Generalized Weighted Fairness and its Application for Resource Allocation in IEEE 802.16e Mobile WiMAX[1,2]

Chakchai So-In, Raj Jain and Abdel-Karim Al Tamimi
Department of Computer Science & Engineering,
Washington University in St. Louis
St. Louis, MO 63130 USA
{cs5, jain, aa7}@cse.wustl.edu

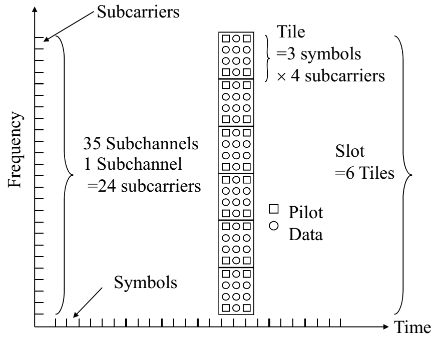
Abstract— **In wireless networks, the definition of fair resource allocation is ambiguous because the wireless channel condition is not constant over time and location. Two main ways to define fair resource allocation are: by number of allocated service time units per user, *temporal fairness*, or the number of transmitted bytes per user, *throughput fairness*. In wireless broadband networks i.e., IEEE 802.16e Mobile WiMAX, the quality of services offered by service providers is associated with the price paid. Similar to a traditional cellular phone system, the users may be required to pay air-time charges. The throughput fairness is favorable to customers with a poor channel condition or those at a far distance. The temporal fairness will be preferred by the customers near the base stations and with good channel condition. In this paper, we define the Generalized Weighted Fairness (GWF) criterion that allows carriers to implement and apply either of the two fairness criteria to a Mobile WiMAX environment. In addition, we show how a scheduling algorithm can use the GWF criterion to achieve a general weighted fair resource allocation in IEEE 802.16e Mobile WiMAX networks. We use Deficit Round Robin with Fragmentation (DRRF) as an example of a scheduling algorithm. Numerical and simulation results are presented to demonstrate the effect of GWF.**

Keywords-component; WiMAX; Mobile WiMAX; IEEE 802.16e; Generalized Weighted Fairness; Scheduling; Resource Allocation; QoS; Temporal Fairness; Throughput Fairness

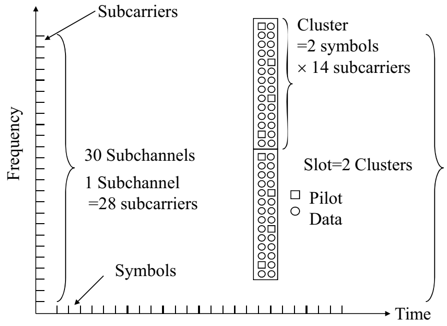
I. INTRODUCTION

The resource allocation problem in IEEE 802.16e Mobile WiMAX [1, 2] is basically how to schedule resources, i.e., number of slots, for each user in each Mobile WiMAX frame. Each slot consists of one subchannel allocated for the duration of some number of OFDM (Orthogonal Frequency Division Multiplexing) 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 subchannelization modes.

For example, in the Partially Used Sub-Channelization (PUSC) scheme, one slot consists of one subchannel over two OFDM symbol periods for downlink and one subchannel over three OFDM symbol periods for uplink. Figure 1 shows details of the slot definition for both uplink (a) and downlink (b) in the PUSC mode.



(a) Uplink Tiles



(b) Downlink Cluster

Figure 1. Symbols, Clusters, Tiles, and Slots (PUSC)

The mapping process from logical subchannel to multiple physical subcarriers is called a permutation. Basically there are two types of permutations: distributed and adjacent. The distributed subcarrier permutation is suitable for mobile users while adjacent permutation is for fixed (stationary) users. The PUSC, discussed above, is one of the distributed permutation modes. Others include Fully Used Subchannelization (FUSC) and Adaptive Modulation and Coding (band-AMC). In this paper, we focus on the PUSC mode commonly used in a mobile wireless environment [3].

The IEEE 802.16e Mobile WiMAX standard supports bi-directional communication as a frequency division duplexing (FDD) and a time division duplexing (TDD). For the FDD, the uplink and the downlink use different frequency bands. For the TDD, the uplink traffic follows the downlink traffic in the time domain. The generalized weighted fairness scheme discussed in this paper can be used for both FDD and TDD systems. However, to keep the discussion focused, we use the TDD system.

978-1-4244-5586-7/10/$26.00

Proceedings of 7th Annual IEEE Consumer Communications and Networking Conference, Las Vegas, NV, January 9-12, 2010

S_i	Number of slots allocated to mobile station *i*
B_i	Number of bytes allocated to mobile station *i*
b_i	Number of bytes per slot for mobile station *i*
N	Number of active mobile stations

Slot Fairness Scheme (temporal fairness):

$$Total\_Slots = \sum_{i=1}^{N} S_i$$

$$S_i = S_j \qquad | i, j \leq N$$

Byte Fairness Scheme (throughput fairness):

$$Total\_Slots = \sum_{i=1}^{N} S_i$$

$$B_i = B_j \qquad | i, j \leq N$$

$$B_i = b_i S_i$$

Figure 2. Formal definitions of slot fairness and byte fairness

Although the standard allows several configurations such as mesh networks and relay networks, our focus is only on 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).

Mobile WiMAX supports several Modulation and Coding Schemes (MCSs) such as Binary Phase Shift Keying (BPSK) and Quadrature Amplitude Modulation (QAM). The BPSK results in *1* bit per symbol and is used for channels in poor conditions. The QAM results in more bits per symbol and is used for more reliable channel conditions. Since the MCS used for a user depends upon the location of the user and varies with time, the slot capacity (number of bits in the slot) is also varied with time and location.

In wireless networks, the link capacity keeps changing over time and distance. The issue of defining fairness depends upon both service provider's and customers' points of views. As a result, in this paper we propose the Generalized Weighted Fairness (GWF) which equalizes the weighted sum of service time units and transmitted bytes. In Mobile WiMAX, again the service time unit is represented as a number of slots allocated to each mobile station. Mobile WiMAX equipment manufacturers can implement generalized fairness. The service providers can then set a weight parameter to any desired value and achieve either slot fairness or throughput fairness or some combination of the two.

The IEEE 802.16e Mobile WiMAX standard [1] defines five classes of service: Unsolicited Grant Scheme (UGS), extended real Time Polling Service (ertPS), real time Polling Service (rtPS), non real time Polling Service (nrtPS), and Best Effort (BE). Each of these has its own set of QoS parameters such as minimum throughput, maximum allowable delay, delay variation (jitter), etc.

The GWF can be applied to all classes of service. However, to simplify the analysis, in this paper our focus is on the BE class which has no minimum throughput requirement, delay, or delay jitter requirements. Thus, the primary goal is to maintain a fair resource allocation among BE flows by providing equal opportunities for all BE flows.

This paper is organized as follows: we first define temporal fairness or slot fairness and throughput fairness or byte fairness in Section II. We introduce the generalized weighted fairness in Section III. In Section IV, we describe steps on how to calculate the weight or the quantum, derived from the GWF equation to be used in the scheduler. Numerical and simulation results are presented in Section V. Finally, conclusions are discussed in Section VI.

II. TEMPORAL AND THROUGHPUT FAIRNESS

Unlike in traditional wired networks, in wireless networks, the available data rate changes with time and location. In IEEE 802.16e Mobile WiMAX networks, different modulation and coding schemes (MCSs) are used depending upon the signal quality.

As briefly described earlier, in wireless networks e.g., wireless local area networks or WLAN there are two ways to define fairness: temporal fairness and throughput fairness [11]. Traditionally, the first method is to allocate equal number of service time units to each user. In Mobile WiMAX networks, the resource allocation is in terms of the number of slots [1]. Therefore, here the unit used to achieve the temporal fairness is "*slot*". In other words, the base station will allocate equal number of slots to each user. The users, who are near the base station and therefore have a very good link, will be able to use good MCSs (multiple bits per symbol) and consequently get good throughput in (bytes per second). This is also called *slot fairness* or *temporal fairness*.

On the other hand, the second method is to allocate equal throughput (or bytes) to each user. The users near the base station will need fewer slots then those far away to transmit the same number of bytes. This is called *byte fairness* or *throughput fairness*. Figure 2 shows formal descriptions for both slot and byte fairness schemes.

III. GENERALIZED WEIGHTED FAIRNESS

As described in Section II, a fairness scheme can be either slot fairness or byte fairness. It can be argued which definition is better. For example, the users want to run certain applications such as VoIP (Voice over IP), which require a certain bit rate and so the users would prefer to get a guaranteed bit rate for their application or a fairshare of the available bit rate in case of best effort service. This is the argument in favor of byte fairness or throughput fairness.

The argument in favor of temporal fairness or slot fairness goes as follows. The service provider has a fixed number of slots and if the user happens to choose a bad location (such as the basement of a building on the edge of the cell), the provider will have to allocate significant number of slots to provide the same quality of service as to a user who is outside and near the base station. Since the providers have no control over the locations of users, they can argue that they will provide the same resources to all users and the throughput observed by the user will depend upon their location.

S_i	Number of slots allocated to mobile station *i*
B_i	Number of bytes allocated to mobile station *i*
b_i	Number of bytes per slot for mobile station *i*
N	Number of active mobile stations
w	Weight parameter between 0 and 1.
M	Number of bytes per slot for the highest MCS

$$Total\_Slots = \sum_{i=1}^{N} S_i$$

$$wS_i + (1-w)B_i / M = wS_j + (1-w)B_j / M \quad | \, i, j \le N$$

$$B_i = b_i S_i$$

Figure 3. Formal description of weighted fairness

Some service providers will prefer throughput fairness while the others will favor the slot fairness. The equipment manufacturers should allow both possibilities. One way to do this is to implement the weighted fairness introduced in this paper. The formal definition of the weighted fairness is as shown in Figure 3.

In general, setting the weight, *w*, to *0* results in byte fairness, and setting it to *1* results in slot fairness. A carrier can select the weight to be either of the two values or any value between *0* and *1*. Note that GWF as defined in Figure 3 uses equal weight for all stations. It is possible to generalize this further by allowing different weights for different users based on their priority or the price paid.

IV. EXAMPLE MODIFICATIONS TO A SCHEDULING ALGORITHM

The Deficit Round Robin with Fragmentation (DRRF) [6] scheduling algorithm is a potential algorithm that can be used to allocate the resources in IEEE 802.16e Mobile WiMAX networks. The DRRF is similar to a traditional deficit round robin algorithm or DRR [8] in that it is used to maintain the throughput fairness over a long period of time. Unlike DRR, the DRRF allows the packet fragmentation transmission in order to utilize the Mobile WiMAX frame. The unit of the quantum size [6, 8] is the number of slots, and is derived from the queue size, MCS, and QoS parameters in order to meet the throughput guarantee [6, 7].

To apply the concept of the generalized weighted fairness (GWF) to the DRRF, basically the quantum is updated according to the GWF equation (See Figure 3). Figure 4 shows pseudo codes, or steps on how to calculate the proper quantum size, in order to achieve a full Mobile WiMAX frame utilization. In other words, in an overload scenario, the sum of all quanta for all active MSs is the total size of the Mobile WiMAX frame. Note that the mobile stations are sorted in ascending order of number of requested slots [6] as a precondition.

In Figure 4, the first step is to compute a temporary weight (derived from the GWF equation) denoted as *MS_i.temp_weight* for each active mobile station *i*. This first step does not consider the number of free slots, queue size, and numbers of requested slots. Therefore, Step 2 calculates the proper quantum with those considerations. The accumulated weight value, denoted as *accum_weight*, is used to update the weighted value of each MS. This step (Step 3) is required so full frame utilization can be achieved. The calculation is constrained by queue size and number of free slots. This value is used to calculate a generalized weighted fair share, *gwf_fair_share*, and in Step 4, the quantum for each mobile station, *$MS_{i\_current}$.quantum*, is updated with the constraint of number of requested slots. Finally, the number of free slots is updated. This loop continues until there are no more free slots or the requested slots for all active MSs have been satisfied.

```
FOR each MS                                                            //1st Step
   $MS_i$.temp_weight = weighted_ISP + (1-
                          weighted_ISP) × $MS_i$.SlotCapacity;
END FOR

FOR each MS                                                            //2nd Step
  accum_weight = 0;
  FOR each $MS_i$ from $MS_{i\_current}$ to $MS_N$
//3rd Step
    accum_weight += $MS_{i\_current}$.temp_weight/ $MS_i$.temp_weight;
  END FOR

  gwf_fair_share = floor(free_slots/ accum_weight);
  IF (gwf_fair_share >= $MS_{i\_current}$.req_slots) THEN              //4th Step
     $MS_{i\_current}$.quantum = $MS_{i\_current}$.req_slots;
  ELSE
     $MS_{i\_current}$.quantum = gwf_fair_share;
  END IF
  Update_FreeSlotCapacity(free_slots, $MS_{i\_current}$quantum);       //5th Step
  $MS_{i\_current}$++;
END FOR
```

Figure 4. Quantum Size Derivation

V. PERFORMANCE EVALUATION

In this section, we present both numerical and simulation results of system throughput with the General Weighted Fairness (GWF) scheme based on the performance evaluation parameters specified in Mobile WiMAX System Evaluation Methodology document and WiMAX profiles [3, 4, 5]. These parameters are shown in Table I.

We consider only downlink resource allocation; the queue size in number of bytes is translated to the number of requested slots. The analysis for uplink is very similar except that the bandwidth requests are used to derive the number of requested slots. However, the base station has no information about individual packet sizes.

TABLE I
PERFORMANCE EVALUATION PARAMETERS

Parameters	Values
PHY	OFDMA
Duplexing Mode	TDD
Frame Length	5 ms
System Bandwidth	10 MHz
FFT size	1024
Cyclic prefix length	1/8
DL permutation zone	PUSC
RTG + TTG	1.6 symbol
DL:UL ratio	2:1 (29: 18 OFDM symbols)
DL Preamble	1 symbol-column
MAC PDU size	Variable length
ARQ and packing	Disable
Fragmentation	Enable
DL-UL MAPs	4 symbol-columns

In this analysis, we use OFDMA PHY with 10 MHz system bandwidth, 5 ms frame, 1/8 cyclic prefix and a DL:UL ratio of 2:1. Note that 1.6 symbol-columns are used for TTG (Transmit to Transmit Gap) and RTG (Receive to Transmit Gap). The number of downlink symbol-columns per frame is 29 [3, 4]. Of these 1 symbol-column is used for preamble and 4 symbol-columns for Frame Control Header (FCH), DL MAP and UL MAP (repetition of 4), leaving 24 symbol-columns for data transmission.

Notice that the size of DL and UL maps depends on the number of mobile stations (MSs). Four symbol-columns can carry seven MSs and FCH. We do not include optional 4 symbol-columns used to transmit Downlink Channel Descriptor (DCD) and Uplink Channel Descriptor (UCD) in this analysis.

In the Partial Usage of Subchannels (PUSC) mode, there are 30 subchannels in the downlink and each slot consists of one channel over a two symbol duration. As a result, there are 30 × (24/2) = 360 downlink slots per frame. In this analysis, there are seven Mobile Stations (MSs). Each MS has only one flow and is mapped to a different MCS used in Mobile WiMAX [5].

Table II lists the system throughput for each MCS. In the table, we show seven MCSs – one for each user. The analysis can be easily extended for more flows or MCSs. In our analysis, the interference is represented as a change of MCS.

Note that for PUSC mode with 10 MHz and 1024 FFT (Fast Fourier Transform), the number of subcarrier-symbol combinations per slot is 56 (Figure 1). Of these, 8 combinations are used as pilots leaving 48 combinations for data. With two bits per symbol in QPSK and half coding rate, this results in 48 bits (or 6 bytes) per slot. Figure 5 shows numerical results for system throughput for various MCSs when the weight values are varied from *0* to *1*.

A. Simulation Configuration

We used a modified version of the ns-2 simulator [9] in which a WiMAX module has been added [10]. In order to show the effect of GWF and validate the numerical results, we used Constant Bit Rate (CBR) traffic over UDP with 5 Mbps source rate. We simulated four mobile stations with BPSK1/2, QPSK1/2, QPSK3/4, and 16QAM1/2 MCSs to represent a variety of wireless channel conditions.

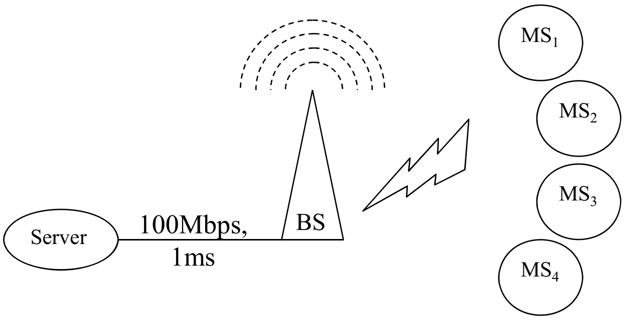


Figure 6. Simulation Topology

TABLE II
DATA THROUGHPUT ANALYSIS (DOWNLINK)

MCSs	Bit/ Symbol	Coding Rate	Bytes/ Slot	Throughput (kbps)
BPSK1/2	1	½	3	1,728
QPSK1/2	2	½	6	3,456
QPSK3/4	2	¾	9	5,184
16QAM1/2	4	½	12	6,912
16QAM3/4	4	¾	18	10,368
64QAM2/3	6	⅔	24	13,824
64QAM3/4	6	¾	27	15,552

All simulations were run from 0 to 22 seconds with 10 seconds of traffic duration. Flows start at 10, 10.02, 10.04, and 10.06 seconds and end after 20 seconds. The packet size is set at 500 bytes. At the base station, there is one queue for each mobile station and each queue is 50 packets long. We set the weight values at 0.0, 0.5, and 1.0.

The link from the base station to mobile station is the only bottleneck (Figure 6). Other simulation parameters are as shown earlier in Table I. Note that we do not use TCP flows because the slow-start feedback mechanism of TCP has a significant impact on the resource usage and it is difficult to isolate its effect from the effect of the scheduling algorithm.

In this simulation configuration, four mobile stations can saturate the link capacity so that the fairness is exercised (there are dropped packets due to buffer overflow at the BS). Although we demonstrate four CBR flows to demonstrate the effect of GWF, with more mobile stations or more numbers of flows, intuitively the effect of weighted fairness will be more obvious. We show the effect of GWF in these ranges using numerical analysis.

B. Simulation Results

We show that DRRF with GWF can achieve throughputs similar to that obtained by our numerical analysis. Figures 7, 8, and 9 show that the throughput follows the GWF with weights of 0, 0.5, and 1.0. Overall, the results are similar to numerical results. Figure 7 shows that all mobile stations achieve *byte fairness* with weight set to 0. Figure 9 shows *slot fairness* with weight 1 - the mobile station throughput is proportional to its channel condition or modulation and coding scheme (MCS). Figure 8 shows the effect of GWF with weighted fairness of 0.5.

Notice that the throughput varies because of Downlink Channel Descriptor (DCD) and Uplink Channel Descriptor (UCD) and other management messages being transmitted as well. In addition, since the packing is disabled, the anticipated number of MAC headers for each allocation is not precise; the overhead here includes both MAC header (6 bytes) and fragmentation subheader (2 bytes).

VI. CONCLUSIONS

In this paper, we have given a general definition of fairness, Generalized Weighted Fairness or GWF. In Mobile WiMAX networks, this definition provides a compromise between *slot fairness* or *temporal fairness* in which case the

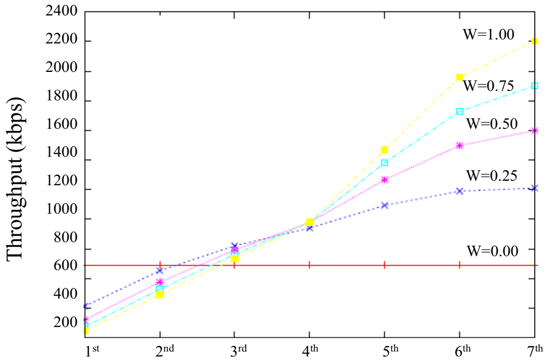


Figure 5. Analysis of throughput for GWF (weight is varied from 0 to 1)

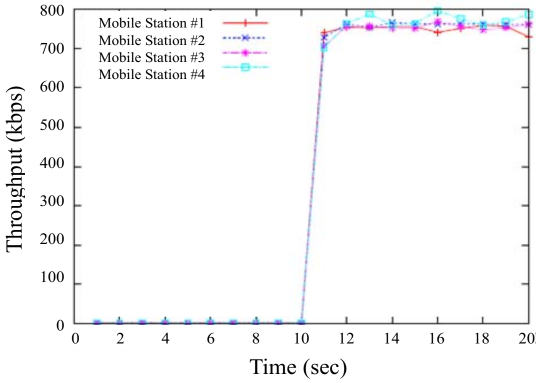


Figure 7. Four-flows throughput for DRRF with GWF, weight is 0.

service provider allocates slots fairly but throughput of the users varies with their locations; and *byte fairness* or *throughput fairness*, in which all users get the same throughput but the service provider has to allocate extra resources for users in poor channel condition. The Generalized Weighted Fairness allows the service providers to set the weight parameter to any value between 0 and 1 and achieve any level of compromise between these two extremes.

In this analysis, we do not consider the overhead such as MAC header and other subheaders. The results from our analysis can be used as an upper bound of the system goodput, the system throughput after overheads. However, we show the simulation results by taking overheads into account. The DRRF was chosen as a scheduling algorithm with the per frame transmission opportunity calculated by the GWF. The results show the throughput is similar to that obtained numerically.

REFERENCES

[1] IEEE P802.16Rev2/D2, “DRAFT Standard for Local and metropolitan area networks,” Part 16: Air Interface for Broadband Wireless Access Systems, Dec. 2007, 2094 pp.

[2] C. So-In, R. Jain, and A. Al-Tamimi, “Scheduling in IEEE 802.16e WiMAX Networks: Key Issues and a Survey,” in IEEE Journal on Selected Areas in Comm., vol. 27, no. 2, pp. 156-171, Feb. 2009.

[3] WiMAX Forum, “WiMAX System Evaluation Methodology V2.1,” Jul. 2008, 230 pp. URL=[http://www.wimaxforum.org/technology]

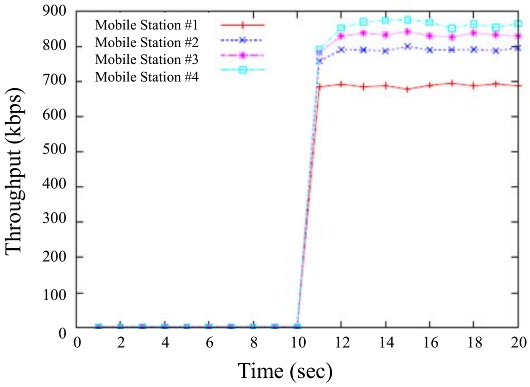


Figure 8. Four-flows throughput for DRRF with GWF, weight is 0.5.

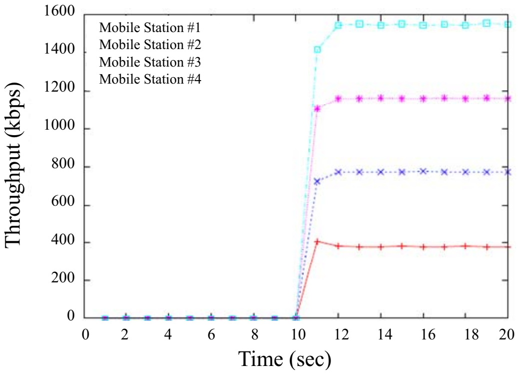


Figure 9. Four-flows throughput for DRRF with GWF, weight is 1.

[4] R. Jain, C. So-In, and A. Al-Tamimi, “System Level Modeling of IEEE 802.16e Mobile WiMAX Networks: Key Issues,” in IEEE Wireless Comm. Mag., vol. 15, no. 5, Oct. 2008.

[5] C. So-In, R. Jain, and A. Al-Tamimi, “Capacity Evaluation for IEEE 802.16e Mobile WiMAX,” To appear in Journal of Comp. Systems, Networks, and Comm. (Special Issue in WiMAX, LTE, and WiFi Internetworking), April. 2010.

[6] C. So-In, R. Jain, and A. Al-Tamimi, “A Deficit Round Robin with Fragmentation Scheduler for Mobile WiMAX,” in Proc. IEEE Sarnoff Symp., 2009, pp. 1-7.

[7] A. Sayenko, O. Alanen, and T. Hamaainen, “Scheduling solution for the IEEE 802.16 base station,” in Int. Journal of Comp. and Telecommunications Networking, vol. 52, pp. 96-115, Jan. 2008.

[8] M. Shreedhar and G. Varghese, “Efficient fair queueing using deficit round robin,” in *IEEE/ACM Trans. On Networking*, vol. 4, no. 3, pp. 375-385, Jun. 1996.

[9] UCB/LBNL/VINT, “Network Simulator – ns (version 2),” Jun. 2007. URL=[http://www.isi.edu/nsnam/ns/index.html]

[10] WiMAX Forum, “The Network Simulator NS-2 MAC+PHY Add-On for WiMAX,” Aug. 2007. URL=[http://www.wimaxforum.org]

[11] Y.P. Fallah and H. Alnuweiri, “Analysis of temporal and throughput fair scheduling in multirate WLANs,” in Computer Networks, vol. 52, no. 16, pp. 3169-3183, Nov. 2008.

---

[1] This work was sponsored in part by grant from Application Working Group of WiMAX Forum [2] “WiMAX,” “Mobile WiMAX,” “Fixed WiMAX,” “WiMAX Forum,” “WiMAX Certified,” “WiMAX Forum Certified,” the WiMAX Forum logo and the WiMAX Forum Certified logo are trademarks of the WiMAX Forum.