Courteous Algorithm: Performance optimization in WiMAX networks

Michel Kadoch, Chafika Tata
Ecole de technologie superieure

Abstract

WiMAX networks have five Quality of Service (QoS) defined to guaranty service differentiation. However, the 802.16 standard does not specify scheduling mechanisms for these networks. This article develops an algorithm named the courteous algorithm which consists of offering scheduling for WiMAX traffics and gives added services to lower classes of traffic without affecting the high priority traffic. The principles of our approach consist in servicing packets of less privileged classes and yet following the strict requirements of the more rigorous real time traffic. A mathematical analysis is performed for the scheduling system. The validation of our model has been realized by simulation. The results of the simulations show that the courteous algorithm is highly recommended in the case of a WiMAX network having a higher volume of nrtPS connections than rtPS.

Keywords: WiMAX, QoS, Scheduling, M/G/1, nrtPS, rtPS

I. Introduction

The new decade is witnessing a rapid evolution of wideband wireless networks, namely WiMAX which has been developed to offer higher throughput access to Internet to rural areas. Two versions of WiMAX have been standardized. There is the IEEE.802.16d for fixed WiMAX networks and the IEEE802.16e for the mobile WiMAX network.

These networks are interconnected following two topologies which are: Point-Multipoint (PMP) topology or mesh network topology. Contrary to the fist topology, the mesh network allows each of its stations to use other stations to forward information instead of always connecting to a base station. In the PMP mode, WiMAX subscriber stations (SS) always communicate with each other by transiting through the base station. Each SS must then establish a connection with the BS. Each connection is unique and is characterized by an identifier known as the connection identifier (CID). Packets transmitted in an uplink are put in a queue and each flow of data is queued in a specific queue corresponding to its class of service.

The IEEE802.16d standard defines four classes of service. The UGS (Unsolicited Grant Service) class supports real time traffic with fixed packet size generated in fixed regular time intervals. The rtPS (Real-Time Polling Service) also supports real time traffic but with packets of variable size generated in fixed regular time intervals. While nrtPS (Non Real-Time Polling Service) deals with traffic that tolerate delays, and generates packets of variable sizes in variable time intervals. The last class, BE (Best Effort) corresponds to traffic that demands no QoS (Quality of Service). The standard has added a fifth class, namely ertPS (Extended Real-Time Polling Service) which is similar to UGS but with variable packet sizes as in rtPS.

Although the IEEE802.16 standard defines the classes of service that highlights the differentiation of traffic types according to priority, it does not describe a scheduling system for uplink and downlink connections. Many solutions have been proposed for priority traffic management in WiMAX networks. The measures proposed in the literature bring some improvement in the management of different type of traffic in WiMAX networks. Their main objective is to conceive models that will reduce waiting delays and loss of packets, and will increase the throughput. Many scheduling algorithms have been used to that effect. There is the WFQ algorithm also with some modifications [10,4,5], has been examined. The throughput allocated to different traffic classes is implemented in a way as to assure service equity for all classes. However it may happen that the allocated resources are not fully utilised. This is an under utilisation of the throughput. Other variations of WFQ are used such as IWFQ [7] or WRR [1,7,12]. Limits of these algorithms are similar to those of WFQ. There is also the EDF algorithm used essentially for the management of queues by [8,12]. It treats with higher priority packets that have the nearest deadline of being destroyed. This algorithm is however not optimal in the non pre-emptive case [2]. There are of course the RR and FQ systems, although they try to guaranty service equity, neglect the fact that some traffics need to be processed in priority since their constraints are with respect to throughput, and packet loss. Other systems [6,9,11] have been proposed in order to compensate for the drawbacks in the above solutions. Nonetheless, most of them favour higher priority traffic.

However, these solutions mostly favour the optimization of the QoS for traffic with high priority class. Even if they allow occasionally some privileges to lower priority classes, the performance of these latter classes remain insufficient, particularly longer delays and consequently loss of packets, mostly due to the fact that the bandwidth is monopolized by higher priority classes.

The objective of this contribution is to reduce the delay and loss of packets of less favoured traffic without affecting the QoS of higher priority classes. We propose a new approach for WiMAX network traffic scheduling. This solution assures the performance optimization in IEEE802.16 networks by allowing packets of less privileged classes to be served within a certain priority. This would be possible if there is sufficient time for the scheduler to treat higher priority class traffic.

Our solution which we name the courteous algorithm, deals with the management of two types of traffic namely rtPS and nrtPS which respectively relate to VoIP and FTP services. However, it is possible to extend our analysis to all traffic classes by considering *k* calls in the network and the arrival of *m* new calls of various types (UGS, rtPS, nrtPS or BE). This can also be applied to mobile WiMAX by considering the service classes corresponding to the different handoff types.

In order to analyse the courteous algorithm, a mathematical model is developed for the different WiMAX network traffic scheduling.

The remaining of this article is organized as follows. Section II lists application conditions of the courteous algorithm. Section III presents a mathematical analysis of the two queues in our scheduling system and the possible extension to multiple queues. Section IV develops the courteous algorithm. The simulation results and its analysis of our work are shown in Section V and section Vi conclude this article.

II. Application Conditions of the courteous algorithm

The WiMAX network assigns a service priority to each packet being queued. The packets are then served according to their priority starting with the higher one. Each type of traffic is assigned its proper queue, each with a FIFO discipline. These queues could be handled with PQ (Priority Queue) or WFQ (Weighted Fair Queue) discipline so that the higher priority class traffic is processed expressively. In our case we apply instead a new discipline, a courteous mechanism that allows nrtPS class packets to be served before rtPS class packets provided that there is sufficient time to serve rtPS class packets without affecting their traffic conditions.

We are implementing this solution for uplinks in the WiMAX base station. This could also be done within relay stations or even in a subscriber station that will act as an access point to another WLAN. The application conditions of the courteous algorithm will have the effect of extending the waiting time of the rtPS packets without affecting its QoS. In order to simplify our study, we will start with the analysis of the management of two queues C_1 and C_2 that relate respectively to two classes of service namely rtPS and nrtPS. Packets of the C_1 class have priority Pr_1, while those of class C_2 have priority Pr_2. The four following conditions must be satisfied in order for packets of class C_2 be served before those of class C_1.

Condition 1

The first condition is that priority of the first queue is higher than that of the second queue or:

$$Pr1 \; > Pr2 \qquad (1)$$

Condition 2

The courteous class is the one that gives up the service in favour of the lower class. This however must not affect its QoS. In other words, the packet loss rate for the courteous class should not be beyond ω_1 which represents the tolerated threshold of packet loss rate for class C_1 traffic. The packet loss probability at time t' for the traffic of class C_1 is η_1.

The time value t' determines the end of the allocated time to the lower class. It is represented by:

$$t' = t + \tau_2 \qquad (2)$$

Where t is the initial execution time and τ_2 is the service time given to the lower class which is also the courteous time. The value of τ_2 is calculated as follows:

$$\tau_2 = \gamma * \mu \qquad (3)$$

Where γ is the number of class C_2 packets that benefit from the courteous time and μ is the average service time. We note that $\eta_1(t') < \omega_1$. This can then be represented as $\eta_1(t' + \tau) < \omega_1$.

Condition 3

This condition relates to the probability of packet loss for class C_2 namely η_2 at time t just before the application of the courteous algorithm. η_2 is the factor that determines if class C_2 traffic needs more bandwidth. In fact, if η_2 is greater than ω_2, the tolerated packet loss for class C_2, then this class would require to be served.

$$\eta_2\ (t) > \omega_2 \qquad (4)$$

Condition 4

This condition is essential for deciding if enough time is available to service class C_2 packets. The time τ_2 required to service class C_2 packets should not exceed the tolerated waiting time ξ_1 of packets of class C_1 giving up their service time.

$$\tau_2 < \xi_1 \qquad (5)$$

If this condition is not satisfied the courteous algorithm cannot be applied for otherwise the rate of packet loss for class C_1 will increase and thus η_1 will exceed which violates condition 2 above.

III. Mathematical Analysis

A. Case of two queues

Figure 1 represents the queuing system considered in our study. It is an M/G/1 queuing discipline with non preemptive priority. There are two queues, namely one for rtPS and the other for nrtPS and a single server. The rtPS queue has a maximum size of K_1. It contains the C_1 class packets. The nrtPS queue has a maximum size K_2. and contains packets of class C_2. Packet arrivals for both queues follow a Poisson process with a mean arrival rate of λ_1 packets/second for class C_1 packets and λ_2 packets/second for class C_2 packets. The total mean arrival rate λ is the sum of both λ_1 and λ_2. The service time 1/μ packets/second is exponentially distributed. The interarrival time for both cases $(1/\lambda_1$ and $1/\lambda_2)$ has an exponential distribution. The mean queue size of both rtPS and nrtPS are respectively L_{q1} and L_{q2} and have both a geometric distribution. The mean number of packets in both queues is respectively L_1 and L_2. The threshold value for each queue, namely Th_1 and Th_2, relates to the rate the queue is filling up and to the packet loss ω_1 and ω_2. A burst has a maximum of R_1 packets for class C_1 rtPS traffic.

When applying the courteous algorithm, the higher priority queue gives up its turn for service for the benefit of the lower priority queue until its queue fills up to the threshold value of Th_1. The expected number of packets for reaching this threshold is γ which is also the number of class C_2 packets that benefit from the courteous time. The waiting time in the queues for both classes is respectively W_{q1} and W_{q2}.

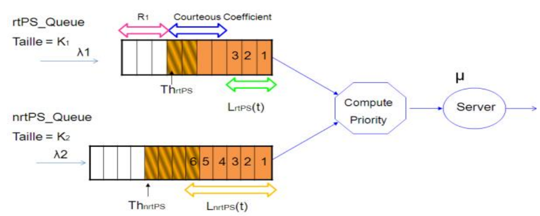

Figure 1. M/G/1 queuing system for the courteous algorithm

Figure 2 represents the state diagram of the M/G/1 model with Priority Queuing. It is the basic model that will be considered in our solution. A state $P_{m,n}$ represents a system with m rtPS packets, n nrtPS packets. While $P_{m+1,n}$ represents a system with $m+1$ rtPS packets and n nrtPS packets.

Arrival λ_1 increments the number of packets in rtPS queue. This would effect a change of state from $P_{i,j}$ to $P_{i+1,j}$. Likewise, a new arrival λ_2 which would increment the nrtPS queue would change the state from $P_{i,j}$ to $P_{i,j+1}$. Once a packet is served, the state changes by retracting to an earlier stage. For example, servicing a priority packet will cause a change in state from $P_{i,j}$ to $P_{i-1,j}$ (if the system permits it); while the service of an nrtPS lower priority packet will move states from $P_{i,j}$ to $P_{i,j-1}$ (again if the move is legal).

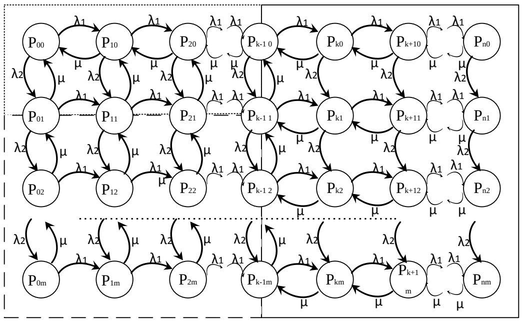

Figure 2 . State transition diagram

The set of formulas, as developed in [3] would be very relevant to our system. The following formulas have been derived from our state diagram in figure 2.

$$L_1 = \left(\frac{\lambda_1}{\mu}\right)\left(1+\rho-\frac{\lambda_1}{\mu}\right) \Big/ \left(1-\frac{\lambda_1}{\mu}\right) \quad (6)$$

$$L_{q1} = \left(\rho\frac{\lambda_1}{\mu}\right) \Big/ \left(1-\frac{\lambda_1}{\mu}\right) \quad (7)$$

$$W_{q1} = \rho / \mu - \lambda_1 \quad (8)$$

$$L_2 = \left(\frac{\lambda_2}{\mu}\right)\left(1+\rho\frac{\lambda_1}{\mu}-\frac{\lambda_1}{\mu}\right) \Big/ (1-\rho)\left(1-\frac{\lambda_1}{\mu}\right) \quad (9)$$

$$L_{q2} = \left(\rho\frac{\lambda_2}{\mu}\right) \Big/ (1-\rho)\left(1-\frac{\lambda_1}{\mu}\right) \quad (10)$$

$$W_{q2} = \rho / (1-\rho)(\mu-\lambda_1) \quad (11)$$

We would further need to compute the courteous coefficient γ and the tolerated waiting time ξ_1 of packets of class C_1.

When Courteous Algorithm is applied, the mean packet waiting time for class rtPS will increase by ξ_1. On the other hand, nrtPS packets will benefit of an equivalent mean waiting time.

The additional waiting time for rtPS will increase the mean queue length and thus the mean number of rtPS packets in the system. This additional waiting value corresponds to the Courteous coefficient γ. Since the average service time is the same for both queues, the number of nrtPS that benefit from this solution is the same as that of rtPS that give up their service. Therefore, the mean queue length for nrtPS and the mean number of packets of this class will be reduced by γ. We can then consider the following system of equations (12).

$$\begin{aligned} L_{rtPS} &= L_1 + \gamma \\ Lq_{rtPS} &= L_{q1} + \gamma \\ Wq_{rtPS} &= W_{q1} + \xi_1 \\ L_{nrtPS} &= L_2 - \gamma \\ Lq_{nrtPS} &= L_{q2} - \gamma \\ Wq_{nrtPS} &= W_{q2} - \xi_1 \end{aligned} \quad (12)$$

Computation of ξ_1 and γ

The following two equations compute ξ_1 and γ.

$$\gamma = (K_1 - ((\lambda_1 + \sigma_{\lambda_1}) * (W_{q1} + \sigma_{W_{q1}}))) - L_{q1} \quad (13)$$

Where σ_{λ_1} being the variance of λ_1 and $\sigma_{W_{q1}}$ the variance of W_{q1}.

We can then specify:

$$\xi_1 = \gamma * \mu \quad (14)$$

Therefore

$$\xi_1 = ((K_1 - ((\lambda_1 + \sigma_{\lambda_1}) * (W_{q1} + \sigma_{W_{q1}}))) - L_{q1}) * \mu \quad (15)$$

Computation of priority

Our queuing model serves the arriving packets in order of priorities. Before serving a queue, the system computes the packet priority of each head of queue and services the one that scores the highest priority.

Packets from class C_1 have priority Pr_1 and packets from class C_2 have priority Pr_2. In this M/G/1 system of queues there are, at all time t, i packets of priority Pr_1 (rtPS queue) and j packets of priority Pr_2 (nrtPS queue) and one packet of priority Pr_{Max} at the server suc as:

$$Pr_{Max} = Max\ (Pr_1, Pr_2) \quad (16)$$

In order to compute Pr_{Max}, it is necessary to determine the values of Pr_1 and Pr_2.

We assume that:

$$0 \le Pr_k \le 1, such\ as\ k = \{1,2\} \quad (17)$$

Given β_1 and β_2, the weight relative to rtPS_Queue and nrtPS_Queue that are reflecting the percentage of bandwidth allocated to classes C_1 and C_2 respectively. Note:

$$\Sigma\beta i \le 1 \mid i = \{1,2\} \quad (18)$$

$$0 \le \beta_2 < \beta_1 \le 1$$

Also, W_{s1} and W_{s2} which are the waiting time in the system of classe C_1 and class C_2 Packets respectively such that :

$$W_{sk} = Wq_k + \mu \mid k = \{rtPS, nrtPS\} \quad (19)$$

Given ψ, the rate of congestion in the transmission channel.

Priority of packet P_k depends on a number of factors, namely :

- The packet class. The normal objective is to give a higher priority to rtPS traffic packets which is feasible based on formula (18).
- The packet system waiting time W_{s1} for the one belonging to class C_1, and W_{s2} for class C_2. Furthermore, packets will increase in priority the longer they wait. This will prevent the loss of packets.
- The rate of congestion ψ in the transmission channel which will have the effect to increase the packet priority as the congestion increases.

These three factors are considered in the computation of priorities as follows:

$$Pr_1 = \begin{cases} \beta_1 + (W_{s1} /(W_{s1} + W_{s2}) + (L_1 + L_2)/(K_1 + K_2 + 1)) & if\ f(W_{s1}, \psi) \leq 1 - \beta_1 \\ \beta_1 & otherwise \end{cases} \quad (19)$$

$$Pr_2 = \begin{cases} \beta_2 + W_{s2} /(W_{s1} + W_{s2}) + (L_1 + L_2)/(K_1 + K_2 + 1) & if\ f(W_{s2}, \psi) \leq 1 - (\beta_1 + \beta_2)\ \ L_1 \geq Th_1 \\ Pr1 & if\ L_{q2} \geq Th_2\ \ L_{q1} < Th_1 \\ \beta_2 & if\ f(W_{s2}, \psi) > 1 - (\beta_1 + \beta_2) \end{cases} \quad (20)$$

Such that :

$$f(W_{s1}, \psi) = W_{s1} /(W_{s1} + W_{s2}) + (L_1 + L_2)/(K_1 + K_2 + 1) \quad (21)$$
$$f(W_{s2}, \psi) = W_{s2} /(W_{s1} + W_{s2}) + (L_1 + L_2)/(K_1 + K_2 + 1) \quad (22)$$

B. Case of n queues

Queuing system :

In this section we consider the queuing system composed of n queues, each, Q_i corresponds to class of service C_i such as $i = 1,2, \dots, n$.

Traffic arrival for class C_i follows a a Poisson process with mean arrival of λ_i packets/sec. Note that λ is the total mean arrival rate where $\lambda = \sum_{i=1}^{k} \lambda_i$

Service for class C_i is exponential with a mean rate of 1/μ packet/s.

The packet interarrival time for class C_i is exponential with a mean time of $1/\lambda_i$ sec.

The mean queue length for Q_i such that $i \in \{1,,n\}$, known as L_{qi} , is geometric. Similarly for L_i, the mean number of packets of class C_i in the system.

Computing ξ_{i-1} :

We assume that there are i classes of traffic where $i = \{1,2, \dots k\}$. We need to compute ξ_{i-1}, the tolerated supplementary waiting time for packets of traffic class C_{i-1} We assume that $Pr_1 > Pr_2 > .. > Pr_k$, such that $k = \{1, .. i, \dots n\}$ each Pr_k is the priority of class k.

The following equation calculates the value of ξ_{i-1}, such that $i \in \{1, .., k\}$.

$$\xi_{i-1} = coeff_{courteous\ (i-1)} * \mu \quad (23)$$

Similar to the two queues we can derive the following results :

$$R_{i-1} = \left(\lambda_{i-1} + \sigma_{\lambda_{i-1}}\right) * \left(Wq_{i-1} + \sigma_{Wq_{i-1}}\right) \quad (24)$$

$$coeff_{courteous\ (i-1)} = K_{i-1} - \left(\left(\lambda_{i-1} + \sigma_{\lambda_{i-1}}\right) * \left(Wq_{i-1} + \sigma Wqi{-}1 - Li{-}1 \right.\right. \quad (25)$$

R_{i-1} represents the maximum packets that can be contained in a burst of class C_{i-1}. It is the product of the maximum packet arrival rate of this class and the maximum waiting time of the same class in the corresponding queue.

$coeff_{courteous\ (i-1)}$ represents the number of packets that give up their service time to packets of class C_i Note that since the mean service time is the same for all packets, whatever the priority, we can consider that $coeff_{courteous\ (i-1)}$ is also the number of packets of class C_i that benefits from the courtesy.

Finally, the supplementary waiting time tolerance in the queue of class C_{i-1}. is given by the following formula :

$$\xi_{i-1} = [K_{i-1} - \left(\left(\lambda_{i-1} + \sigma_{\lambda_{i-1}}\right) * \left(Wq_{i-1} + \sigma_{Wq_{i-1}}\right)\right) - L_{i-1}] * \mu \quad (26)$$

Note that :

- K_{i-1}: is the queue size of the courteous packets.
- λ_{i-1} : is the mean arrival rate of packets of class C_{i-1}
- $\sigma_{\lambda_{i-1}}$: is the variance of λ_{i-1}
- Wq_{i-1}: is the means waiting time of packets of class C_{i-1} in their queue
- $\sigma_{Wq_{i-1}}$: is the variance of Wq_{i-1}
- μ is the mean service time of a packet. All packets have the same mean service time.

Considering the analysis made in (Gross 1998), with respect to M/G/1 queue with many non preemptive priorities, we can consider the following equations in the computation of Wq_{i-1}, the mean waiting time in queue Q_{i-1}, and its length L_{q_i-1} :

$$Wq_{i-1} = \frac{\sum_{k=1}^{i-1} \frac{\rho_{k-1}}{\mu}}{(1-\sigma_{i-1}) * (1-\sigma_i)} \quad (27)$$

$$\sigma_r = \sum_{k=1}^{r} \rho_k < 1 \quad (28)$$

$$L_{q_{i-1}} = \frac{\lambda_{i-1} \sum_{k=1}^{n} \rho_k / \mu}{(1-\sigma_{i-1})(1-\sigma_i)} \quad (29)$$

Computing the priority – Generalization to traffic of class i

In addition to the two formulas (19) and (20) that give the values of the two first top priorities, we consider the following formulas that will compute the other lower priorities.

$$Pr3 = \begin{cases} \beta_3 + W_{s3} / (W_{s1} + W_{s2} + W_{s3}) + (L_1 + L_2 + L_3)/(K_1 + K_2 + K_3 + 1) & \\ \quad if \ f(W_{s3}, \psi) \leq 1 - (\beta_1 + \beta_2 + \beta_3) \ and \ L_2 \geq Th_2 & \\ Pr2 & if \ L_3 \geq Th_3 \ and \ L_2 < Th_2 \\ \beta_3 & if \ f(W_{s3}, \psi) > 1 - (\beta_1 + \beta_2 + \beta_3) \end{cases} \tag{30}$$

⋮

$$Pri = \begin{cases} \beta_i + W_{si} / (\sum_{j=1}^{i} W_{sj}) + (\sum_{j=1}^{i} L_j)/(\sum_{j=1}^{i} K_j + 1) & \\ \quad if \ f(W_{si}, \psi) \leq 1 - ((\sum_{j=1}^{i} \beta_j)) \ et \ L_{i-1} \geq Th_{i-1} & \\ Pr_{i-1} & \text{if } L_i \geq Th_i \ \ L_{i-1} < Th_{i-1} \\ \beta_i & if \ f(W_{si}, \psi) > 1 - ((\sum_{j=1}^{i} \beta_j)) \end{cases} \tag{31}$$

IV. Structure of the Courteous Algorithm

Figure 3 shows the Courteous Algorithm as applied to a WiMAX network. The two types of traffic considered are rtPS for traffic of class C_1 and nrtPS for traffic of class C_2. The corresponding queues are respectively rtPS_Queue and nrtPS_Queue.

As was mentioned previously, ω_1 and ω_2 represent, respectively, thresholds for rate of rtPS and nrtPS packets loss tolerated in the system. These values are respectively proportional to Th_1 and Th_2. These later are queues filling thresholds for rtPS_Queue and nrtPS_Queue.

For the system illustrated in figure 1, we apply the computation of the priorities for rtPS and nrtPS packets in the queues. This is the result of the function in the algorithm "Priority Computation" shown in figure 3. This procedure indicates whose turn it is to be served, rtPS or nrtPS.

If rtPS is to be served the nit can be courteous by giving up its turn of service if the threshold Th_1 is not reached. This is indicated by the value of $coeff_{courteous}$ greater or equal to 1.

For every iteration, the number of nrtPS packets served is indicated by the counter cpt such that its value is less or equal to $coeff_{courteous}$, the expected number of packets up to the threshold Th_{rtPS} given a tolerated packet loss of ω_1. The initial value of cpt is 1.

$Nbr_{pk\_courteous}$ identifies the total number of rtPS courteous packets in a predefined time interval. It is also, by definition, equal to the number of nrtPS benefiting from this solution. This will allow us to evaluate the performance of our algorithm. Thresholds R_1, R_2, Th_1, Th_2 make it possible to compute $coeff_{courteous}$, and allow to decide if the application of our algorithm is feasible.

```
Compute_Threshold (R1, R2, Th1, Th2 )
    Nbr_pk_courteous = 0
    If Lq_rtPS ≠ 0 et Lq_nrtPS ≠ 0 then
    Priority Computation
        If Turn = 1 then
            If Lq_rtPS < Th1 then
                If Lq2 > Th2 then
                    Compute coeff_courteous
                    Cpt = 1
                    While cpt ≤ coeff_courteous
                      Serve (nrtPS_Queue)
                      cpt = cpt + 1
                      Nbr_pk_courteous=Nbr_pk_courteous+1
                    End While
                End If
                Else
                    Serve (rtPS_Queue)
            End If
            Else
                Serve (nrtPS_Queue)
        End If
    End If
```

Figure 3 Courteous Algorithm applied to WiMAX network: Two classes of service.

The following two cases indicate the conditions when rtPS will not give up its service in favour of nrtPS :

- The nrtPS packet loss probability is not important. This is shown by values below Th_2.
- The service time for nrtPS is greater than the tolerated supplementary waiting time for an rtPS packet. This is shown by $\tau_2 > \xi_1$.

The function « Priority Computation » in the algorithm, calculates the priority of rtPS and nrtPS packets. This function then returns a value « Turn » that indicates within the algorithm witch packet of the two classes C_1 or C_2 will be served first.

Figure 3 represents the function Compute_Threshold (R_1, R_2, Th_1, Th_2) used in the Courteous algorithm. It calculates values for Th_1 and Th_2 corresponding to thresholds for tolerated packet loss probabilities ω_1 and ω_2 in relation to the queues rtPS_Queue and nrtPS_Queue, respectively. Indeed, Th_1 and Th_2 will assure that the conditions for applying the courteous algorithm are satisfied.

$Fpr_1 = f(W_{s1}, \psi)$
$Fpr_2 = f(W_{s2}, \psi)$
If $Fpr1 > 1 - \beta_1$ then $Pr_1 = \beta_1$
else
$Pr_1 = \beta_1 + Fpr_1$
End If
If $Fpr2 > 1 - \beta_1 - \beta_2$ Then $Pr2 = \beta 2$
else
$Pr_2 = \beta_2 + Fpr_2$
End If
$MaxPr = Max\ (Pr_1, Pr_2)$
If $MaxPr = Pr1$ Then $Turn = 1$
else $Turn = 2$
End If

Figure 4 Compute_Priority Function.

$R_1 = (\lambda_1 + \sigma_{\lambda 1}) * (W_{q1} + \sigma_{Wq1})$
$R_2 = (\lambda_2 + \sigma_{\lambda 2}) * (W_{q2} + \sigma_{Wq2})$
$Th_1 = K_1 - R_1$
$Th_2 = K_2 - R_2$

Figure 5 Compute_Threshold Function.

This procedure also calculates values for R1 required in finding $coef\ f_{courteous}$ which represents the number of packets that can jump their turn of service in favour of class nrtPS. The wait time is the supplementary tolerated waiting time ξ_1.

IV. Simulation approach

A set of scenarios have been realized, representing a system of M/G/1 queues with non pre-emptive priorities for uplinks within a base station of a WiMAX network. The scheduling algorithms CPQ, PQ and WFQ have each been applied separately while considering the same basing parameters such as packets arrival rates for rtPs et nrtPS, Queues size and their weight, the maximum size of a voice traffic burst, the maximum waiting time for rtPS packet, etc.
Each scenario takes into consideration a maximum waiting time for voice traffic. The simulation was done with delays of 10 ms and 20 ms. This has been applied to the three models namely PQ, WFQ and CPQ.
The mean interarrival rate for rtPS is of the order of 0.0001s and that of nrtPS is equal to 0.00001s. The sample size is 10000.

Numerical results:

Table 1 summarizes the most important numerical results for the 3 algorithms CPQ, WFQ and PQ. Data in the table show that PQ is optimal for the mean waiting time in rtPS queue, as well as for the average size, while CPQ performs better for the mean waiting time for nrtPS as well as for its size. WFQ results are in between those of the two previous algorithms.
Note that the mean waiting time in the voice queue for CPQ is greater than found in the other models. This is explained by the fact that a great number of rtPS packets give their turn of service to other classes, thus increasing their waiting time but without reaching the rtPS tolerated maximum delay. This is reflected by the fact that no voice packet is lost in this case. The nrtPS packets that benefit from this courtesy is of the order of 82.49% of all arriving nrtPS packets.

Table 1 Numerical results

Results	CPQ	WFQ	PQ
rtPS (voice) mean # of packets in queue	4.964	0.13289	0.10695
nrtPS (data) mean # of packets in queue	6.2788	6.879	6.8855
rtPS (sec) mean queueing delay	**5.006e-005**	**1.3857e-006**	**1.1153e-006**
nrtPS (sec) mean queueing delay	**6.3893e-005**	**7.1723e-005**	**7.1798e-005**
# of rtPS packets arrived	900	900	900
# of nrtPS packets arrived	9102	9100	9100
# of rtPS lost packets	0	0	0
# of nrtPS lost packets	890	1115	1116
# nrtPS packets advantaged	7509		
% of packets advantaged	**82.49%**		

Graphical results

Study of rtPS queue length

Figure 6 shows the queue length for rtPS. The smallest queue is found with PQ followed by WFQ. CPQ is in fact the method where the longest queue for rtPS is found. In this latter instance it is noted that 80% of FTP packets are served instead of voice packets. It should however be noted that the increase of the queue length in CPQ for voice packets does not affect the QoS of this service since the packet loss rate is null.

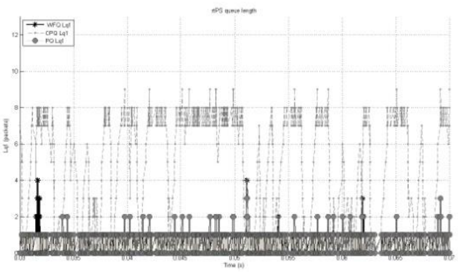

Figure 6 rtPS queue length.

Study of nrtPS queue length

Figure 7 shows that the smallest queue length for nrtPS is obtained with the CPQ approach. This is indeed the result of the courtesy made by the queues of higher priority to lower queues. WFQ is the next better approach because it gives a chance to lower priority queues to be served. PQ on the contrary only gives service to lower priority queue when higher priority queues are empty.

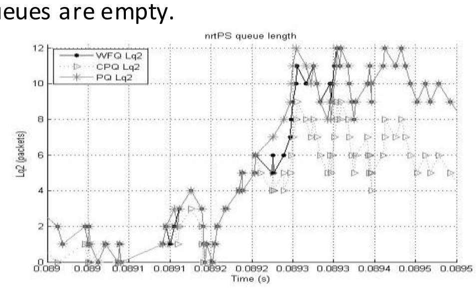

Figure 7 nrtPS queue length.

Study of the waiting time of rtPS queue

The additional waiting time in CPQ (figure 8) imposed to voice packets is important because of the increase of nrtPS packets benefiting of this courtesy. It is clear that PQ as well as WFQ have a better performance for rtPS queue for which the waiting time is very low.

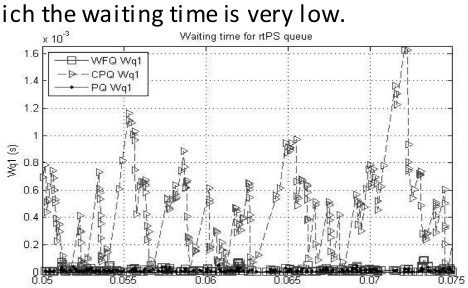

Figure 8 Waiting time for voice queue.

Study of the waiting time of nrtPS queue

Figure 9 shows that the best possible waiting time for nrtPS is possible with the CPQ model. It is no surprise to find the longest waiting time for this king of traffic within the PQ model. WFQ model is nearer to PQ.

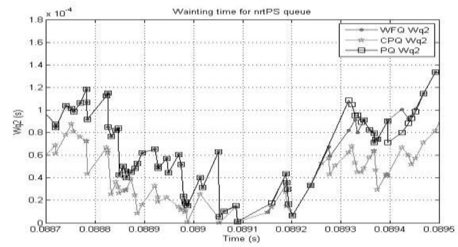

Figure 9. Waiting time for nrtPS queue.

Study of nrtPS packet loss rate

From figure 10 it is shown that the packet loss in nrtPS is smallest with the CPQ approach. The courtesy offered by higher priority queues reduces data packet loss.

The courteous algorithm thus offers the opportunity to reduce considerably the data waiting time, nrtPS queue size, and, most important, reduces the packet loss rate. This important result shows that the system will be subject to less congestion caused by data retransmission due to a high packet loss rate.

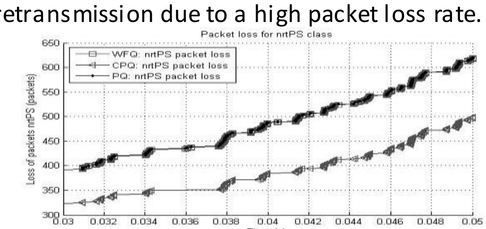

Figure 10 Packet loss for nrtPS class.

Contribution of the courteous algorithm

In order to obtain reliable and representative results, we will consider the ratio of the mean arrival rate of nrtPS traffic over the total mean arrival rate, i.e. $r_{\lambda_2} = {\lambda_2}/{(\lambda_1 + \lambda_2)}$. This parameter could indicate when the proposed algorithm would be mostly recommended and under what mean arrival rate of nrtPS traffic to apply it.

Table 2 Contribution of CPQ in terms of nrtPS traffic rate

Max delay for rtPS	10 ms	10 ms	20 ms	20 ms	20 ms
λ_1(pack/sec)	2500	250	5000	5000	10000
λ_2(pack/sec)	10000	10000	20000	40000	100000
r_{λ_2}	80 %	97.5%	80%	88.8%	90.9%
Contrib. of CPQ	7.5758%	2.1561%	13.26%	22.24%	82.49%

As shown on table 2, two standard maximum delays tolerated for voice, namely 10 ms and 20 ms have been used in our scenarios. The maximum delay for rtPS packets is in fact that of voice packets. In our scenarios these values are threshold, once exceeded packets are destroyed.

Figure 11 gives an indication of the positive implication of the proposed algorithm. The figure shows the two cases of maximum delays ($\boldsymbol{Dmax}$) of 10 ms and 20 ms for rtPS traffic.

Figure 11 indicates in the case of $\boldsymbol{Dmax}$ rtPS of 10ms, as the nrtPS arrival rate increases, the rtPS packets are becoming less courteous. This result stems from the fact that too short a delay (10 ms) does not give enough time for rtPS packets to give up some of their service time. This implies a reduction of ξ_1 representing the additional waiting time of the courteous packets. Reduction of ξ_1 also implies the reduction of $\boldsymbol{coeff_{courteous}}$ representing the number of data packets that could have been served.

As for the case of $Dmax$ rtPS of 20ms, the proposed algorithm is more favourable to an increase in of nrtPS traffic. By increasing the tolerated maximum delay for voice packets increases the value of ξ_1. As long as this value is not reached, voice packets can wait. $coeff_{courteous}$ is influenced by ξ_1.. Its value will also increase. The percentage of nrtPS packets privileged by CPQ will therefore increase.

The proposed algorithm is highly recommended when nrtPS traffic is important with respect to rtPS traffic. The rtPS traffic should have a higher tolerated maximum delay.

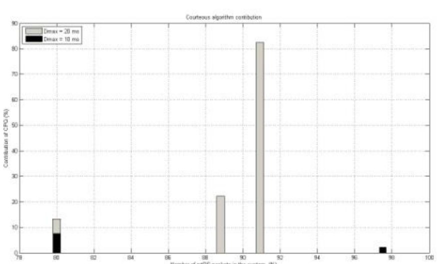

Figure 11 Courteous algorithm contribution.

V. CONCLUSION

WiMAX networks offer the possibility to expand service to isolated areas. It gives a very performant solution in terms of throughput, economic reacheability compared to wired alternative. What is required is a base station (BS) and users equipped with indoor or outdoor modems. Coverage of a BS can reach 50 Km.

QoS is however a crucial point in this network. Work has been undertaken in solving a number of issues namely mobility, admission control, scalability, scheduling and more.

Most of the research work has developed algorithms in conformity with IEEE802.16 standard, favouring rtPS traffic and giving less importance to nrtPS requirements. Real time traffic is definitely important because of its sensibility to delay. This is why most of the research work is polarized in optimizing response time to rtPS traffic.

Our proposal offers an alternative while still striving to offer the best service to rtPS traffic. It in fact shows that the service to nrtPS can substantially be improved while maintaining a high QoS standard for rtPS traffic. Our proposal improves indirectly the overall traffic since it contributes to the reduction of the packet loss rate, reduction of its retransmission, and reduction of congestion with the reduction of non real time traffic in the WiMAX environment.

Our proposal is offering a scheduling system for different traffic types in IEEE802.16 networks. Differently from existing algorithms our approach improves the scheduling of lower priority traffics while responding to the higher priority constraints. The principle in our approach consists in substituting service of packet of high priority with service to lower priority traffic whenever possible. This is in fact possible when the maximum packet loss rate of the courteous class has not been reached. Furthermore, the supplementary waiting time imposed on higher priority packets should not exceed the maximum tolerated waiting time.

Our mechanism is applicable to upstream channels to the base station. It may also be implemented in a user station used as an access point for other SS. Our study has focalized on two types of traffics namely rtPS and nrtPS. The results of our work can also be applied to the three types of traffic, rtPS, nrtPS and BE. It may also be applicable to subclasses within these established classes such as current calls, new calls and signalling calls for mobile WiMAX. This approach could also be extended to heterogeneous networks such as WiMAX/WiFi.

Consequently, scalability issue causes no problem to our approach.

A mathematical model corresponding to our system has been developed. This model allowed us to calculate mean waiting time, buffer size, additional waiting time for courteous packets, and the maximum burst size as well as packet priority. This has been validated with simulation using Matlab by comparing our results with those of the two scheduling algorithms PQ and WFQ. In some instances our results are similar to WFQ and better than PQ and in other cases it demonstrates an improved performance. It is shown that the proposed algorithm is best suited when there is an increase in nrtPS traffic. A very encouraging result shows that our approach has allowed courteous service to more than 80% of lower priority packets. This does not deter service to high priority traffic which is comparable to that of WFQ.

VI. References

[1]Belghith, A. and L. Nuaymi (2008). Comparison of WiMAX scheduling algorithms and proposals for the rtPS QoS class. EW 2008. 14th European Wireless Conference: 1-6.

[2]Ekelin, C. (2006). Clairvoyant non-preemptive EDF scheduling. 18th Euromicro Conference on Real-Time Systems. Germany 23-29.

[3]Gross, D. and C. M. Harris (1998). Fundamentals of queueing theory. New York : Wiley.

[4]Lin, J.-C., C.-L. Chou, et al. (2008). Performance Evaluation for Scheduling Algorithms in WiMAX Network. AINAW 2008. 22nd International Conference on Advanced Information Networking and Applications - Workshops 25-28 March 2008

[5]Lo, S.-C. and Y.-Y. Hong (2008). A novel QoS scheduling approach for IEEE 802.16 BWA systems ICCT 2008. 11th IEEE International Conference on Communication Technology 46-49.

[6]Prasath, G. A., C. P. Fu, et al. (2008). QoS scheduling for group mobility in WiMAXICCS 2008. 11th IEEE Singapore International Conference on Communication Systems 1663-1667.

[7]Qu, G., Z. Fang, et al. (2008). A new QoS Guarantee Strategy in IEEE802.16 MAC Layer ICPCA 2008. Third International Conference on Pervasive Computing and Applications. 1: 157 - 162

[8]Ruangchaijatupon, N., L. Wang, et al. (2006). A Study on the Performance of Scheduling Schemes for Broadband Wireless Access Networks. ISCIT '06. International Symposium on Communications and Information Technologies: 1008 - 1012.

[9]Thaliath, J., M. M. Joy, et al. (2008). Service class downlink scheduling in WiMAX COMSWARE 2008. 3rd International Conference on Communication Systems Software and Middleware and Workshops: 196 - 199

[10]Vinay, K., N. Sreenivasulu, et al. (2006). Performance evaluation of end-to-end delay by hybrid scheduling algorithm for QoS in IEEE 802.16 network IFIP International Conference on Wireless and Optical Communications Networks: 5 pp.-5

[11]Wang, Y., S. Chan, et al. (2008). Priority-Based fair Scheduling for Multimedia WiMAX Uplink Traffic. ICC '08. IEEE International Conference on Communications 301-305.

[12]Xie, X., H. Chen, et al. (2008). Simulation studies of a Fair and Effective Queueing algorithm for WiMAX resource allocation ChinaCom 2008. Third International Conference on Communications and Networking in China: 281-285.