A CROSS LAYER SCHEDULING SCHEME FOR WIMAX

CHANDRASEKARAN V\textsuperscript{1}, NAGARAJAN N\textsuperscript{2},
\textsuperscript{1} Anna University, Department of ECE, Chennai, India
\textsuperscript{2} Coimbatore Institute of Engineering and Technology, Coimbatore, TN, India

E-mail: \textsuperscript{1}vccpt@yahoo.in, \textsuperscript{2}swekalnag@gmail.com

ABSTRACT

In this paper, we propose a novel scheduling scheme that will provide quality-of-service (QoS) support in the downlink (DL) of a WiMAX system. The proposed Subchannel mapping Scheduling Algorithm completely exploits the physical layer properties of the Orthogonal Frequency Division Multiple Access (OFDMA) in order to improve the throughput of the WiMAX network at the same time improving the QoS of the various services guaranteed by the network. The proposed scheduling scheme has three parts. The first stage of the scheduling shame is the optimized queuing model that will carry out effective mapping between the admitted service flows and the packet arrival statistics. The performance of the algorithm is improved further by the enhancement scheme which would further improve the performance of the proposed scheme through fuzzy logic design. At the last stage we have the slot allocation process which completes the process. The cross-layer approach followed in the scheduling scheme provides effective slot allocation under various channel conditions. It also offers more options for the network to cater delay-sensitive multimedia traffic by reducing the delay in the network. Simulation results have demonstrated that the proposed scheduling scheme improves the throughput and also reduces the delay of the WiMAX network.

Key words: Mobile WiMAX, Orthogonal Frequency Division Multiple Access (OFDMA), Scheduling, Quality-Of-Service.

1. INTRODUCTION

One of the key features of the IEEE 802.16m Mobile WiMAX system is its strong quality-of-service (QoS). IEEE 802.16m Mobile WiMAX provides multiple QoS classes for voice, video, and data applications [1], [2], [3]. To meet QoS requirements especially for voice and video transmissions with delay and delay jitter (delay variation) constraints, the key issue is how to allocate resources among contending users. That is why there are many papers on designing resource allocation algorithms for IEEE 802.16m Mobile WiMAX [3].

The resource allocation in IEEE 802.16m Mobile WiMAX is in units of number of slots. Each slot consists of one subchannel allocated for the duration of some number of orthogonal frequency division multiplexing (OFDM) symbols. The number of subcarriers in the subchannel and the number of OFDM symbols in the slot depend upon the link direction—uplink (UL) or downlink (DL) and the subchannelization mode. For example, in the partially used sub-channelization (PUSC) mode, one slot consists of one subchannel over two OFDM symbol periods for downlink and one subchannel over three OFDM symbol periods for uplink [2], [3],[4].

The IEEE 802.16m Mobile WiMAX standard supports bi-directional communication by both frequency-division duplexing (FDD) and time-
division duplexing (TDD). For FDD, the uplink and the downlink use different frequency bands. For TDD, the uplink traffic follows the downlink traffic in time domain. The UGS algorithm discussed in this paper can be used for both FDD and TDD systems. However, to keep the discussion focused, we use a TDD system. Although the standard allows several configurations such as relay networks, our focus is only on the point to multipoint network configuration. Thus, a base station (BS) is the single resource controller for both uplink and downlink directions for all mobile stations (MSs).

IEEE 802.16m Mobile WiMAX offers five classes of service: unsolicited grant service (UGS), extended real-time polling service (ertPS), real-time polling service (rtPS), non-real-time polling service (nrtPS), and best effort (BE) classes. UGS is designed for constant bit rate (CBR) traffic with strict throughput, delay, and delay jitter constraints. ertPS is a modification of UGS for voice with silence suppression. rtPS is designed for variable bit rate voice, video, and gaming applications that have delay constraints. nrtPS is for streaming video and data applications that need throughput guarantees but do not have delay constraints (the packets can be buffered). BE is designed for data applications that do not need any throughput or delay guarantees. Note that in practice, the carrier may provide some levels of service guarantee for BE traffic.

These five service classes can be divided in two main categories: on real-time and real-time. nrtPS and BE are in the first category; UGS, rtPS, and ertPS are in the second category. For the first category, common schemes can directly apply such as weighted fair queue (WFQ) and a variation of round robin (RR) since there are no hard constraints on delay and delay jitter[2]–[7]. On the other hand, real-time services have strict constraints on these parameters. This makes scheduling difficult in trying to meet the delay constraint and tolerate the delay jitter with optimal throughput.

UGS is one of the real-time services. Basically, UGS traffic provides a fixed periodic bandwidth allocation. Once the connection is setup, there is no need to send any other requests. UGS is designed and used commonly for constant bit rate (CBR) real-time traffic such as leased-line digital connections (T1/E1) and voice over IP (VoIP). The main QoS parameters are maximum sustained rate, maximum latency, and tolerated jitter (the maximum delay variation).

The performance of the system is completely dependent on the resource allocation or the scheduling scheme. However, in WiMAX the choice of the scheduling scheme is left to the user. Most of the existing research allocation or scheduling mechanisms that are available in literature tend to work well on themselves. However, they do not combine well with physical layer in order to maximize the throughput of the system. We shall list a few recent citations to prove the same. In [8], the authors have attempted to evaluate the performance of Weighted Fair Queuing (WFQ) scheduling algorithm and Opportunistic Weighted Fair Queuing (OWFQ) and have proposed an enhanced opportunistic based WFQ scheduling algorithm called channel and duration aware WFQ (cd-WFQ). The proposed algorithm reduces the average delay by giving priority to packet with longer queuing duration. Although the algorithm performs better in terms of average delay and fairness compared to OWFQ the algorithm does not understand the principle of slot allocation WiMAX which leads to its failure.

A downlink scheduling scheme called adaptive priority-based downlink scheduling (APDS) for providing QoS guarantees in IEEE 802.16 networks is proposed in [9]. The proposed scheme in [9] concentrates on two factors, namely, Priority assignment and resource allocation. The slots are allocated based on the different service-type connections and their QoS requirements. Though the proposed scheme in [9] addresses the need for starvation of resources it clearly not cross layered.

The authors in [10] propose three-tier QoS service architecture and scheduling schemes to provide QoS supports in WiMAX networks with point-to-multi-point (PMP) topology. The distinct feature of the proposal is that QoS provisioning has
been dynamically distributed to three tiers and implemented independently by both base station (BS) and subscriber stations (SSs). Though the proposed scheme improves the performance of the WiMAX network in terms of packet delay and packet loss rate, the scheme achieves the same at the cost of decreased throughput and loss of QoS features.

A new scheduler with call admission control was proposed based on Latency-Rate (LR) server theory and with system characteristics as specified by the system standard using the Wireless MAN-OFDM (Orthogonal Frequency Division Multiplexing) air interface in [11]. The proposed cross layer scheduling algorithm had two challenges though. First, the cross layer approach has lost its significance in the present context as WiMAX has migrated to OFDMA from OFDM. Second, the proposed approach was not suitable on Mobile Stations (MS) as the scheduler calculates the time frame (TF) in order to maximize the number of stations allocated in the system while managing the delay required for each user.

[12] proposes a new technique called swapping min-max (SWIM) for scheduling that not only meets the delay constraint with optimal throughput, but also minimizes the delay jitter and burst overhead. However, the proposed scheme in [12] is only suitable for UGS scheme as it lets the other services starve for resources.

The authors in [13] have examined the various design issues that are associated with multimedia downlink scheduling in the multicast/broadcast-based WiMAX system and have proposed an end-to-end framework that includes source coding, queue prioritization, flow queuing, and scheduling. However, the proposed framework uses OFDM whereas the WiMAX has already migrated to OFDMA systems and the benefits of the OFDMA system are very high.

The authors in [14] have proposed a multicast services-based scheduling (MSBS) algorithm that improves the energy efficiency of both unicast and multicast services, while satisfying the quality of service requirements of the MSSs in 802.16m wireless networks. MSBS schedules the packets in such a way that each packet is transmitted before its deadline and the energy consumed by the MSSs is reduced by minimizing the number of state transitions by the MSSs. However [14] focuses more on energy optimization rather than enhance bandwidth utilization.

One of the closest and best cross layer approaches in WiMAX was proposed in [15]. Usually the OFDMA resource allocation algorithms aim to determine which users to schedule, how to allocate subcarriers to them, and how to determine the appropriate power levels for each user on each subcarrier. The authors of [15] were able to demonstrate a pretty good cross layer approach that satisfied the demands of OFDMA based cross layered approach. However, one serious issue that the authors failed to address was the varying channel qualities. In WiMAX it is impossible to have the same channel quality for all the channels which makes the algorithm proposed in [15] suitable only for synthetic networks.

As discussed, there are a few algorithms available in market that use a cross layer approach for effective scheduling in WIMAX systems and even more most the cross layer approaches compromise on one factor or more in order to achieve its goal. This warrants the need for an effective cross layer approach that will ensure effective bandwidth utilization at the same time providing fair resource allocation across all the services of WiMAX system. In this paper, we propose a novel scheduling scheme that will exploit the physical layer properties of the WiMAX network in order to maximize the throughput of the network at the same time guaranteeing other QoS parameters. The proposed scheme combines well with the OFDMA slot allocation mechanism thereby improving the resource utilization. Combined with a proper Call Admission Control procedure the scheduling scheme would ensure maximize utilization of the precious bandwidth.

2. ALGORITHM:

In this paper, we propose a new scheduling scheme called the subchannelisation mapping algorithm
that understands and exploits on how the physical resource allocation is carried out in WiMAX systems. The WiMAX system uses Orthogonal Frequency Division Multiple Access (OFDMA) as the modulation scheme. In order to reduce the signaling overhead subcarriers are grouped into resource blocks (RBs). For the sake of conceptual clarity, the RSs are referred to as scheduling blocks (SBs) in future in this paper.

For the sake of conceptual, we first introduce the various symbols and terminologies used in the algorithm before explaining the working principle of the algorithm. Each SB has duration of 1 ms and consists of typically 12 or 14 OFDMA symbols. The OFDMA symbols are represented as $R_{OF}$. Let $M$ be the total number of subcarriers in the channel. $M(S) \leq M$ is the data carrying channels, where $S = 1, 2, \ldots, R_{OF}$. Let $R_{K(c)}$ be the code rate associated with the modulation scheme, $K \in (1, 2, \ldots, K)$; $R_{i}$ is the constellation size of the modulation scheme $K$ and $T_{s}$ is the OFDMA symbol duration. The bit rate $r_{k}$ for a single SB is given by

$$r_{k} = \frac{R_{K(c)} \log_{2} R_{K}}{T_{s} R_{OF}} \sum_{S=1}^{R_{OF}} M(S) \tag{1}$$

Let $S$ be the number of users using the channel at the same time and $R_{i}$ be the number of SB’s that are being transmitted at a particular time slot. Let $R_{i}$ be the subset of $R_{i}$ SB, whose channel quality index values are to be reported back by the particular user through the CQICH of the WiMAX system. The size of $R_{i}$ is determined by the feedback obtained through the CQICH.

The channel qualities of all the subcarriers are indicated by real scalar or vector sent back by user ‘$i$’. Let $m_{i,n} \in \{1, 2, \ldots, \} R_{i}$ be the real scalar or vector representation of the channel quality for a particular user ‘$i$’ in a SB. Let $Q_{i,m_{i,n}} \in \{1, 2, \ldots, \} n$ be the index of the modulation scheme that can be supported by user ‘$i$’ for the particular scheduling block having a channel quality value $m_{i,n}$ such that

$$Q_{i,m_{i,n}} = \max_{n \in N_{i}} \{ Q_{i,m_{i,n}} \}
$$

Scheduling blocks whose channel qualities that have been reported back are assigned to the current modulation scheme which is being served.

Our significant feature of the WiMAX downlink system is that all the scheduling blocks that are assigned to a particular user in a particular time frame are allocated the same modulation scheme. Say for instance, a particular modulation scheme is to be used for user ‘$i’$, and then only certain scheduling blocks are assigned to that user. For example, if $R_{i} = 4$ and

$$\begin{align*}
Q_{i,m_{i,2}} &< Q_{i,m_{i,3}} \\
Q_{i,m_{i,3}} &< Q_{i,m_{i,4}} \\
Q_{i,m_{i,4}} &< Q_{i,m_{i,5}}
\end{align*} \tag{3}$$

Suppose if the modulation scheme $K = Q_{i,m_{i,2}}$ is being served, then only SB’s $n=3$ and 5 can be allocated to the particular user because only that block has got enough channel quality to support that modulation scheme. If a particular channel does not have the required value of channel quality than the WiMAX station will allocate a higher modulation scheme to that particular user to meet the QoS requirements of that user. Thus an optimal value of ‘$K’ is required to maximize the bit rate for user ‘$i’.

The proposed algorithm has two stages. First the optimization process is discussed and the enhancement scheme which would further improve the performance of the proposed scheme through fuzzy logic design is elaborated followed by the slot allocation procedure.

The algorithm first determines the modulation scheme and then allocates the slots to the particular user ‘$i’ so as to ensure that the maximum data rate $R_{i}$ is achieved for the given channel quality.

Let $Q_{i,m_{i,n}} \in \{ Q_{i,m_{i,n}} \}, n \in N_{i}$

Let $Q_{i,m_{i,n}} = \max_{n \in N_{i}} \{ Q_{i,m_{i,n}} \}$
And let $a_i = [a_{i,1}, a_{i,2}, ... ... a_{i,q_{\text{max}}(i)}]$ be the modulation scheme vector for user $i'$. 

$$a_{i,j} = \begin{cases} 0, & \text{if user is not assigned } \text{Mod. sch} \\ 1, & \text{if user is assigned the Mod. sch} \end{cases}$$

(4)

Where Mod. sch indicates modulation scheme

The optimal value of $a_i$ which can maximize the total bit rate for the user $i'$ is given by

$$a_{i,j}^{\text{max}} = \max_j a_i \sum_{k=1}^{q_{i,\text{max}}(m_{i,n})} a_{i,k} r_k$$

(5)

Which means that the channel quality index is higher than lower modulation schemes with higher data rate may be assigned to the customer. On the other hand when the channel experiences sufficient degradation, then the system resorts to provide higher modulation schemes to the user thereby reducing the data rate which is given by

$$\sum_{k=1}^{q_{i,\text{max}}(m_{i,n})} a_{i,k} = 1; Y_i$$

(6)

$$a_{i,k} \in \{0,1\}; Y_{i,k}$$

(7)

Equation (5) is possible if the selected bit rate for a particular slot is less than $m_{i,n}$ can support. Equation (6) ensures that the assigned modulation scheme for a particular user can take only values between 1 and $Q_{\text{max}}(i)$. The optimization problem can be easily addressed as follows:

Let $A^{(i)}$ be an $N_{i} * Q_{\text{max}}(i)$ matrix with $(n,m)^{th}$ element such that

$$\left\{ r_{n,m}^{i} = r_{m}^{i}; m = 1,2, ... Q_{i,\text{max}}(m_{i,n}) \right\}$$

The sum of the elements in the $m^{th}$ column can be given by

$$b_{m}^{(i)} = \sum_{n=1}^{N_{i}} r_{n,m}^{(i)}$$

(8)

The optimal modulation scheme for a particular user $i'$ is

$$k^{*} = \arg \max_{l \leq k \leq Q_{i,\text{max}}^{(i)}} b_{m}^{(i,v)}$$

(9)

And the corresponding maximum bit rate is $b_{m}^{(i)}$. Once the optimization is finished and the modulation scheme determined we enter into the process of slot allocation.

Let us now discuss the slot allocation mechanism for the proposed scheme. The slot allocation mechanism intelligently computes the priorities among the users to maximize the system throughput and also to improve the various QoS requirements. The following inputs are considered before the slot allocation process starts. Let $J$ represent the number of base station. In order to solve the slot allocation problem and data scheduling process, it is imperative to determine the Boolean control variable $\left(x_{k,M}^{N}\right)$ of the proposed framework which will ensure the allocation of the slots to the subscribers.

$$\left(x_{k,M}^{N}\right) = \begin{cases} 1, & \text{if } A^{*} \\ 0, & \text{otherwise} \end{cases}$$

$A^{*}$ is the $m^{th}$ slot used by the $k^{th}$ subchannel allocated by the $n^{th}$ MS.

The total number of slots that have been allocated by the network is computed by the expression $N * K * M$. The estimation of the variable $x_{k,M}^{N}$ is done assuming many parameters or constraints.

1) It is absolutely necessary to ensure that a particular slot $M'$ not be assigned to more than one MS.

$$\sum_{N=1}^{J} x_{k,M}^{N} \leq 1(k = 1,K; m = 1,M)$$

(10)

2) Fixing the number of slots to a particular MS should ensure that the BS provides the required transmission rate which is guaranteed by equation (2).

3) Scheduling condition of a single burst for the particular MS is given by
\[ x_{j,M}^N = x_{j,M}^N(j - z + 1) - \sum_{i=1}^z x_{j,M}^N(i - z) \leq 0 \quad (11) \]

Conditions 14 and 15 are used when the number of slots needed to transmit the overhead data exceeds the actual slots that are available for transmission. The slot allocation process is expected to meet the optimality criterion defined by

\[ \min \sum_{n=1}^J \sum_{k=1}^M \sum_{m=1}^N x_{k,M}^N \quad (17) \]

3. RESULTS AND DISCUSSIONS:

We also performed extensive simulations to evaluate the performance of the proposed subchannel mapping algorithm. To this end, we have evaluated the performance of the algorithm on OPNET. We have also implemented model based downlink resource allocation framework and the weighted DRR (WDRR) in OPNET which allows selectively running these schemes to compare the performance of these schemes with that of our proposed scheme. The WDRR is a well-known scheme designed for providing QoS for delay-sensitive multimedia applications in many wired and wireless systems, including WiMAX. At each scheduling instant, priority is given to packets that have earlier deadlines. WDRR is a modified version of DRR [15], which essentially is an approximation of WFQ [17], [18]. The authors of [19] also found that, in a mobile WiMAX system, the channel-adaptive design of WDRR offers better performance over the standard DRR. We therefore compare our proposed scheme to WDRR instead of the standard DRR. In order to further confirm and validate our results we have further compared the performance of our algorithm with a latest algorithm that id gaining popularity namely the model based downlink scheduling algorithm [20]. The WDRR algorithm and the model based downlink scheduling algorithm will be referred as conventional algorithm and scheduling algorithm or MAM in the Figures indicating the results.

We simulate a single IEEE 802.16m WiMAX cell with PMP architecture, in which a single BS communicates with multiple MSs. The MSs are assumed to be uniformly distributed across the cell. OFDMA-based air interface is used to establish the connection between the BS and the
MSs. We have simulated 11000 WiMAX frames for each of the simulation experiments. For simplicity, we assume that each MS is associated with one DL service flow. MCS profiles for individual MSs are randomly chosen from a collection of 10000 trace files containing frame-to-frame link adaptation information generated from a separate IEEE 802.16m link-level simulator that assumes ITU Pedestrian B3-km/h mobility model [21]. It also assumes a 2×2 multiple-input–multiple-output (MIMO) model in the DL with two transmit antennas at the BS and two receive antennas at each MS. The assumed MCS levels and their corresponding per-slot data capacities are listed in Table II. MCS levels 1, 2, and 3 employ quaternary phase-shift keying 1/2 with repetition 6, 4, and 2, respectively. The MIMO mode for MCS levels 4–11 is space–time block coding, whereas MCS levels 12–19 adopt spatial multiplexing.

3.1 Throughput:

Figure 1 illustrates that the throughput of the proposed subchannelisation mapping scheduling algorithm has increased by approximately 100 percent with respect to the WDRR algorithm by 23 percent with respect to the model based scheduling algorithm. This is because the WDRR algorithm is more of a modified version of the fair scheduling algorithm and does not account for any policy making decisions. It just aims to carry out the slot allocation procedure. On the other hand though the model based scheduling algorithm has policy decisions before allocation of slots it does not combine well with the features of the OFDMA. One such feature of the proposed scheduling algorithm that is not used in the model based scheduling algorithm is the power compatibility feature of the OFDMA system wherein slots with different power levels can be allocated to users at different locations thereby reducing the need to avoid occasional switching of the modulation schemes. The improvement in throughput by 100 percent and 23 percent with respect to the WDRR and the model based algorithm clearly indicates that the proposed scheduling algorithm combines well with the physical layer properties of WiMAX.

3.2 Delay:

In the next set of simulations, we evaluate how effectively the proposed scheduling algorithm can meet QoS guarantees for real-time multimedia applications such as streaming video, audio/video conferencing, and VoIP. These delay-sensitive applications need their packets to be transmitted within a certain delay limit and can only tolerate a
very limited amount of packet losses. In our analysis, we assume a homogeneous traffic scenario, where each MS is associated with a real-time service flow characterized by identical traffic parameters. Figure 2 clearly illustrates that the delay experienced by the delay sensitive applications is higher in the model based scheduling scheme and in the WDRR scheme. The proposed scheduling mechanism proactively manages the backlogs at the connection queues and is targeted toward meeting the packet delay guarantees, in contrast to the other mode based scheduling schemes where the goal is minimizing overall packet delay in the system. Hence the proposed scheduling scheme manages to reduce the packed delay comprehensively. It can thereby achieve a lower system-wide packet delay for more number of service flows, as is evident in Figure 2. In addition, the proposed scheduling scheme achieves an increased throughput.

3.3 Voice MOS, Jitter and Packet Delay Variation:

Although delay and throughput are two very important parameters that determine the performance of the overall WiMAX network care has to taken that the proposed scheme does not achieve the same at the cost of other QoS parameters and hence we go ahead and also evaluate the other necessary QoS parameters that are associated with voice and video. First, we evaluate the performance of voice by computing the Mean Opinion Score of Voice for all the three scheduling algorithms. It has been clearly illustrated in Figure 3 that the MOS value achieved by the proposed scheduling algorithm is comparable to the WDRR scheme and model based scheduling scheme which clearly concludes that the proposed scheme achieves fair resource scheduling. Another important QoS factor that is associated with voice is Voice Jitter. From Figure 4 it is clearly evident that the subchannelisation mapping scheduling algorithm produces a slightly lower jitter than the other two algorithms. This slight decrease in value of voice jitter may be attributed to the decrease in packet delay variation experienced by voice packets that is illustrated in Figure 5. As explained earlier the proposed scheduling scheme manages to effectively pool the resources and also do effective slot allocation to reduce the delay of the network thereby improving the various QoS Factors. Figures 3, 4 and 5 clearly indicate that the improved throughput of the system is not achieved at the cost of other QoS demands.
3.4 Video Packet Delay Variation:

Similar to the way in which the voice packets were served it is evident from Figure 6 that the improved throughput does not affect the QoS of video packets either. It can be seen from Figure 6 that the proposed subchannelisation mapping scheduling scheme works well in allocating resources to the video packets as well thereby reducing the packet delay variation in a video conferencing application.

4. CONCLUSION:

In this paper, we have proposed a novel scheduling algorithm for downlink services of IEEE 802.16m mobile WiMAX systems. At first, we have developed a queuing optimization model that relates important performance measures of an admitted service flow to its packet arrival statistics and a parameter in the packet scheduling scheme followed by enhancement scheme and a slot allocation mechanism. The proposed scheduling scheme is a cross layer approach that utilizes the defined parameters of the algorithm well with the channel quality and the backlog information to schedule packets and allocate OFDMA slots among downlink QoS connections. The proposed model-based approach offers simplicity of design for the online component, flexibility to the service providers in defining QoS parameters, and better compatibility to the IEEE802.16m standard. Instead of complexities involved with installing separate scheduling schemes within the same resource allocation framework for real-time and non-real-time traffic, the proposed subchannelisation mapping scheduling scheme offers an efficient unified way to handle both types of traffics and their diverse QoS requirements. We have run simulation experiments to evaluate the performance of our proposed scheme in providing QoS for real-time multimedia. Compared with WDRR scheduling and the model based scheduling scheme the proposed scheduling scheme achieves improved throughput without compromising the other QoS parameters. The proposed scheme works well for real time traffic. However, it would be interesting to evaluate the performance of the algorithm on non-real time traffic scenarios.

REFERENCES:


