QoS requirements; at the same time, it uses the wireless bandwidth efficiently by exploiting multiuser diversity among connections with different kinds of services and finally make an optimum trade off between throughput and fairness. Simulation results show that our scheduler can guarantee connections' QoS demands, avoid starvation of low-priority service class, and achieve better fairness degree.

REFERENCES

[1] IEEE P802.16-REVd/D5-2004: "Air Interface For fixed Broadband Wireless Access Systems"

[2] IEEE Standard 802.16 Working Group, IEEE standard for local and metropolitan area networks part 16: air interface for .fixed broadband wireless access systems (revision of IEEE standard 802.16-2001), 2004

[3] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer modeling of adaptive wireless links for QoS support in multimedia networks" , in Proc. of 1st Intl. Conf. on QoS in Heterogeneous Wired/Wireless Networks (QShine), pp. 65-75, Dallas, TX, Oct. 18 - 20, 2004.

[4] Q. Liu,, X. Wang, and G. B. Giannakis, "A cross-layer scheduling algorithm with QoS support in wireless networks", IEEE Trans. On Veh. Technol., vol.. 55, no. 3, pp. 839-847, May. 2006..

[5] C. Chien, M.B. Srivastava, R. Jain, P. Lettieri, V. Aggarwal, R. Sternowski, "Adaptive radio for multimedia wireless links", IEEE J. Select. Areas Commun. 17 (1999) 793-813.

[6] S. Shakkottai and A. Stolyar, "Scheduling algorithms for a mixture of real-time and non-real-time data in HDR" , in Proc. of the 17th International Teletraf.c Congress (ITC-17), Salvador da Bahia, Brazil, Sept. 24-28, 2001.

[7] R. Jain, D. Chiu and W. Haweu, "A Quantitative Measure of Fairness and Dicrimination for Resource Allocation in Shared Computer System", Research Report, TR-301, September, 1984