ABSTRACT

Wireless networks today support multiple classes of service with a variety of Quality of Service requirements. This paper analyzes the impact of a novel preemptive priority based Prioritized Service resource allocation model with Queuing (PSQ) for preempted traffic. It describes this model in comparison to the traditional Complete Partitioning, Complete Sharing and Hybrid Sharing schemes and conducts a performance evaluation using both numerical techniques and simulation. It also compares two different scenarios of calls that are resumed after preemption. The results indicate the usefulness and limitations of the PSQ analytical model in comparison to the simulations.

KEY WORDS

Prioritized Service with Queuing (PSQ); Multi-server Queuing network; Complete Sharing; Wi-Fi; WiMAX networks.

1. Introduction

Ensuring quality of service in an integrated network with multiple classes of traffic continues to be a challenging research topic. With future wireless networks, it is expected that multiple sources of traffic from a variety of different networks will continue to demand limited channel resources.

Historically, a number of schemes have been proposed for resource allocation in an integrated environment [1–5]. [1] was one of the earlier works formulating a two dimensional Markov model for analyzing a mixed media cellular system. [2] studied the performance of Dynamic Partition and Dual-Threshold Bandwidth Reservation schemes. [3] analyzed a priority based call admission control (CAC) scheme in a mobile environment, while [4] analyzed the blocking probabilities in multi-rate, multi-class systems where each request has different resource requirements. [4] has also analyzed Complete Sharing, a Guaranteed Minimum as well as Prioritized Preemptive policies. [5] has studied bandwidth reservation policies in a mixed WiMAX, Wi-Fi network.

In recent years similar analysis has been done with UMTS, Wi-Fi and WiMAX networks [6–9]. A Prioritized Resource Sharing scheme for allocating channels to voice calls in WiMAX and Wi-Fi Integrated Networks has been proposed and analyzed by Andrews, Kondareddy and Agrawal [6]. In [7] a heterogeneous environment with UMTS, Wi-Fi and WiMAX has been analyzed, while [8] has modeled a prioritized vertical handoff scheme in integrated UMTS/WLANs and [9] models a prioritized channel allocation scheme in mobile wireless networks.

However in none of the papers reviewed has the issue of waiting time for calls preempted (i.e. put on hold) been considered. Our paper extends the Prioritized Resource Sharing scheme of [6] by considering not only the blocking probability criteria but also by developing a model for analyzing the queuing delays faced by calls put on hold.

Most of the schemes studied are built upon the concepts drawn from the traditional Complete Sharing, Complete Partitioning, and Hybrid Sharing schemes as described in [6]. In section 2, we briefly describe these earlier schemes, and then propose the Prioritized Service with Queuing (PSQ) scheme. In section 3 the analytical model is presented and a methodology for finding its numerical solution. Section 4 describes the simulation model of the PSQ scheme and two possible scenarios that are used for handling calls upon their resumption after preemption. Section 5 presents the simulation results achieved. Finally section 6 presents the conclusions.
2. Prioritized Service with Queuing (PSQ)

As in [6], we assume a tightly coupled architecture with an integrated WiMAX and Wi-Fi network. The WiMAX network serves as the backbone to connect Wi-Fi hotspots, in which Wi-Fi users voice calls are aggregated and communicate to the WiMAX base station via a Wi-Fi-WiMAX Bridge. Such calls are referred to as Wi-Fi calls. Those calls that emerge directly from WiMAX subscribers are referred to as WiMAX calls. Channels can be allocated to either of the two classes of traffic, which we are WiMAX calls or Wi-Fi calls on the basis of the following schemes.

Complete Sharing (CS)
The CS scheme assumes that both types of calls (WiMAX and Wi-Fi) compete equally for all the available channels on a first come first served basis.

Complete Partitioning (CP)
The CP scheme assumes that of the \( N \) available channels, \( N_M \) channels have been reserved for WiMAX calls and \( N_F \) channels have been reserved for Wi-Fi calls, with \( N_M + N_F = N \). There is no sharing of channels in this case.

Hybrid Sharing (HS)
The HS scheme assumes that of the \( N \) available channels, \( N_M \) channels have been reserved for WiMAX calls and \( N_F \) channels have been reserved for Wi-Fi calls, while \( N_S \) channels are shared on an equal priority, on a first come first served basis, with \( N_M + N_F + N_S = N \).

Prioritized Sharing (PS)
The PS scheme assumes that the \( N \) available channels, \( N_M \) channels have been prioritized for WiMAX calls and \( N_F \) channels have been prioritized for Wi-Fi calls, with \( N_M + N_F = N \). The PS scheme is fully described in [6] and hence for the sake of brevity is only briefly described here as follows:

“Prioritized means that the channels can be accessed when they are free but a user can be terminated and queued to accommodate other class of users if certain pre-defined criteria are met.”

For a complete description please refer to [6]. However in the analysis that followed in [6], although the blocking probability was adequately addressed, the issue of the terminated call being queued and the associated queuing delay was not addressed. Hence in our analysis we propose the following PSQ channel allocation scheme.

The PSQ scheme assumes that as in the PS scheme, of the \( N \) available channels, \( N_M \) channels have been prioritized for WiMAX calls and \( N_F \) channels have been prioritized for Wi-Fi calls, with \( N_M + N_F = N \) (i.e. no common equally shared channels). However, any call that is preempted due to another call with priority is then queued (i.e. put on hold). It may be noted that at any given time calls of only one particular class, (i.e. either Wi-Fi or WiMAX) will be queued. Once a free channel is available due to the departure of the prioritized call, queued calls are served based on the existing prioritized sharing allocation scheme. It may be noted that neither traffic type has any absolute priority over the other; the overall prioritization of a traffic type depends on the number of prioritized channels \( N_M \) and \( N_F \).

3. Analytical Model of PSQ

The Prioritized Service with Queuing can be modeled as a two dimensional Markov chain as depicted in Figure 1, whose state space and transition rates are described below. The model assumes that WiMAX and Wi-Fi calls are Poisson and arrive at an average rate of \( \lambda_M \) and \( \lambda_F \) calls per second. The call durations are exponential with an average duration of \( 1/\mu_M \) and \( 1/\mu_F \) seconds.

Let each state be represented uniquely by the parameters: \((i,j)\), where \((i,j)\) represents the total number of WiMAX and Wi-Fi calls present in the system respectively. However in certain states some calls will be queued due to preemption. We refer to these states as Preempted States, and the rest of the states will be represented as Normal States.

To simplify the transition rate equations, as will become evident subsequently, we introduce the following two parameters for the Preempted States: \( \{p_{ij}, q_{ij}\} \). These represent the number of queued WiMAX and Wi-Fi calls respectively, and they clearly depend on the state \((i,j)\). \( p_{ij} \) and \( q_{ij} \) can be expressed in terms of \((i,j)\) using the following functions:

\[
p_{ij} = F_p(i,j) = \begin{cases} 0 & \text{if } i + j \leq N \\ i + j - N & \text{else if } i > N_M \\ 0 & \text{otherwise} \end{cases} \tag{3.1}
\]

\[
q_{ij} = F_q(i,j) = \begin{cases} 0 & \text{if } i + j \leq N \\ i + j - N & \text{else if } j > N_F \\ 0 & \text{otherwise} \end{cases} \tag{3.2}
\]
\( (i,j)[p_{ij}, q_{ij}] \) : Normal or Preempted States;
\( (i,j) = \text{Total # of (WiMAX,Wi-Fi) Calls;} \)
\( [p_{ij}, q_{ij}] = \# \text{ of Queued (WiMAX, Wi-Fi) calls in this state} \)

Note for Normal states \( [p_{ij} = 0, q_{ij} = 0] \).

Figure 1. Markov Model Transitions

\( (i,j)[p_{ij}, q_{ij}] \) : Preempted States;
\( (i,j) = \text{Total # of (WiMAX,Wi-Fi) Calls;} \)
\( [p_{ij}, q_{ij}] = \# \text{ of Queued (WiMAX, Wi-Fi) calls in this state} \)

\( (i,j) \) : Normal States;
\( (i,j) = \text{Total # of (WiMAX,Wi-Fi) Calls} \)
(Note: here \( [p_{ij} = 0, q_{ij} = 0] \) and hence not indicated)

Total number of channels: \( N = 4 \)
\# of ch. prioritized for Wi-Fi: \( N_F = 2 \)
\# of ch. prioritized for WiMAX: \( N_M = 2 \)
\# of shared channels: \( N_S = 0 \)

Figure 2. Markov Model for \( N=4, N_F=2, N_M=2 \) case
It may be noted that the states where both \( p_{ij} = 0 \) and \( q_{ij} = 0 \) correspond to the Normal States and where either is greater than zero correspond to the Preempted States. As an example of the above, if \((i+j)\leq N\) all the channels have not been used hence the system must be in a Normal State. If \((i+j) > N\) and \(i > N_m\), then the system must be in a Preempted State, with \((i+j-N)\) WiMAX calls queued.

In order to determine the steady state equilibrium probabilities of the states we shall resort to a numerical solution based on matrix analysis. This is next described.

Let \(S\) be the possible set of states \((i, j)\) in the Markov model. The total number of states in \(S\) may be represented by \(N_T\), and expressed as a function \(N_M\) and \(N_F\) as:

\[
N_T = (N_M + N_F + 1)(N_F + 1) + (N_M + 1)N_M
\]  

(3.3)

Now let us define the following operators \(\delta_M(i, j)\), \(\delta_F(i, j)\) which are zero when further WiMAX and Wi-Fi calls are blocked, respectively:

\[
\delta_M(i, j) = \begin{cases} 
1 & \text{if } p_{i+1,j} = 0 \\
0 & \text{if } p_{i+1,j} > 0 
\end{cases}
\]  

(3.4)

\[
\delta_F(i, j) = \begin{cases} 
1 & \text{if } q_{i,j+1} = 0 \\
0 & \text{if } q_{i,j+1} > 0 
\end{cases}
\]  

(3.5)

Based on the above, it is now possible to understand the transition rates as depicted in Figure 1. As an example, consider the transition rate from state \((i, j)\) to \((i-I, j)\): as the number of WiMAX calls in service is given by \((i-p_i)\), hence the departure rate of WiMAX calls is given by \((i-p_i)\mu_M\). Similarly the transition rate from state \((i, j)\) to \((i+1, j)\) is as follows: if state \((i+1, j)\) is a Preempted state with WiMAX calls queued in it, then clearly any further WiMAX calls will be blocked, as represented by \(\delta_M(i, j) = 0\); otherwise the arrival rate is simply \(\lambda_M\).

Figure 2 shows an example of the Markov Chain for the case \(N=4, N_F=2, N_M=2, N_G=2\). We can see from (3.3) that the number of states \(N_T=21\). Of these, 6 states are preempted states, with either \(p_j\) or \(q_j\) is greater than zero.

In order to find a matrix solution to the two dimensional Markov Chain, we first map the two-dimensional state-space \((i, j)\) to a single dimension using any arbitrary function. As an example of the above we choose the mapping function, \(n = M(i, j) = \text{maps} (i, j) \text{ to } n \in 0, 1, ..., N_T - 1\) as follows:

\[
M(i, j) = \begin{cases} 
(N+1)j+i & \text{if } j \leq N_F \\
(N+1)(N_F+1) + (N_M+1)(j-N_F-1)+i & \text{if } j > N_F 
\end{cases}
\]  

(3.6)

We can now use the following sequence of steps to convert the problem into a standard rate matrix which can then subsequently be solved by matrix methods. This is outlined below.

Creating and Solving the Rate Matrix \(Q\)

**Step a:** Create rate matrix \(Q(n, m)\), based on the Markov Model of Figure 1, using the following algorithm:

**Rate Matrix Creation Algorithm:**

for \((i = 0 \text{ to } N)\)

{  
if \((i \leq N_m)\) then \(j_{\text{max}} = N\)  
extern value
else \(j_{\text{max}} = N_F\)
}

for \((j = 0 \text{ to } j_{\text{max}})\)

{  
\(n = M(i, j)\)  
\(p_j=F_s(i, j)\)  
\(q_j=F_q(i, j)\)  
}

next \(i\)

Step b: Set the diagonal entries:

\[
Q(n, n) = -\sum_{m \neq n} Q(n, m)
\]  

(3.7)

Step c: Create transpose matrix \(Q^T\) from matrix \(Q\)

Step d: Let

\[
\hat{Q}(n, m) = \begin{cases} 
Q^T(n, m) & 0 \leq m \leq N_F - 1 \\
1 & m = N_F - 1 
\end{cases}
\]  

(3.8)

Step e:

\[
b_n = \begin{cases} 
0 & 0 \leq n < N_T - 1 \\
1 & n = N_T - 1 
\end{cases}
\]  

(3.9)
Step f: Solve matrix equation: \( \hat{Q}P = b \) for the unknown elements of \( P: P_n \ e \in (0, \ldots, N_f - 1) \)

Hence \( P_n \) represents, in one-dimension, the solution to the desired steady state probabilities of the Markov Chain of Figure 1.

4. Simulation Model of PSQ

In order to verify the analytical results, simulations were conducted using OPNET modeler [10]. Figure 3a depicts the basic network topology used. It comprises of two transmitting nodes depicting a source initiating Class 0 calls and another source initiating Class 1 calls. Class 0 and Class 1 generically refer to the two classes of traffic present, using the correspondence to WiMAX and Wi-Fi traffic respectively.

Figure 3b shows the detailed model of the Hub. The key aspect of the Hub is modeled in the queuing component labeled as ‘Queue_PSQ’. This implements the PSQ model which was implemented by making the necessary changes in the multiserver queue component provided by OPNET.

Once a call is preempted and later is resumed, several possible scenarios are possible for the resumed call duration. The ‘Queue_PSQ’ implements two of these options as described below:

Preemptive Priority – Type A
The Type A scenario assumes that once a call is resumed after being preempted, the call duration for the continuing call is based on the original duration of the call that was preempted. (Any portion of the original call that was preempted is assumed will have to be repeated in both the cases.)

Preemptive Priority – Type B
The Type B scenario makes the approximation that once a call is resumed after being preempted, a new call duration is chosen based on the original call length distribution. Hence this is as a memory-less scenario where the size of the original call that was preempted is lost and not used. Due to the memory-less property of Markov Chains, the models of Figures 1 and 2 are in fact based on this assumption. Clearly this is an approximation to the more realistic scenario depicted by Type A.

The results of the analytical model versus the simulations are first shown for Type A and subsequently for Type B.

The simulations, whose results are shown in Figures 4a, 4b and 5a and 5b, assume a small system with a total of 10 channels, with 7 prioritized for WiMAX and 3 prioritized for Wi-Fi services, as summarized in Table 1. All packet sizes (\( P \)) are assumed to be of average length 250 bytes. The bandwidth (\( B \)) assumed for service of the packets is 100 kbps. Hence the packetization delay for both WiMAX and Wi-Fi calls is denoted by \((1/\mu)\) and is 20 milliseconds:

\[
1 / \mu = 8P / B = 250 \times 8 / 100 = 20 \text{ ms} \quad (4.1)
\]

The WiMAX packet arrival rate is fixed at 500 packets per second, resulting in an offered load of 10.

\[
\rho_M = \lambda_M / \mu = 500 \text{ pps} \times 20 \text{ ms} = 10 \quad (4.2)
\]

In all the simulations the Wi-Fi packet arrival rate is varied from 0 to 1000 pps, resulting in a maximum Wi-Fi offered load of 20:

\[
\rho_F = \lambda_F / \mu = 1000 \text{ pps} \times 20 \text{ ms} = 20 \quad (4.3)
\]

5. Simulation Results

Figure 4a shows the call Probability of blocking \( P(B) \) for the Wi-Fi and WiMAX services using the analytical method (‘Thry’) and simulation (‘Sim’) using Type A preemptive priority. As the Wi-Fi load saturates the network, the blocking probabilities approach the equivalent blocking probabilities of the Complete Partitioning (CP) case with 7 channels reserved for WiMAX and 3 channels for Wi-Fi. This can be verified with the call blocking probability for the CP case, given \( M \) channels:

\[
P(B)_{CP} = \left( \rho^M / M! \right) \sum_{i=0}^{M} \rho^i / i! \quad (5.1)
\]

As can be verified, \( P(B)_{CP} \) for Wi-Fi with a maximum load of 20 and with 7 channels is 86% and similarly \( P(B)_{CP} \) for WiMAX with \( \rho_M = 10 \), and \( M=7 \) channels is 41%. These correspond to the limiting values shown in Figure 4a. It may be noted here that the analytical results for Wi-Fi slightly underestimate the simulation results. This will be explained shortly.

Figure 4b shows the mean waiting time for the preempted calls that have been put on hold of both the services. Here the mean waiting time of the analytical results greatly under-estimate the simulation results for the Wi-Fi calls. This and the earlier discrepancy in the blocking probability can be explained once we see the Type B results below.

Figure 5a and 5b similarly show the analytical and simulation results for Type B preemptive priority. We now notice that the earlier discrepancy between the simulation and analytical results for Wi-Fi calls has been completely removed, and the simulation results now match with the
earlier analytical results. It is now evident why this discrepancy existed in the Type A simulations. As was explained earlier, the Type A preemption as implemented in the simulations is the more realistic form of preemption as compared to Type B. However the analytical model assumed a memory-less operation in both the cases. Hence the Type B simulation which now modeled the memory-less preemption was expected to match with the memory-less analytical model, which it did, as depicted in Figures 5a and 5b.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Channels (N)</td>
<td>10</td>
</tr>
<tr>
<td>Channels prioritized for WiMAX (N_M)</td>
<td>7</td>
</tr>
<tr>
<td>Channels prioritized for Wi-Fi (N_F)</td>
<td>3</td>
</tr>
<tr>
<td>Bandwidth per channel (B)</td>
<td>100 Kbps</td>
</tr>
<tr>
<td>Average Call duration (1/µ_M, 1/µ_F)</td>
<td>20 ms</td>
</tr>
<tr>
<td>WiMAX call arrival rate (λ_M)</td>
<td>0.5 per sec</td>
</tr>
<tr>
<td>Wi-Fi call arrival rate (λ_F)</td>
<td>Varied</td>
</tr>
</tbody>
</table>

Table 1. Network Parameters

6. Conclusion

This paper has presented a new resource allocation scheme referred to as Prioritized Service with Queuing (PSQ). A two-dimensional Markov model has been presented for evaluation of queuing delays and blocking probabilities for a network with two classes of voice traffic that share channels based on the PSQ allocation. Simulation analysis has indicated that there are several policies that can be applied to calls on their resumption after preemption. The policy used can have a significant impact on the queuing delay. One such policy has been shown to be accurately modeled by the proposed analytical model.

Results also indicate that the queuing delays for calls on hold can become significant, depending on the load on the network. However the analytical model may be used if the PSQ mechanism is implemented, to limit the offered traffic load on the network such that queuing delays for calls on hold are kept within acceptable limits. In such scenarios the PSQ mechanism may be effective in ensuring better utilization of resources as compared to other proposed earlier schemes for resource sharing.

References

Figure 3a. Network with Type 0 (WiMAX) and Type 1 (Wi-Fi) traffic.

Figure 3b. Hub with Multi-Server Queue

Figure 4a. Probability of Blocking with Preemptive Priority of Type A

Figure 4b. Mean Waiting Time with Preemptive Priority of Type A

Figure 5a. Probability of Blocking with Preemptive Priority of Type B

Figure 5b. Mean Waiting Time with Preemptive Priority of Type B