Joint Packet Scheduling and Radio Resource Assignment for WiMAX Networks

NAN XU[1], WEIJIA JIA[2]

[1] *Department of Computer Science, City University of Hong Kong, Hong Kong SAR, China*
[2] *Future Networking Center, CityU ShenZhen Research Institute, City University of Hong Kong, Hong Kong SAR, China*

The IEEE 802.16 standard defines the QoS signaling framework and various types of service flows, but left the QoS based Packet Scheduling and Radio Resource Assignment undefined. This paper proposes a novel joint Packet Scheduling and Radio Resource Assignment algorithm for WiMAX Networks. Our algorithms can effectively assign the suitable slots to meet the QoS requirements of the different service type flows while taking the throughput and fairness into considerations. The effectiveness of our algorithms have been demonstrated through extensive analysis and simulation data. The results show that our algorithms greatly improve the throughput with relatively low complexity.

Key words: IEEE 802.16, Packet Scheduling, Radio Resource Assignment, QoS.

1 INTRODUCTION

WiMAX (IEEE 802.16) has recently been standardized for Broadband Wireless Access (BWA) technologies. BWA is becoming increasingly important in Wireless Metropolitan Area Networks (WMANs). Point-to-multipoint (PMP) mode in the IEEE 802.16 Media Access Control (MAC) protocol allows the Base Station (BS) to directly communicate with Subscriber or Mobile Stations (MSs). The IEEE 802.16 provides broadband access for the MSs based on OFDM technique. The IEEE 802.16 standard defines the QoS signaling framework and various types of service flows, but the actual QoS mechanisms such as Packet Scheduling and Resource Assignment Algorithms for these service flows are not mentioned. Therefore, the joint considerations of efficiency for Packet Scheduling and Radio Resource Assignment is very challenging because the Packet Scheduling for individual MSs should be considered together with resource utilization, which has not been well addressed.

This paper addresses these issues from the fundamental bases of both Packet Scheduling and Radio Resource Assignment while maximizing system throughput. To guarantee QoS requirement among five type service flows, we propose a novel hybrid Queuing Analytical Model for Packet Scheduling. In this model, there are two phases in the Packet Scheduling scheme. In phase one, depending on different QoS requirements in each sub-service-queue, we design specific packet selection algorithms for each sub-service-queue respectively. For phase two we propose the improved weighted fair queuing (WFQ) algorithm for Packet Scheduling among each sub-service-queue considering the different service priority.

The IEEE 802.16 MAC supports two types of resource grant mode which are Grant per Connection (GPC) and Grant per SS (GPSS). For GPC, resource is granted to a connection by BS individually. However, for GPSS mode, a portion of the available resource is granted to each of the MS, which is responsible for assigning resource among the corresponding connections. According to the feature of GPSS, BS is acting as the master controller of the entire system. However, it just assigns the whole resource to each MS that is successful in resource request. The resource is assigned by BS on a subset of total radio resource available to meet the users' demands, and the target of resource assignment is to maximize resource utilization. Hence, we design a combined two layered resource assignment scheme both on BS and MS and propose heuristic algorithms in the scheme to optimize system throughput based on QoS guarantee.

The rest of this paper is organized as follows: Section 2 summarizes related work. Section 3 describes the system model and the channel model. In Section 4, we discuss Queuing Analytical Model for Packet Scheduling considering the QoS constraints for different service types. Section 5 designs two heuristic resource assignment algorithms. The numerical results are presented in Section 6. We conclude the paper in Section 7.

2 RELATED WORK

Packet Scheduling and Radio Resource Assignment play an important role in providing QoS support to various wireless networks, especially for WiMAX networks. Not only there are very few literatures on jointly considering Packet Scheduling and Radio Resource Assignment for wireless networks built on the IEEE 802.16 standard [1], but also very few analytical models proposed for the system working in the GPSS mode.

For Packet Scheduling, [11] and [8] show that the hybrid algorithms (EDD along with WFQ) in a node gives better performance for real time services instead of EDD only. And [10] developed new scheduling algorithms for the IEEE 802.16d OFDMA/TDD based broadband wireless access system, in which radio resources of both time and frequency slots are dynamically shared by all users. However, it does not consider the wireless channel model but just a probability model for performance analysis. [7] considered some scheduler structures that are executable in environments of multiple traffic classes and multiple frequency channels. The scheduler has less flexibility in scheduling according to the frequency selection. In [5] a cross-layer scheduling algorithm was developed at the MAC layer for multiple connections with diverse QoS requirements. However, the scheduler does not consider scheduling multiple connections each time and the fairness issue.

For Radio Resource Assignment, a queuing-theoretic and optimization-based model was presented in [6] for radio resource management in IEEE 802.16 wireless networks. The system model uses a single carrier air-interface and GPC mode which are not the primary technology in the IEEE 802.16 Networks. In [3], it defines a flow metric that dynamically measures the extent to which a flow merits bandwidth allocation. Four types of interrelated resource allocation problems in OFDMA WMANs has been considered in [2]: dynamic subcarrier allocation, adaptive power allocation, connection admission control, and capacity planning. In [4] the FASA algorithm solved the problem of finding a suitable sub-channel and power joint allocation method for multiple users in 802.16e OFDMA/TDD cellular systems. For downlink resource management, an integrated APA-CAC downlink resource management framework has been proposed for WiMAX networks in [9].

3 SYSTEM MODEL AND CHANNEL MODEL

We consider a single BS serving multiple MSs through an TDMA/TDD access mode. MIMO (Multiple Input and Multiple Output) and OFDM tech-

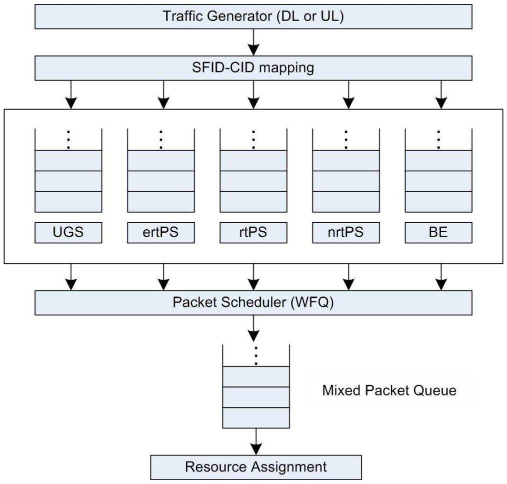

FIGURE 1
System model

niques are implemented in physical layer.

3.1 Description of System Model

The system model is illustrated in Fig. 1. The MAC layer of the IEEE 802.16 includes classifying external network packets generated from uplink and downlink and associating them to the proper MAC Service Flow Identifier (SFID) and Connection ID (CID). Then packets are inserted into different queues with an assigned CID in the MAC layer after the SFID-CID mapping. Packet Scheduler supports the appropriate packets handling mechanisms WFQ for packets transport according to each type of service, including UGS, ertPS, rtPS, nrtPS and BE. Through the Packet Scheduler, each packet from WFQ is queued into another queue called Mixed Packet Queue (MPQ). Resource assignment module retrieves packets from the Mixed Packet Queue and utilizes certain resource assignment algorithm to match them with appropriate time slots as defined by the DL-MAP or UL-MAP sent by the BS. The DL-MAP and UL-MAP are specify information about resource assignment made for each MS on downlink/uplink. Thus each MS knows when and how long to receive from and transmit data to the BS.

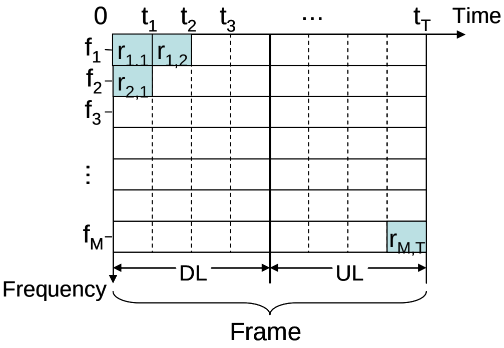

FIGURE 2
The frame structure

3.2 Channel Model

The frame structure in OFDM mode is represented in Fig. 2. There are total M sub-frequencies. In each sub-frequency, they are divided into T time slots according to TDMA techniques. Therefore, there are total $M \times T$ time slots. We mark the time slot in i-th sub-frequency and j-th slot by (i, j), for $i = 1, 2, \ldots, M; j = 1, 2, \ldots, T$.

A time slot can be treated as a channel. Let the real transmission rate in channel (i, j) be $r_{i,j}$. Different time slots have different transmission rates according to their modulation and coding schemes. Let the maximum and minimum real transmission rates of a time slot are C_{max} and C_{min} bits/s. Let $X_{i,j}(l) = k$ when $C_{min} + k(C_{max} - C_{min})/K \leq r_{i,j} < C_{min} + (k+1)(C_{max} - C_{min})/K$ ($k = 0, 1, \ldots, K-1$). Then $X_{i,j}(l)$ is the variable about the real transmission rate states of channel (i, j) in l-th frame. When $X_{i,j} = k$, the real transmission rate is uniformly distribution in the range $[C_{min} + k(C_{max} - C_{min})/K, C_{min} + (k+1)(C_{max} - C_{min})/K]$, for $k = 0, 1, \ldots, K-2$, and uniformly distribution in the range $[C_{min} + (K-1)(C_{max} - C_{min})/K, C_{max}]$, for $k = K-1$. Assume that the real transmission rate of channel (i, j) is unchanged during a time slot period. That is, the state $X_{i,j}(l)$ of the real transmission rate of channel (i, j) is stable during a time slot period. The detail modeling process is described in Appendix [Description of Channel Model].

4 QUEUING ANALYTICAL MODEL FOR PACKET SCHEDULING

The main goal of Queuing Analytical Model is to guarantee system QoS requirements. In WiMAX networks, for each packet with certain SFID and CID, like Fig. 1, it should enter and be queued in a specific sub-service-queue depending on its service type. Then it is selected and put into the head of the queue based on some specific algorithms. After that, the WFQ selects it from the queue, and insert it to the MPQ for transmission. Thus the Packet Scheduling can be divided into two phases: scheduling in sub-service-queues, and scheduling in WFQ.

4.1 Phase 1: Scheduling in Sub-service-queues

Suppose there are L_i packets in queue Q_i, where $i \in \{UGS, ertPS, rtPS, nrtPS, BE\}$. We aim to select a packet q_k^i from Q_i and insert it to the head of Q_i, therefore q_k^i can be a candidate packet and be scheduled by WFQ. Depending on the QoS type, different considerations are involved thus the selection algorithms are quite different.

UGS The UGS class does not send bandwidth requests. The BS periodically provides real-time and fixed bandwidth allocation. So in the UGS queue, no much complicated considerations are demanded. The simple FIFO (First In First Out) algorithm can meet the requirements. The packets are scheduled in the sequences of their arrivals into the queue.

ertPS/rtPS In the case of the ertPS/rtPS class, delay is considered to be one of the most important requirements, therefore the EDF (Early Deadline First) can be applied to select the packet with the highest delay priority. Let d_l be the delay requirement of the l-th packet $q_l^{ertPS/rtPS}$ in the ertPS/rtPS queue, and t_l be the time that $q_l^{ertPS/rtPS}$ has waited in $Q_{ertPS/rtPS}$. Let $\Delta t_l = d_l - t_l$. The objective of the EDF algorithm is to find the l-th packet $q_l^{ertPS/rtPS}$ with the minimum Δt_l.

nrtPS The nrtPS class is designed to support delay-tolerant streams consisting of variable-sized packets for which a minimum data rate is required. The delay requirement is not as tight as rtPS. Packets in the nrtPS queue Q_{nrtPS} can be scheduled with a WFQ algorithm, which considers fairness among all the nrtPS connections on the packet level. The scheduler keeps a record of actual transmission rate r_c for each nrtPS connection, and each connection is assigned a weight w_c according to its traffic demands and minimum reserved transmission rate, i.e.,

$$w_c = \frac{R_{c,nrtPS}^{min}}{\sum_{c \in nrtPS} R_{c,nrtPS}^{min}} \tag{1}$$

where $R_{c,nrtPS}^{min}$ denotes the minimum bandwidth transmission rate of the c-th connection of nrtPS service type. We give

$$\frac{1}{p_c} = \frac{r_c/w_c}{\sum r_c} \tag{2}$$

where p_c denotes the priority of packets belong to the c-th nrtPS connection. Each time the scheduler selects a packet with the maximal p_c from the nrtPS queue and insert it to the head of Q_{nrtPS}.

BE Various algorithms considering fairness based on service time/traffic, etc. Let r_c be the actual transmission rate of the c-th connection, t_c be the latest time that the packets belonging to c-th connection was scheduled recently. The priority of the c-th connection's packet p_c can be formulated as

$$p_c = \lambda \cdot t_c + \frac{1}{1 + r_c \cdot v} \tag{3}$$

(1) Fairness of Time. If $\lambda = 1$ and $v = 0$, equation 3 denotes that the scheduler selects the connection being scheduled least recently. Statistically, it makes all connections be scheduled with the same scheduling counts therefore achieves fairness among different BE connections. (2) Fairness of Transmission Rate. If $\lambda = 0$ and $v = 1$, equation 3 indicates that the connection with the smallest traffic rate will be scheduled first. Statistically, it keeps all the connections transmitting with the same transmission rate.

4.2 Phase 2: Scheduling in WFQ

We use WFQ algorithms to illustrate the Packet Scheduling. Base on the different QoS requirement and the delay requirement of the five sub-service-queues. We give weight $p_i = w_i, i \in \{UGS, ertPS, rtPS, nrtPS, BE\}; w_{UGS} > w_{ertPS} > w_{rtPS} > w_{nrtPS} > w_{BE})$ to each type of sub-service-queue and $\sum_{i\in\{UGS,ertPS,rtPS,nrtPS,BE\}} w_i = 1$.

There are two QoS parameters in WFQ algorithm, which are inherent weight of each type of sub-service-queue w_i and wait time of packet t_i. Let d_i be the delay requirement of UGS, ertPS, rtPS, nrtPS and BE. Then we can give the following expression.

$$p_i = \frac{e^{1/(d_i - t_i)} \times w_i}{\sum_{i\in\{UGS,ertPS,rtPS,nrtPS,BE\}} e^{1/(d_i - t_i)} \times w_i} \tag{4}$$

It represents the adaptive weight of the five sub-service-queues. When $d_i \rightarrow t_i$, the delay will become the main factor in this algorithm. Otherwise, w_i is the main factor. Then the system is more fair.

If pick up packet by algorithm at time t, arrival packet of MPQ is UGS, ertPS, rtPS, nrtPS, BE with probability

$$P_i = \frac{p_i(t)}{\sum_{i \in \{UGS, ertPS, rtPS, nrtPS, BE\}} p_i(t)} \tag{5}$$

where

$$p_i(t) = \begin{cases} p_i & \text{queue is nonempty} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

5 RADIO RESOURCE ASSIGNMENT

5.1 Assumptions and Objective

After Packet Scheduling, the scheduled packets should be assigned to suitable radio resource(eg.time slots) for data transmission. We can formulate the resource assignment problem as an integer optimization problem that uses an objective function given in Equation(7).

For simplicity of problem formulation, we give the following assumptions:

1. Each time pick up one packet from MPQ for allocation;
2. Each packet can only be divided into two parts (It can easily be extended to K parts);
3. Each time slot can only be assigned to one packet or part of one packet if splitted;
4. Assume that each QoS type has a upper bound capacity $C_i, i \in \{UGS, ertPS, rtPS, nrtPS, BE\}$.

In a frame, based on the channel model, the real transmission rate of time slot (i,j) is $r_{i,j}$ $(i = 1, 2, \ldots, M; j = 1, 2, \ldots, T)$. The transmission rate of the packet being transmitted in time slot (i,j) in current frame is $b_{i,j}$ ($b_{i,j} \leq r_{i,j}$). Our objective is

$$\max \sum_{i=1}^{M} \sum_{j=1}^{T} b_{i,j} \tag{7}$$

s.t. $b_{i,j} \leq r_{i,j}$, $\sum_{(i,j) \in UGS} b_{i,j} \leq C_{UGS}$, $\sum_{(i,j) \in ertPS} b_{i,j} \leq C_{ertPS}$, $\sum_{(i,j) \in rtPS} b_{i,j} \leq C_{rtPS}$, $\sum_{(i,j) \in nrtPS} b_{i,j} \leq C_{nrtPS}$ and $\sum_{(i,j) \in BE} b_{i,j} \leq C_{BE}$

5.2 Heuristic Assignment Algorithms

In this section, in order to achieve the objective function, we propose two different heuristic algorithms to allocate time slots for packets with the purpose of gaining higher time slot utilizations. These two algorithms are both iterative and in each step they pick only one packet from the MPQ for scheduling. The first one is named *Packet Splitting* (PS) while the other is *Slots Combination* (SC).

Packet Splitting Algorithm

Main idea of this algorithm is presented as follows. Each time a packet is picked from the head of the MPQ, one or two time slots are selected to allocate for it. According to the packet's size, the time slot with the minimal transmission rate that is able to transmit the packet is selected. If a packet is too large to be transmitted by any time slot, it is considered to be split and transmitted by two time slots. Firstly the maximal available time slot is selected and transmits a portion of the packet, the transmitted proportion is equal to the transmission capacity of the time slot. As for the remaining part of the packet, another proper time slot is required. This time slot should be as small as possible, but large enough to transmit the remaining part of the packet. The PS algorithm is detailedly described in Algorithm 1.The computation complexity of time slot sorting is $O(nlog_2n)$ and that of searching the proper time slot for assignment is log_2n. Thus the computation complexity of PS Algorithm is $O(nlog_2n)$.

Slots Combination Algorithm

This algorithm is quite different from the PS Algorithm on the aspect of time slot selection. It consists of three steps: combination, sorting, and allocation. In the first step, a combination operation is performed. Each two time slots are merged into a pair, referring to Line 3 in Algorithm 2. Hence there are $C^2_{M \cdot T}$ time slot-pairs for the $M \cdot T$ time slots. In the second step, the combined time slot-pairs are sorted together with the original time slots in an ascending order, therefore there are $C^2_{M \cdot T} + M \cdot T$ elements in the array. In the third step, the packet picked from the head of MPQ is allocated with a single or pair time slot whose transmission capacity is as small as possible but is able to cover the packet's size. One more thing should be noticed is that, after a time slot or time slot pair was allocated, all the time slots or time slot pairs consisting of it should be removed from the array. Description of the SC Algorithm is presented in Algorithm 2. The computation complexity of SC is $O(n^2)$, thus the computation complexity of SC Algorithm is also $O(n^2)$.

Algorithm 1 The Packet Splitting Algorithm

```
procedure PACKETSPITTNG(transmission rate $r_{i,j}$, packets in MPQ)
    sort $r_{i,j}$ $(i = 1, 2, \dots, M; j = 1, 2, \dots, T)$ from max to min;
    $S \leftarrow \{(i, j) : i = 1, 2, \dots, M; j = 1, 2, \dots, T\}$
    while $S \neq \varnothing$ and MPQ is nonempty do
        pick the first packet $q$ from MPQ
        $a \leftarrow$ size of $q$
        $r_{max} \leftarrow \max\{r_{i,j} : (i, j) \in S\}$
        if $a \geq 2 \cdot r_{max}$ then
            the packet cannot be transmitted in current frame
        end if
        if $r_{max} \leq a \leq 2 \cdot r_{max}$ then
            $(i_0, j_0) \leftarrow r_{max}$
            assign channel $(i_0, j_0)$ to this packet
            $S \leftarrow S - \{(i_0, j_0)\}$
            $a \leftarrow a - r_{i_0,j_0}$
            $(i_1, j_1) \leftarrow \min\{r_{i,j} : r_{i,j} \geq a, (i, j) \in S\}$
            assign channel $(i_1, j_1)$ to this packet
            $S \leftarrow S - \{(i_1, j_1)\}$
        else
            $(i_2, j_2) \leftarrow \min\{r_{i,j} : r_{i,j} \geq a, (i, j) \in S\}$
            assign channel $(i_2, j_2)$ to this packet
            $S \leftarrow S - \{(i_2, j_2)\}$
        end if
    end while
end procedure
```

Algorithm 2 The Slots Combination Algorithm

```
procedure SLOTSCOMBINATION(transmission rate $r_{i,j}$, packets in MPQ)
    $R_1 \leftarrow \{r_k^1 : k = 1, 2, \dots, M \times T\}$ where $r^1_{(i-1)\cdot T+j} \leftarrow r_{i,j}$
    Combine each two slots to a slot pair with transmission rate of $r^2_{m,n} \leftarrow r^1_m + r^1_n$
    $R_2 \leftarrow \{r_k^2 : k = 1, 2, \dots, C^2_{M \times T}\}$
    $R = R_1 \bigcup R_2$, $r_k$ is the $k$-th element in $R$ $(|R| = C^2_{M \times T} + M \times T)$
    sort $r_k$ $(k = 1, 2, \dots, C^2_{M \times T} + M \times T)$ from min to max;
    $S \leftarrow \{(i, j) : i = 1, 2, \dots, M; j = 1, 2, \dots, T\}$
    while $S \neq \varnothing$ and MPQ is nonempty do
        pick the first packet $q$ from MPQ
        $a \leftarrow$ size of $q$
        find the minimum $r_k \in R$ where $r_k \geq a$
        if more than 1 $r_k$ found then
            select the $r_k$ with combined slots
        end if
        if $r_k$ is transmission rate of combined slots $(i, j)$ and $(m, n)$ then
            assign channel $(i, j)$ and $(m, n)$ to this packet
            $S \leftarrow S - \{(i, j), (m, n)\}$
            remove all $r_l \in R$ where $r_l$ is combined by $(i, j)$ and $(m, n)$
        else
            $(i, j)$ is the channel whose transmission rate is $r_k$
            assign channel $(i, j)$ to this packet
            $S \leftarrow S - \{(i, j)\}$
            remove all $r_l \in R$ where $r_l$ is combined by $(i, j)$
        end if
    end while
end procedure
```

6 SIMULATION AND EVALUATION

6.1 Experiment Setup

In order to evaluate the performance of our proposed resource assignment algorithms, we set up a java based simulator for our experiments. In the simulator, we define a scheduler to assign slots, in which we realize four resource assignment algorithms, i.e., FIFO, Random, PS and SC.

6.2 Implementation and Metrics

We implement the PS algorithm and SC algorithm to perform resource assignment. In addition, the FIFO algorithm and Random algorithm are also implemented as benchmarks. In the FIFO algorithm, time slots are selected in the sequence according to their positions in the frame, and then assigned to packets. In the Random algorithm, a packet is assigned with one or two time slots which are selected randomly from the frame.

The time slot utilization is defined as the ratio of the sum of all transmitted packets' size over the total transmission capacity of all time slots in a frame, i.e.,

$$\tau = \frac{\sum_{k=1}^{K} s_k}{\sum_{i=1}^{M} \sum_{j=1}^{T} r_{i,j}} \tag{8}$$

where K denotes number of the packets being transmitted in current frame, and s_k is the size of the k-th packet.

6.3 Parameters and Results

We run two set of simulations for the four algorithms: FIFO, Random, PS, and SC. Packet sizes are randomly distributed between 128 to 512 bits, but with different distributions. In the first set of simulations it is uniform distribution while in the second set it is Gaussian distribution. The total generated packets number is 5000. In a frame, the frequency number is 5 and slot number is 100, and the frame count is 100.

The range of the transmission rate of each time slot is normally from 256 to 512 bits per slot. Generally the value of transmission rates (named states) are discrete, therefore the number of states is limited. Assume that there are K different transmission rates, the value set of transmission rates is $R = \{256, 256 + \Delta r, 256 + 2 \cdot \Delta r, \ldots, 512\}$ where $\Delta r = \lfloor (512 - 256)/K \rfloor = \lfloor 256/K \rfloor$. The transmission rate of each time slot in a frame is described in Channel Model (Detail in section 3.2). In the simulations, we set the value of K to $2, 4, 8, \ldots, 128, 256$ in order to observe the performance of the four

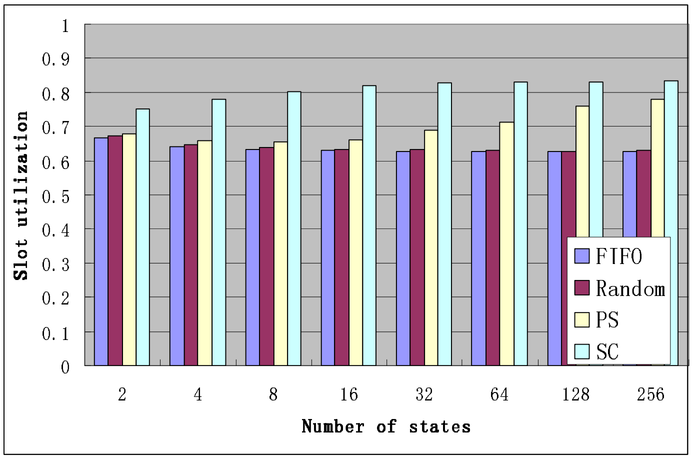

FIGURE 3
Slot utilization with different resource assignment algorithms (Uniform distribution)

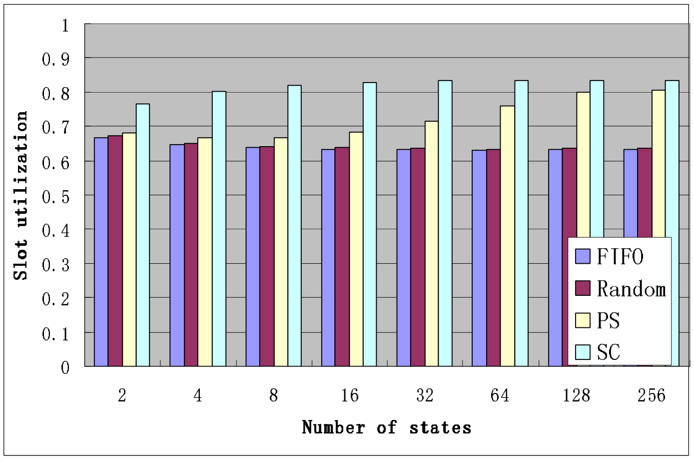

FIGURE 4
Slot utilization with different resource assignment algorithms (Gaussian distribution)

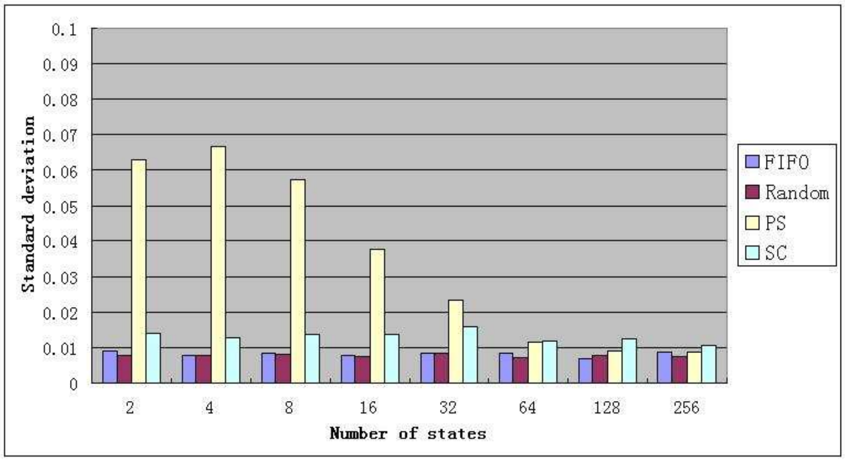

FIGURE 5
Standard deviation with different resource assignment algorithms (Uniform distribution)

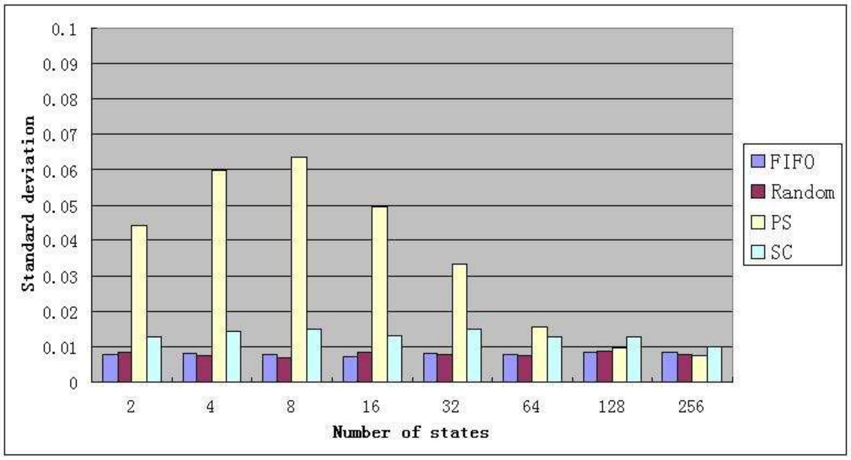

FIGURE 6
Standard deviation with different resource assignment algorithms (Gaussian distribution)

algorithms in different conditions. We run the simulation 100 times to get the mean values for performance evaluations.

Simulation results are shown in Fig. 3, Fig. 4, Fig. 5 and Fig. 6 where packet sizes are in unform distribution and Gaussian distribution respectively. In the first two figures, X axis denotes the number of states, while Y axis denotes the slot utilizations of FIFO, Random, Packet Splitting and Slots Combination. In the next two figures, X axis remains the same, but Y axis denotes the standard deviation of these four resource assignment algorithms.

6.4 Result Analysis

It is observed from Fig. 3 and Fig. 4 that among all the four algorithms, the SC algorithm shows the most mean slot utilization. Also the PS algorithm performs more better than FIFO and Random algorithms, especially when the number of states is large. The performance of FIFO algorithm and Random algorithm seems almost the same. In Fig. 5 and Fig. 6, the standard deviation of PS algorithm is larger than the other three algorithms in the first five states. However, along with the increasing of the number of states, the standard deviation of PS algorithm dropped rapidly. In the last three states, the standard deviation of PS algorithm is almost the same with the other three algorithms, or even less. That means the more states, there is less performance variation of PS algorithm from the mean slot utilization. Considering all these factors, SC algorithm has the most mean slot utilization with relatively small standard deviation, which shows the best performance.

Another thing should be noticed is that as the number of states increasing, the time slot utilization of our two algorithms (i.e., PS and SC) also rises. This is because more states mean more candidate time slots to choose, therefore the scheduler has more opportunities to choose the smallest but large enough time slots to allocate for a given packet.

7 CONCLUSIONS

We have presented a joint Packet Scheduling and Radio Resource Assignment based on GPSS framework for IEEE 802.16 networks. For Packet Scheduling, we propose a novel hybrid Queuing Analytical Model which has two scheduling phases. The two phased design could beautifully achieve the system QoS requirements both on inner of each Sub-service-queues and the relation among them. Following the QoS guaranteed Packet Scheduling, the combined two layered resource assignment scheme based on GPSS is proposed to greatly enhance the system throughput. Take uniform distribution for example, SC algorithm greatly enhance the utilization of radio resource from 63% up to 83% on average compared to FIFO and Random algorithms. We find that SC algorithm not only has amazing performance on improving system throughput, but also has relatively low complexity (The computation complexity is $O(n^2)$).

In summary, the joint Packet Scheduling and Radio Resource Assignment based on GPSS mode can guarantee the QoS requirement, at the same time, it provides a unified resource assignment to greatly maximize system throughput.

ACKNOWLEDGMENTS

The paper is supported by CityU Strategic Research Grant No. 7008110, ShenZhen Basic Research Grant No. JC200903170456A and China NSF No. 61070222/F020802.

REFERENCES

[1] IEEE standard for local and metropolitan area networks part 16: Air interface for fixed broadband wireless access systems. Technical report, IEEE std 802.16TM-2004.

[2] Syed Hussain Ali, Ki-Dong Lee, and Victor C.M. Leung. (Febrary 2007). Dynamic resource allocation in OFDMA wireless metropolitan area networks. *IEEE Wireless Communications*, 14(1):6–13.

[3] Harsha Gowda, Ramya Lakshmaiah, Manjot Kaur, Chandrashekar Mohanram, Manjeet Singh, and Shashidhara Dongre. (Febrary 2007). A slot allocation mechanism for diverse QoS types in OFDMA based IEEE 802.16e systems. In *The 9th International Conference on Advanced Communication Technology*, volume 1, pages 13–17. IEEE.

[4] SangJun Ko and KyungHi Chang. (Febrary 2007). Capacity optimization of a 802.16e OFDMA/TDD cellular system using the joint allocation of sub-channel and transmit power. In *The 9th International Conference on Advanced Communication Technology*, volume 3, pages 1726–1731. IEEE.

[5] Qingwen Liu, Xin Wang, and Georgios B. Giannakis. (May 2006). A cross-layer scheduling algorithm with QoS support in wireless networks. *IEEE Transactions on Vehicular Technology*, 55(3):839.

[6] Dusit Niyato and Ekram Hossain. (November 2006). A queuing-theoretic and optimization-based model for radio resource management in ieee 802.16 broadband wireless networks. *IEEE Transactions on Computers*, 55(11):1473–1488.

[7] Won-Hyoung Park, Sunghyun Cho, and Saewoong Bahk. (June 2006). Scheduler design for multiple traffic classes in OFDMA networks. In *International Conference on Communications (ICC)*, pages 790–795. IEEE.

[8] R. Perumalraja, J. Jackson Juliet roy, and S. radha. (September 2006). Multimedia supported uplink scheduling for IEEE 802.16d OFDMA network. In *Annual India Conference*, pages 1–5. IEEE.

[9] Bo Rong, Yi Qian, and Kejie Lu. (June 2007). Integrated downlink resource management for multiservice WiMAX networks. *IEEE Transactions on Mobile Computing*, 6(6):621–632.

[10] Vandana Singh and Vinod Sharma. (April 2006). Efficient and fair scheduling of uplink and downlink in IEEE 802.16 OFDMA networks. In *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE.

[11] K. Vinay, N. Sreenivasulu, D. Jayaram, and D. Das. (April 2006). Performance evaluation of end-to-end delay by hybrid scheduling algorithm for QoS in IEEE 802.16 network. In *International Conference on Wireless and Optical Communications Networks*. IEEE.

[Description of Channel Model]

Assume that $X_{i,j}(l)$ is a Finite state Homogeneous Markov Chain (FHMC) with states $0, 1, \ldots, K-1$ having

$$\left\{\begin{array}{lcl} P_{k,k+1} & = & \alpha_{i,j}(k=0,1,\ldots,K-2) \\ P_{k,k-1} & = & \beta_{i,j}(k=1,2,\ldots,K) \\ P_{k,k} & = & \gamma_{i,j}(k=1,2,\ldots,K-1) \\ P_{0,0} & = & 1-\alpha_{i,j} \\ P_{K,K} & = & 1-\beta_{i,j} \end{array}\right. \tag{9}$$

where $\alpha_{i,j}+\beta_{i,j}+\gamma_{i,j}=1, \alpha_{i,j}\neq\beta_{i,j}$, and $\alpha_{i,j},\beta_{i,j},\gamma_{i,j}\geq 0$, for $i=1,2,\ldots,M; j=1,2,\ldots,T$. The state transition diagram of $X_{i,j}(l)$ is show in Fig. 7.

And the transition matrix of $X_{i,j}(l)$ is

$$P_{i,j}=\left(\begin{array}{cccccc} 1-\alpha_{i,j} & \alpha_{i,j} & 0 & \ldots & 0 & 0 \\ \beta_{i,j} & \gamma_{i,j} & \alpha_{i,j} & \ldots & 0 & 0 \\ 0 & \beta_{i,j} & \gamma_{i,j} & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \gamma_{i,j} & \alpha_{i,j} \\ 0 & 0 & 0 & \ldots & \beta_{i,j} & 1-\beta_{i,j} \end{array}\right)_{K\times K}$$

Let the limiting distribution of $X_{i,j}(l)$ be

$$\Pi_{ij}=(\pi_{i,j,0},\pi_{i,j,1},\ldots,\pi_{i,j,K-1}) \tag{10}$$

where $\pi_{i,j,k}$ is the limiting probability of $X_{i,j}(l)$ for $k=0,1,\ldots,K-1$. Then Π_{ij} is the unique nonnegative solution of

$$\left\{\begin{array}{l} \Pi_{ij}=\Pi_{ij}\times P_{ij} \\ \Pi_{ij}\times\mathbf{1}'=\Pi_{ij} \end{array}\right. \tag{11}$$

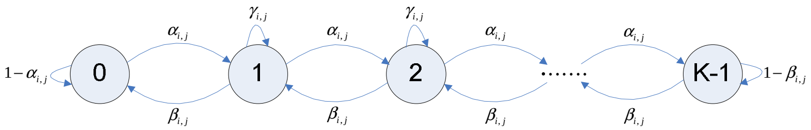

FIGURE 7
State transition diagram of channel (i,j)

where $\mathbf{1} = (1,1,\ldots,1)_K$. That is

$$\left\{\begin{array}{ll} \pi_{i,j,0} = (1-\alpha_{i,j})\pi_{i,j,0} + \beta_{i,j}\pi_{i,j,1} & \text{k=0} \\ \pi_{i,j,k} = \alpha_{i,j}\pi_{i,j,k-1} + \gamma_{i,j}\pi_{i,j,k} & \\ +\beta_{k,j}\pi_{i,j,k+1} & \text{k=1,2,\ldots,K-2} \\ \pi_{i,j,K-1} = \alpha_{i,j}\pi_{i,j,K-2} & \\ +(1-\beta_{i,j})\pi_{i,j,K-1} & \text{k=K-1} \\ \sum_{k=0}^{K-1}\pi_{i,k} = 1 & \end{array}\right. \tag{12}$$

Then

$$\pi_{i,j,k} = (\frac{\alpha_{i,j}}{\beta_{i,j}})^k \times \frac{1-\frac{\alpha_{i,j}}{\beta_{i,j}}}{1-(\frac{\alpha_{i,j}}{\beta_{i,j}})^K} \tag{13}$$

Assume that channel (i,j) is in state $k$in l-th frame. That is, $X_{i,j}(l) = k$ Let $Y_{i,j}$ be the time of channel (i,j) holding on state k after l-th frame. That is, channel (i,j) will first leaving state k in $l + Y_{i,j}(l)$-th frame after l-th frame, Then we can get

$$\begin{array}{ll} & P\{Y_{i,j}(l) = n\} \\ = & P\{X_{i,j}(l+1) = k, X_{i,j}(l+2) = k, \ldots, \\ & X_{i,j}(l+n-1) = k, \\ & X_{i,j}(l+n) \neq k | X_{i,j}(l) = k\} \\ = & \prod_{m=0}^{n-2} P\{X_{i,j}(m+1) = k | X_{i,j}(m) = k\} \\ & \times P\{X_{i,j}(n) \neq k | X_{i,j}(n-1) = k\} \\ = & \left\{\begin{array}{ll} (1-\alpha_{i,j})^{n-1}\alpha_{i,j} & k = 0 \\ \gamma_{i,j}^{n-1}(1-\gamma_{i,j}) & k = 1,2,\ldots,K-2 \\ (1-\beta_{i,j})^{n-1}\beta_{i,j} & k = K-1 \end{array}\right. \end{array}$$

Therefore, $Y_{i,j}(l)$ follows geometric distribution and independent with l.