A combined resource management and admission control scheme for optimizing uplink performance of M-WiMAX systems

Georgios Theodoridis *, Fotini-Niovi Pavlidou

Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece

**A R T I C L E   I N F O**

Article history:
Received 12 August 2010
Received in revised form 15 April 2011
Accepted 10 May 2011
Available online 20 May 2011

Responsible Editor: Ing.W. Kellerer

Keywords:
802.16e
Admission control
M-WiMAX
Resource management
Uplink

**A B S T R A C T**

According to the AMC (Adaptive Modulation-Coding) mechanism of the M(obile)-WiMAX, the amount of data per modulation symbol depends on the perceived Signal-to-Noise Ratio, while, based on the OFDMA (Orthogonal Frequency Division Multiple Access) technique, each communication entails multiple concurrent transmissions. Therefore, in the uplink case, where each terminal has a maximum power limitation, the assignment of additional subcarriers does not always correspond to a proportional datarate increment, since the power dispersion over a larger set of symbols can cause the transition to a lower AMC level. In this framework, the present paper introduces a novel RRM (Radio Resource Management) algorithm that dynamically alters the number of subcarriers to be allocated to each connection so as to achieve spectrum saving via augmenting the bits per symbol ratio. Furthermore, in order to mitigate the extensive fluctuations imposed by the novel RRM scheme, an adequate minimum-complexity CAC (Connection Admission Control) scheme is proposed, which succeeds in accurately estimating the bandwidth requirements of the active flows by incorporating theoretical calculations along with a sampling process. It is proven analytically as well as through simulations that the new combined RRM-CAC strategy increases substantially the system capacity, without violating the Quality-of-Service guarantees.

© 2011 Published by Elsevier B.V.

1. Introduction

During the last decades, data communications have emerged as an essential, if not leading, factor for the majority of business, scientific and social activities. As a result, the provision of ubiquitous as well as high-speed Internet access shortly became a requisite of primary importance. Therefore, although the broadband era had been initially boosted by fiber optics and DSL (Digital Subscriber Line), the research activity soon focused on the development of a reliable and cost-efficient wireless and eventually mobile access system that could support connectivity of advanced quality beyond the limitations of the wired infrastructure. As an outcome, the IEEE 802.16 family of standards, widely known under the term WiMAX (Worldwide Interoperability for Microwave Access), was introduced in 2004 [1] as the most promising wireless solution to the last-mile problem. As a matter of fact, the general motivation behind WiMAX was to expand the ultimate success of IEEE 802.11 (Wi-Fi) over a larger coverage area and develop a system capable of providing Broadband Wireless Access (BWA) in a Metropolitan Area Network (MAN) manner.

In addition, since the early version of IEEE 802.16 supported solely fixed terminals, in 2005, the 802.16e amendment was presented, incorporating all the necessary changes and enhancements for supporting mobility [2]. Based on its specifications, M(obile)-WiMAX is expected to achieve datarates of 70 Mbps for a distance that shall range up to 15 kms, while the requested Quality-of-Service (QoS) shall be guaranteed on per connection basis. In order to maintain such a high performance against the harsh mobile conditions, M-WiMAX Physical layer utilizes a set of
the most innovative technologies, such as OFDMA (Orthogonal Frequency Division Multiple Access), Adaptive Modulation-Coding (AMC), power control and MIMO (Multiple Input/Output) techniques, which allow the service of the contemporary resource demanding Internet applications within the boundaries of the spectrum scarcity [3,4].

However, although the aforementioned mechanisms are particularly chosen so as to ensure operational stability for the whole range of potential propagation environments, this dynamic configuration of the signal transmission process introduces as a side-effect corresponding fluctuations to the bandwidth requirements of the forwarded data-flows. Moreover, in contrast to the detailed description of the Physical and MAC (Medium Access Control) layers, the exact protocols implemented at the IEEE 802.16e upper layers are not strictly defined by the standard, but they have been left open to be decided on per vendor basis, so as to facilitate the maximum possible flexibility of the aggregate network. Consequently, the optimization of the Radio Resource Management (RRM) and Connection Admission Control (CAC) routines under these disparate conditions, is regarded as a top priority, yet challenging, task.

In this framework, the cross-layer interaction between the Physical and the Data Link layer is studied thoroughly in [5,6]. Specifically, in order to incorporate a fairness factor in the RRM procedure and avoid the exclusion of the connections with harsh propagation environment, the system’s revenue is defined as an ascending and concave utility function of the connections’ datarate. Following the same principles, in [7,8] every flow’s utility function is equal to its datarate multiplied by a weighting factor, which in turn can be dependent on several parameters according to the policy of the Service Provider. Moreover, [9–12] propose an RRM algorithm that allocates the system’s resources according to the privileges of every QoS class as well as their so far performance. The latter info is acquired via monitoring the statistics of the scheduler’s queues. Alternatively, so as to realize maximum exploitation of the network resources, the authors of [13] suggest an RRM routine that allocates the available spectrum to the connections of greater datarate requirements and better radio link conditions. Similarly, [14] sets also the priorities of each flow on the basis of the achievable bit per symbol ratio, taking however into account the fact that a single transmission may present different performance among the several OFDMA subcarriers.

Especially for the case of the uplink subsystem, power saving needs also to be taken into consideration by the RRM mechanism, since the Mobile Stations (MSs) operate on battery and therefore the lower their energy consumption is the further their operating duration is prolonged. For this purpose, [15] introduces a trade-off between the amount of allocated spectrum and the transmission power. In particular, this study is based on the fact that, under conditions of bandwidth surplus, the throughput degradation, which shall be caused by a potential decrease in transmission power due to the AMC technique, can be fully compensated by the assignment of extra bandwidth. In addition, for a multi-cell M-WiMAX architecture, this decrease of the transmission power shall also lower the inter-cell interference and therefore a higher level of resource utilization is expected to be realized because of the increment of the Signal-to-Interference Noise Ratio (SINR) [16].

Nevertheless, apart from the aforementioned factors, the performance of the M-WiMAX uplink is further complicated by the functionality of the OFDMA scheme. Specifically, since every Mobile Station (MS) is bounded by a maximum transmission power, the reception SNR is reversely proportional to the number of occupied subcarriers. Thus, there is no linear relationship between the amount of allocated bandwidth and the perceived datarate, as the assignment of additional subcarriers can result in the transition to a lower AMC, i.e. lower percentage of data per symbol ratio, taking however into account the fact that a higher AMC level is taken up, then the imposed symbol-rate degradation is significantly alleviated by the augmentation of the bits per symbol ratio, while, on the other hand, as long as the same AMC level is kept, any extra subcarrier assignment is exactly analogous to the provided datarate increment. This way, not only substantial spectrum saving is achieved, but the QoS constraints of the serviced flows are also satisfied. Moreover, in order to make virtue of the enhanced RRM efficiency towards the ultimate goal of maximizing the cumulative network’s revenue, a suitable CAC algorithm is introduced. The novel CAC mechanism is especially designated for tracking down the bandwidth requirements of the active connections despite the dynamic alterations that are caused by the herein defined RRM routine. For this purpose, a dual approach, which utilizes real-time sampling along with off-line analytical computations, is implemented. As a matter of fact, such a hybrid CAC scheme manages to ensure the precise calculation of the average spectrum demands while maintaining minimum processing and signaling overhead. In consequence, under the combined RRM-CAC scheme, the system’s overall capacity is highly risen with no impact on the communications’ quality.

The rest of the paper is organized as follows: Section 2 provides an overview of the key characteristics of the IEEE 802.16e Physical layer and Section 3 studies the exact RRM procedure at the uplink subsystem. Section 4 presents the proposed RRM routine; the effectiveness of the algorithm is mathematically established, while the conditions of maximum efficiency are also theoretically determined. Furthermore, the proposed CAC algorithm is described thoroughly in Section 5. Section 6 includes the evaluation process along with the simulation results and in Section 7 the final conclusions are drawn.

2. M-WiMAX – Physical layer

2.1. OFDMA

Multiple access to the IEEE 802.16e common radio channel is achieved via the implementation of OFDMA
technique. Along the frequency axis, the available bandwidth \((W_{\text{tot}})\) is divided into numerous \((C_{\text{tot}})\) orthogonal and partially overlapping subcarriers. From these \(C_{\text{tot}}\) subcarriers only a number of \(C_{\text{dat}}\) is used for the transmission of the users’ data, while the rest are reserved for signaling purposes. In parallel, for higher resource granularity, the time-line is divided into sequential timeframes of \(T_{\text{frm}}\) duration and each timeframe is further split into \(L_{\text{slos}}\) timeslots, in a manner that is similar to the TDMA (Time Division Multiple Access) technique. Furthermore, the two-dimension combination of \(C_{\text{sch}}\) subcarriers and \(L_{\text{sch}}\) timeslots forms the minimum bandwidth allocation unit, i.e. one subchannel (Fig. 1). The grouping of subcarriers into a subchannel is called permutation and it can be carried out in two different modes, according to which one subchannel can cover either distributed (diversity permutation) or adjacent (contiguous permutation) subcarriers [4]. Finally, since each timeslot holds a single modulation symbol, the overall pool of resources to be managed is equal to the aggregate number of symbols (timeslots) per timeframe, i.e. \(C_{\text{dat}}L_{\text{slos}}\), which corresponds to the maximum available symbol-rate per Base Station (BS)

\[
SR_{\text{max}} = \frac{C_{\text{dat}}L_{\text{slos}}}{T_{\text{rms}}} \text{ (symbols/s/s)}.
\] 

(1)

Despite the frequency overlap, the subcarriers’ orthogonality allows the error-free demultiplexing of the simultaneous signals via a RFFT (Reverse Fast Fourier Transformation) procedure. Hence, besides the obvious advantage of augmenting (even doubling) the amount of exploitable bandwidth, the utilization of OFDMA also speeds up the transponders’ reconfigurability, since no zone-pass filters are any longer necessary and all demultiplexing processes are carried out via the execution of adequate software. Moreover, because of the fact that each subcarrier occupies a very narrow frequency band, every transmission is considered to be subject to flat fading [17].

2.2. AMC

Besides OFDMA, one of the pivotal technologies utilized at the Physical layer of M-WiMAX, is the AMC mechanism, which allows the system to maintain high performance under the unpredictable and usually harsh conditions of the mobile channel. Specifically, depending on the SNR level at the receiver-end, the transponder automatically adopts one the seven available combinations of modulation and coding that are predefined by the standard. At this point it must be mentioned that, for ease of reference, each one of these seven successive Modulation-Coding Levels shall be hereafter abbreviated as MCL and shall be denoted with \(MC\). Based on the aforementioned AMC functionality, if \(B(e)\) is the number of user-data bits per symbol (except for coding overhead) that are forwarded under the \(e\)th MCL \((MC = e)\), then, in order to sustain a datarate of \(DR\) bits/s, the symbol-rate must be equal to

\[
SR = \frac{DR}{B(e)}. \quad (2)
\]

Moreover, in case SNR falls below \(SN_{\text{thr}, e}\) \((MC = 0)\), no transmission takes place so as to avoid the waste of resources for a communication which is bound to fail.

As it becomes evident from (2), the system’s capacity cannot be statically foreseen, since, even for given datarate demands, the necessary resource allocation to each MS derives as a function of its SNR. The set, \(E\), of available MCLs are summarized in Table 1 along with the corresponding SNR thresholds \((SN_{\text{thr}, e})\) and data-bits per symbol \((B(e))\), \(\forall e \in E \equiv \{1,\ldots, 7\} [14]\).

3. M-WiMAX – analysis of uplink RRM

According to Section 2.1, due to the fact that the transponder transmits simultaneously over multiple orthogonal subcarriers, at each time instant the available power is shared among the number of concurrently emitted symbols. Hence, in the case of the M-WiMAX downlink subsystem where the BS acts as the transponder, the transmitted power per symbol is equal to

\[
S_{\text{tx}} = \frac{S_{\text{tx,BS}}}{C_{\text{tot}}} \quad (3)
\]

where \(S_{\text{tx,BS}}\) is the aggregate power of the Base Station. As a result, the bandwidth requirements of every connection \((SR)\) can be computed directly from (2) and Table 1 as long as the connection’s SNR and requested datarate \((DR)\) are known.

On the contrary, such a straightforward calculation cannot be realized for the uplink case, because of the power limitation that is imposed on per-terminal basis.

![Fig. 1. General structure of an OFDMA timeframe.](image)

<table>
<thead>
<tr>
<th>AMC index ((e))</th>
<th>Modulation coding</th>
<th>Bits per symbol ((B(e)))</th>
<th>Required SNR ((SN_{\text{thr}, e})) in dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1</td>
<td>BPSK &amp; 1/2</td>
<td>0.5</td>
<td>6.4</td>
</tr>
<tr>
<td>2</td>
<td>QPSK &amp; 1/2</td>
<td>1</td>
<td>9.4</td>
</tr>
<tr>
<td>3</td>
<td>QPSK &amp; 3/4</td>
<td>1.5</td>
<td>11.2</td>
</tr>
<tr>
<td>4</td>
<td>16QAM &amp; 1/2</td>
<td>2</td>
<td>16.4</td>
</tr>
<tr>
<td>5</td>
<td>16QAM &amp; 3/4</td>
<td>3</td>
<td>18.2</td>
</tr>
<tr>
<td>6</td>
<td>64QAM &amp; 2/3</td>
<td>4</td>
<td>22.7</td>
</tr>
<tr>
<td>7</td>
<td>64QAM &amp; 3/4</td>
<td>4.5</td>
<td>24.4</td>
</tr>
</tbody>
</table>
In particular, if $C_{\text{occ}}$ is the number of subcarriers occupied by an MS, then, similarly to (3)

$$S_{\text{Tx}} = \frac{S_{\text{Tx,MS}}}{C_{\text{occ}}},$$

where $S_{\text{Tx,MS}}$ is the maximum transmission power of an MS. Thus, given that

$$\langle SN\rangle_{dB} = \langle S_{\text{Tx}}\rangle_{dB} - (A_{\text{ag}})_{dB} - (N_0)_{dB},$$

the SNR at the receiver-end (BS) is computed as a function of $C_{\text{occ}}$

$$\langle SN\rangle_{dB} = \left(\frac{S_{\text{Tx,MS}}}{C_{\text{occ}}}\right)_{dB} - (A_{\text{ag}})_{dB} - (N_0)_{dB},$$

where $A_{\text{ag}}$ and $N_0$ denote the aggregate attenuation and the AWGN (Additive White Gaussian Noise) power level, respectively. For the rest of the paper all SNR expressions shall be in dBs.

Interpreting (6), the SNR is reversely proportional to $C_{\text{occ}}$ ($S_{\text{Tx}}/C_{\text{occ}}$). Therefore, in order to maximize the SNR, the subchannel assignment takes place at horizontal strips, which means that a connection is assigned successive subchannels along the same row until the end of each timeframe. This procedure is carried out repeatedly at the subsequent subchannel rows, till the target datarate is reached (Fig. 1) [4]. However, despite these precautions that are taken during the resource allocation procedure, the danger of falling to a lower MCL as additional subcarriers are overtaken still exists.

In order to give an illustrative example of the exact RRM procedure that takes place in the M-WiMAX uplink, let H be the set of all the flows that are being serviced from a single BS at time instant t. Then, $\forall h \in H$, according to (6), $SN_h$ is a function of $C_{h,\text{occ}}$. On the other hand, $SN_h$ alterations can cause changes to the bits per symbol ratio ($B(e)$) via the AMC mechanism and thus the amount of assigned resources may need to be re-estimated considering the new $B(e)$ value. Hence, $C_{h,\text{occ}}$ is also backwards related to $SN_h$ through (2). Due to this bidirectional association between $SN_h$ and $C_{h,\text{occ}}$, the analytical computation of the connections’ spectrum requirements proves to be impossible. Alternatively, an iterative routine is executed so as to compute $SN_h$ as well as the corresponding MCL for the whole set of $C_{h,\text{occ}}$ values. Eventually, h’s datarate is expressed as a function of $C_{h,\text{occ}}$ and the RRM algorithm chooses the value of $C_{h,\text{occ}}$ that satisfies the requested QoS ($DR_{h,\text{req}}$).

Being more specific, according to Section 2.1, the minimum resource allocation unit is equivalent to one subchannel, which spans across $C_{\text{sch}}$ subcarriers. Therefore, $\forall h \in H$ the domain of $C_{h,\text{occ}}$ is equivalent to the set $J$

$$C_{h,\text{occ}} \in \{0, C_{\text{sch}}, \ldots, j \ldots, C_{\text{occ}}\},$$

where j is an integer multiple of $C_{\text{sch}}$ obeying to the subchannelization rule ($\forall j \in J: j = q \cdot C_{\text{sch}}, q \in \mathbb{N}$).

So, if a connection h occupies all the timeslots of j subcarriers, the connection’s datarate shall be

$$DR_h(j) = \frac{j \cdot L_{\text{bit}} \cdot B(MC(j))}{T_{\text{frm}}} \text{ bits/s.}$$

In (8), $MC(j)$ denotes the MCL that is used ($MC \in E$, Table 1) when the number of simultaneously assigned subcarriers is equal to $j$ and $B(MC(j))$ denotes the corresponding number of forwarded user-data bits per symbol. Thus, the maximum datarate that the connection h can achieve is

$$DR_{h,\text{max}} = \max\{DR_h(j)\}, \forall j \in J.$$ \hspace{1cm} (9)

The number of subcarriers for which connection h reaches the maximum datarate $DR_{h,\text{max}}$ is denoted with $a_h \in J$. Based on (8) and (9), the datarate requirements ($DR_{h,\text{req}}$) of h can be satisfied if and only if

$$DR_{h,\text{req}} \geq DR_{h,\text{req}}$$

or equivalently if

$$\forall j \in J: DR_h(j) \geq DR_{h,\text{req}}.$$ \hspace{1cm} (11)

Under this condition, the number of subcarriers to be allocated to $h$ ($C_{h,\text{occ}} = a_h$) is computed by

$$a_h = \min\{j \in J: DR_h(j) \geq DR_{h,\text{req}}\}$$

and the necessary symbol-rate ($SR_{h,\text{req}}$) for providing the contractual QoS is calculated from

$$SR_{h,\text{req}} = \frac{DR_{h,\text{req}}}{B(MC(a_h))} \text{ symbols/s.}$$ \hspace{1cm} (13)

Correspondingly, the number of transmitted bits per timeframe is

$$D_{h,\text{req}} = DR_{h,\text{req}} \cdot T_{\text{frm}}$$

and the number of transmitted symbols per timeframe is

$$S_{h,\text{req}} = SR_{h,\text{req}} \cdot T_{\text{frm}}.$$ \hspace{1cm} (15)

Otherwise, if (11) cannot be satisfied due to either harsh propagation status or augmented datarate requirements, the transmission is carried out at the maximum achievable bitrate that is determined by (9). In this case, the symbol-rate is

$$SR_{h,\text{req}} = \frac{DR_{h,\text{req}}}{B(MC(a_h))} \text{ symbols/s.}$$ \hspace{1cm} (16)

4. Proposed RRM

4.1. Detailed description

As it becomes evident from Section 3, for the uplink sub-system, the allocation of additional resources beyond the number of timeslots that comprise a single timeframe (expansion towards additional subcarriers) does not guarantee the proportional increment of the connection’s data-rate. On the contrary, due to the SNR degradation that is induced by the dispersion of the transmission power over a larger amount of concurrent symbols, an RRM strategy which focuses solely on the pursuit of $DR_{h,\text{req}}$ may ultimately lead to an unjustifiable waste of spectrum without any substantial data-rate revenue. From this point of view, below, it is being presented an RRM algorithm which, by taking into account the alterations of $B(MC(j))$ against $j$, determines the optimal number of subcarriers to be assigned to each...
connection. This way the best possible exploitation of the scarce radio resources is guaranteed [18].

To this end, let us assume that, instead of the requested number of subcarriers \( r_h \), flow \( h \) is allocated \( i \) subcarriers \((i < r_h, i \in J)\). Then, the consequent increase (gain) in bits per timeframe is calculated by the equation

\[
GD_h(i, r_h) = (DR_h(i) - DR_h(r_h) \cdot T_{frm}, \quad i \in J.
\]

(17)

Although, because of the aforementioned interdependence between the variables \( SN_h \) and \( Ch_{out} \), the sign (positive/negative) of \( GD_h(i, r_h) \) cannot be securely predicted in advance, it is generally expected that, by decreasing the number of subcarriers from \( r_h \) to \( i \), the datarate shall be also degraded \((GD_h(i, r_h) < 0)\). Moreover, the corresponding reduction (gain) in symbols per timeframe \((GS_h(i, r_h))\) shall always be positive

\[
GS_h(i, r_h) = (SR_h \cdot IS - DR_h(i)) \cdot T_{frm}. \quad i \in J.
\]

(18)

Equivalently, based on (2) and (8), \( GD_h(i, r_h) \) and \( GS_h(i, r_h) \) can be expressed as

\[
GD_h(i, r_h) = i \cdot B(MC(i)) \cdot L_{tot} - D_{h, rq}, i \in J, \quad GS_h(i, r_h) = S_{h, rq} - i \cdot L_{tot}, i \in J.
\]

(19)

(20)

As an outcome, the problem of optimizing the number of subcarriers allocated to each connection, is reduced to choosing the appropriate value for \( i \) so as \( GS_h(i, r_h) \) to be maximized, while \( GD_h(i, r_h) \to 0 \). To perform this task, we introduce two new variables:

- \( m_h \in J \). It is the maximum number of subcarriers for which \( MC(m_h) = MC(r_h) + 1 \), i.e.

\[
m_h = \max\{j \in J : MC(j) = MC(r_h) + 1\}.
\]

(21)

According to its definition in (21), \( m_h \) presents two distinctive attributes:

i. The value of \( m_h \) is always lower than the number of subcarriers requested \((r_h)\) by the connection \( h \). Thus, following the subchannelization rule

\[
m_h \leq r_h - C_{sch}.
\]

(22)

ii. The \( m_h \) exists if and only if

\[
MC(r_h) < MC(C_{sch}).
\]

(23)

This means that the necessary MCL for the requested datarate to be reached \((MC(r_h))\) is lower than the MCL for the minimum subcarrier allocation \((MC(C_{sch}))\), i.e. a single subchannel row.

- \( M_h \in J \). It is the maximum number of subcarriers for which \( MC(M_h) = MC(r_h) \), i.e.

\[
M_h = \max\{j \in J : MC(j) = MC(r_h)\}.
\]

(24)

In analogy with \( m_h \), \( M_h \) has the following properties:

i. The value of \( M_h \) is greater or equal to \( r_h \)

\[
M_h \geq r_h.
\]

(25)

ii. In order \( M_h \) to be computable, the requested datarate must be reachable, which means that (11) must be satisfied.

The definition of \( m_h \) and \( M_h \) in reference to \( r_h \), as it is given by (21) and (24), is graphically illustrated in Fig. 2.

Eventually, according to the proposed algorithm, instead of assigning \( r_h \) subcarriers for the transmission to take place at the speed of \( DR_h(r_q) \), the following procedure is implemented:

- During the first timeframe, flow \( h \) is allocated \( m_h \) subcarriers (the whole set of timeslots for these \( m_h \) subcarriers), so as the user-data to be forwarded with rate

\[
DR_h(m_h) = \frac{m_h \cdot L_{tot} \cdot B(MC(m_h))}{T_{frm}} < DR_h(r_q)
\]

(26)

while the symbol transmission rate shall be

\[
SR_h(m_h) = \frac{m_h \cdot L_{tot}}{T_{frm}} < SR_h(r_q).
\]

(27)

- For the following \( k \) frames, flow \( h \) fully occupies \( M_h \) subcarriers. Hence, the datarate reaches the value of

\[
DR_h(M_h) = \frac{M_h \cdot L_{tot} \cdot B(MC(M_h))}{T_{frm}} > DR_h(r_q)
\]

(28)

and the symbol speed is

\[
SR_h(M_h) = \frac{M_h \cdot L_{tot}}{T_{frm}} > SR_h(r_q).
\]

(29)

The value of \( k \) is adequately chosen so as the provision of the agreed QoS to be restored within the time duration of the \( k + 1 \) frames, i.e. the aggregate datarate gain for the \( k + 1 \) timeframes must be equal to zero. This way, the datarate loss which is imposed during the first timeframe is fully compensated.

The notion behind this mechanism lies within the fact that the gain in symbols during the first timeframe \((GS_h(m_h, r_h))\) is higher than the corresponding loss in user-data \((GD_h(m_h, r_h))\), due to the upgrade of the bits per symbol ratio \((MC(m_h) = MC(r_h) + 1)\). Moreover, for the \( k \) subsequent timeframes, the bandwidth overallocation is proportional to the datarate augmentation, since

Fig. 2. Schematic description of \( r_h, m_h \) and \( M_h \).
MC(Mh) = MC(\(r_h\)) (B remains unaltered). Hence, it is expected that cumulatively, at the end of the \(k + 1\) frames, although the bulk of forwarded bits shall be equal to the requested amount, the number of utilized symbols shall be significantly reduced.

In the case that either \(m_h\) or \(M_h\) does not exist, the advanced RRM routine is not executed and the transmission is carried out at the speed of \(DR_{\text{req}}\) and \(DR_{\text{rate}}\) respectively, exactly as it is supposed by the basic scheme of Section 3.

For ease of reference, the proposed RRM algorithm shall be hereafter denoted as RRM\(_{ul}\) (Radio Resource Management UpLink).

### 4.2. Theoretical analysis and proof

In this subsection, the positive contribution of RRM\(_{ul}\) is proved mathematically, while its performance is also studied against different traffic conditions.

Specifically, at first, it shall be shown that, in comparison with the basic bandwidth assignment model described in Section 3, the utilization of the proposed mechanism minimizes the amount of resources that are necessary for guaranteeing the QoS requirements of the serviced flows. To do so, let us denote with \(GD_{\text{ag}}(k_h + 1)\) the aggregate gain in transmitted bits for flow \(h\) during the repetition period of our algorithm \((k_h + 1)\) frames. Then,

\[
GD_{\text{ag}}(k_h + 1) = \sum_{i=1}^{k_h+1} GD_h(i, r_h)
\]

and due to (19)

\[
GD_{\text{ag}}(k_h + 1) = (m_h \cdot B(MC(m_h)) \cdot L_{tot} - D_{h,req}) + k_h \cdot (M_h \cdot B(MC(M_h)) \cdot L_{tot} - D_{h,req}).
\]

Nevertheless, by definition, \(GD_{\text{ag}}(k_h + 1) = 0\). Thus, if we solve (31) for \(D_{h,req}\), we have that

\[
D_{h,req} = \frac{m_h \cdot B(MC(m_h)) + k_h \cdot (M_h \cdot B(MC(M_h)))}{k_h + 1}. \tag{32}
\]

On the other hand, the corresponding aggregate gain in symbols for the same interval of the \(k_h + 1\) frames is

\[
GS_{\text{ag}}(k_h + 1) = \sum_{i=1}^{k_h+1} GS_h(i, r_h)
\]

and due to (20)

\[
GS_{\text{ag}}(k_h + 1) = (S_{h,req} - m_h \cdot L_{tot}) + k_h \cdot (S_{h,req} - M_h \cdot L_{tot}). \tag{33}
\]

Moreover, using (13)–(15), \(GS_{\text{ag}}(k_h + 1)\) is further simplified to the expression

\[
GS_{\text{ag}}(k_h + 1) = (k_h + 1) \cdot \frac{D_{h,req}}{B(MC(r_h))} - (m_h + k_h \cdot M_h \cdot L_{tot}). \tag{34}
\]

Finally, substituting \(D_{h,req}\) with the value that has been calculated from (32) and taking advantage of the fact that, by definition, \(MC(M_h) = MC(r_h)\), (35) results in the equation

\[
GS_{\text{ag}}(k_h + 1) = m_h \cdot \left(\frac{B(MC(m_h))}{B(MC(r_h))} - 1\right) \cdot L_{tot}. \tag{36}
\]

Hence, according to (36), \(GS_{\text{ag}}(k + 1)\) is always greater than zero, since, by the definition of \(m_h\) in (21), \(MC(m_h) > MC(M_h) \Rightarrow B(m_h) > B(r_h)\). Therefore, taking into account the fact that each flow’s duration can be regarded as the aggregate of numerous, discrete, successive intervals of \(k + 1\) frames, we reach the conclusion that the necessary amount of bandwidth under the herein proposed RRM routine is always lower than the corresponding resources occupied under the basic scheme.

In the following step we shall pinpoint the conditions under which the effectiveness of RRM\(_{ul}\) is maximized. In detail, on the basis of the so far analysis, each time the novel algorithm is executed, the bandwidth gain is equal to \(GS_{\text{ag}}(k_h + 1)\). Thus, the average spectrum saving is dependent on two factors: (i) the output of (36) and (ii) how frequently the advanced RRM mechanism is triggered.

As far as the maximization of the first factor is concerned, (36) shows that the gain in transmitted symbols rises linearly versus \(m_h\). However, due to (21) and (8), \(m_h\) is limited by

\[
m_h \leq \frac{DR_{\text{req}} - T_{frm} \cdot L_{tot}}{MC(MC(r_h)) - C_{sch}}. \tag{37}
\]

Consequently, in order \(m_h\) to reach higher values, greater \(DR_{\text{req}}\) and minimum \(B\) is necessary. Furthermore, \(GS_{\text{ag}}(k_h + 1)\) is also affected by the \(B(MC(m_h))/B(MC(r_h))\) ratio, which is calculated from Table 1.

In parallel, the frequency under which the novel RRM scheme is triggered, coincides with the probability that both \(m_h\) and \(M_h\) are computable, i.e. the conditions described in both (23) and (11) must be satisfied.

As far as \(m_h\) is regarded, trying to quantitatively approximate the statement of (23), firstly we calculate the SNR degradation that shall be caused by the hypothetical assignment of extra subcarriers in comparison with the minimum subcarrier allocation of \(C_{sch}\). Using (6), if \(j = q \cdot C_{sch}, q \in \mathbb{N}\) is the number of subcarriers occupied by a random connection, then

\[
(SN)_{db}(j) = (SN)_{db}(C_{sch}) - 10 \cdot \log_{10}q. \tag{38}
\]

where \((SN)_{db}(j)\) and \((SN)_{db}(C_{sch})\) is the SNR when the number of subcarriers is equal to \(j\) and \(C_{sch}\) respectively. Hence, based on Table 1, for the whole set of \(SN_h\) values (apart from \(SN_{thr,4} - 3 \leq SN_h < SN_{thr,4}\) and \(SN_h \geq SN_{thr,7} + 3\), a 3 dB SNR decrease caused by the potential provision of one extra subchannel row \((q = 2)\) shall result in falling to the immediate lower AMC level. Therefore, for this vast majority of cases \(\forall SN_h \neq SN_{thr,4} - 3, SN_{thr,4} \cup SN_{thr,7} + 3, + \infty)\), (23) can be reduced to

\[
DR_{\text{req}} > DR_{\text{sch}} \iff r_h > C_{sch}. \tag{39}
\]

In this respect, Table 2 presents the maximum datarate that can be forwarded via a single subchannel row \((C_{row} = C_{sch})\), for the seven different MCLs that are supported by M-WiMAX (\(DR_{max, e} = \{1, \ldots, 7\}\)). The values are calculated using (8) for \(C_{sch} = 14\) and \(L_{tot} = 42\) [3]. Implementing (39) in conjunction with Table 2, we deduce the conclusion that, given a flow \(h \in \mathbf{H}\) with nominal rate equal
to $DR_{rq} \in [DR_{max,l}, DR_{max,e}]$, the $m_h$ limitation in (39) is not violated as long as $SN_h$ is lower than $SN_{thr,e}$. To make the rationale behind this theorem more evident, let us point out the reverse approach: if $SN_h > SN_{thr,e}$, while $DR_{rq} \in [DR_{max,l}, DR_{max,e}]$, then there is no point in further decreasing the number of occupied subcarriers ($C_{occ}$), since the datarate requirements of the connection are served even within the boundaries of a single subchannel row ($C_{sub}$).

The exactly opposite conditions must be valid in order the $M_h$ prerequisite to be satisfied. In particular, the $SN_h$ must be high enough so as $DR_{acc} > DR_{h,req}$ which is obviously more probable as lower the value of $DR_{h,req}$ becomes.

Consequently, incorporating all the above logical outcomes into one common criterion, the improvement that is introduced to the system’s performance due to the implementation of the novel RRM is expected to be maximized when

$$DR_{rq} \rightarrow DR_{max,e}, \quad \forall e \in \mathbb{E}. \quad (40)$$

This way the contradictory limitations of $m_h$ and $M_h$ are simultaneously served in the best possible manner:

- As long as $DR_{rq}$ is even marginally kept above $DR_{max,e}$, the $m_h$ prerequisite is satisfied for the whole set of terminals for which $SN_h < SN_{thr,e}$.
- The probability that $\exists M_h \in \mathbb{J}$ is still increased, since the datarate requirements are lowered to the minimum value that is allowed by the $m_h$ boundary. Of course, the higher the $DR_{max,e}$ limit rises, the broader becomes the range of the terminals that are excluded from the execution of $RRMUL$ due to the $M_h$ prerequisite.

5. Proposed CAC

The new RRM algorithm that was proposed in the previous section manages to guarantee the resource saving on per connection basis. Nevertheless, in order to exploit the advanced performance of $RRMUL$ towards the maximization of the overall system’s capacity, it is necessary that an adequate Connection Admission Control (CAC) mechanism is also implemented. The goal of the novel CAC scheme ($CAC_{UL}$) is to harmonize with the functionality of $RRMUL$ in a way that it is capable of capturing the new reduced bandwidth requirements of the active flows. As a result, $CAC_{UL}$ succeeds in increasing the number of simultaneously served connections, i.e. the cumulative bulk of data that is actually forwarded by the network at each time instant.

<table>
<thead>
<tr>
<th>AMC index ($e$)</th>
<th>Maximum datarate for $C_{acc} + C_{ch}$ (in kbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58.8</td>
</tr>
<tr>
<td>2</td>
<td>117.6</td>
</tr>
<tr>
<td>3</td>
<td>176.4</td>
</tr>
<tr>
<td>4</td>
<td>235.2</td>
</tr>
<tr>
<td>5</td>
<td>352.8</td>
</tr>
<tr>
<td>6</td>
<td>470.4</td>
</tr>
<tr>
<td>7</td>
<td>529.2</td>
</tr>
</tbody>
</table>

Based on the detailed description of Section 4.1, $RRMUL$ imposes a variable rate of data transmission even upon the connections whose datarate and channel quality are constant. Therefore, the CAC algorithm must be able to calibrate these fluctuations so as to pinpoint the actual spectrum demands (SR) of the active flows. To this aim, according to the usual measurement-based approach, the number of symbols that are transmitted per connection at each timeframe is returned as feedback from the Physical layer to the CAC mechanism. Finally, through the statistical processing of the collected samples, the average bandwidth needs of every connection are precisely calculated [19].

However, such a CAC scheme has always to deal with the issue of optimizing the configuration of the sampling procedure, i.e. (i) the history depth to be taken into account and (ii) the filtering criteria and weighting factors to be used along with the collected samples. In this framework, let $H$ be the set of active connections at the time instant $r$ of a new connection’s arrival. The distance of each connection $h \in H$ is practically divided into numerous sequential timeframes at which $h$ can lie in one of the following states:

ST1. Execution of $RRMUL$ routine. The symbols are transmitted with rate of either $SR_h(m_h)$ or $SR_h(M_h)$, as these are defined by (27) and (29), respectively.

ST2. Routine $RRMUL$ Cannot be executed, because $\not\exists m_h \in \mathbb{J}$ that satisfies (21). The symbols are transmitted with rate of $SR_{h,req}$, which is calculated by (13).

ST3. Routine $RRMUL$ Cannot be executed, because $\not\exists M_h \in \mathbb{J}$ that satisfies (24). The symbols are transmitted with rate of $SR_{h,thr}$, which is calculated by (16).

Thus, the bandwidth demands of a connection $h$ with constant datarate ($DR_{h,req}$) can present abrupt fluctuations due to three factors:

- The phenomenon of toggling between the states ST1, ST2 and ST3 is constant as well as highly stochastic, since it is dictated by the changes of MCL, which in turn follow the corresponding alterations of the wireless link’s quality ($SN_h$).
- By definition of $RRMUL$, even for measurements within the same ST1 state, significant deviations can take place, as $SR_h(m_h) \ll SR_h(M_h)$.
- The values of $SR_h(m_h), SR_h(M_h), SR_{h,req}$ and $SR_{h,thr}$ are not constant for the whole duration of a flow, but they are highly dependent upon the SNR.

As a result, in order to acquire a safe estimation about the mean aggregate spectrum requirements at the uplink subsystem, it is necessary to execute a thorough sampling of each connection’s SR. Specifically, the maintained history must be extended enough so as all the additional fluctuations introduced by the $RRMUL$ module to be alleviated. Moreover, for the same reason, besides the width of the sampling window, the sampling process ($T_{samp}$) must also be dense enough so as to cover all the sequential timeframes ($T_{samp} = T_{fms}$). Nevertheless, such a bulk of sampling info leads to intense signaling and processing overhead,
which can eventually cause the degradation of the system’s efficiency. As a solution, when a connection \(h \in H\) is in state ST1, the proposed CAC algorithm does not acquire \(h\)’s statistics through real-time measurements, but \(h\)’s SR is computed analytically.

In particular, for every data-flow \(h \in H\) that lies in state ST1, the gain in symbols for these \(k_h + 1\) timeframes can be calculated from (36). Hence, in comparison with the bandwidth occupied under the basic RRM scheme \((S_{h, r_{B}})\), the mean symbol gain per timeframe is

\[
G_{S_{h, frm}}(k_h + 1) = \frac{G_{S_{h, ag}}(k_h + 1) \cdot T_{frm}}{k_h + 1}.
\]

where \(k_h\) is in turn computed from (32) for a known value of \(D_{h, rq}\). Therefore, the average symbol-rate during the execution of RRM\(_{UL}\) (state ST1) is equal to

\[
SR_{h, RMS_{UL}} = SR_{h, rq} - \frac{G_{S_{h, frm}}(k_h + 1)}{T_{frm}}.
\]

As an outcome, \(\forall h \in H\) at time instant \(t\), \(h\)’s mean spectrum requirements \((SR_{h, av})\) are computed as the average value of the last \(N_{samp}\) Samples of its symbol transmission rate:

\[
SR_{h, av} = \frac{1}{N_{samp}} \sum_{k=1}^{N_{samp}} SR_{h, k}\text{ symbols/s}.
\]

where \(SR_{h, samp}\) is the aforementioned set of \(h\)’s last \(N_{samp}\) samples and every \(SR_{h, k} \in SR_{h, samp}: k = \{1, \ldots, N_{samp}\}\) is equal to:

\[
SR_{h, k} = \begin{cases} SR_{h, RMS_{UL}}, & \text{when } h \text{ lies in ST1,} \\ SR_{h, rq}, & \text{when } h \text{ lies in ST2,} \\ SR_{h, ac}, & \text{when } h \text{ lies in ST3.} \end{cases}
\]

Finally, the new candidate connection (cc) arriving at time instant \(t\) shall be admitted if and only if:

\[
\sum_{h=1}^{H} SR_{h, av} + SR_{cC, rq} \leq SR_{max}.
\]

where \(SR_{max}\) is defined in (1).

6. Simulation and performance evaluation

6.1. Simulation model

The performance of the RRM and CAC modules is assessed through a simulation software that has been developed by the authors in C++ and it is capable of realistically recreating the network environment as well as implementing the RRM and CAC mechanisms. The reference architecture considers the downlink channel of a single-cell M-WiMAX system, with radius of 1.5 km and minimum distance 20 m. The MSSs move randomly within the cell boundaries with a constant speed \(s\), which is uniformly distributed between 4 km/h and 50 km/h (from urban pedestrians to suburban vehicular users). Furthermore, every MS is supposed to carry out only one data connection per time instant. Regarding the emulation of the propagation environment, the overall attenuation \((A_{agg})\) of the transmitted signal derives as the sum of two separate factors [17]: (i) the attenuation due to distance \((A_{dL})\) and (ii) the shadowing \((A_{sh})\) due to temporal absence of LOS (Line Of Sight). \(A_{agg}\) is a function of the distance \(d\) (in km) between the MS and the BS

\[
(A_{dL})_{dB} = 128 + 37.6 \cdot \log_{10}d
\]

and shadowing (in dB) has been found to follow normal distribution with mean value equal to 0 dB and standard deviation of 8 dB [6]. The rest of the Physical layer parameters of the simulated IEEE 802.16e system, as these are defined in [2,3], are summarized in Table 3.

All the connections are considered to have the same data rate requirements \((DR_{req})\), while their arrival and servicing pattern follows a Poisson behavior, i.e. both the connections’ interarrival time \((I)\) and duration \((D)\) are exponentially distributed with mean values equal to \(I\) and \(D = 300\) s, respectively. Moreover, in order to be able to simulate disparate traffic conditions, we introduce \(DR_{agg,nm}\) as the average rate of the aggregate traffic that is being directed into the network, normalized to the network’s maximum capacity. In detail, let \(DR_{max}\) be the maximum data rate that can be forwarded by the system

\[
DR_{max} = \frac{C_{dat} \cdot L_{dat} \cdot B(7)}{T_{frm}} = 27.216\text{ Mbps}
\]

and let \(DR_{agg}\) be the mean value of the cumulative data rate requirements at each time instant

\[
DR_{agg} = \frac{D}{T} \cdot DR_{req}.
\]

Then, \(DR_{agg,nm}\) is defined as

\[
DR_{agg,nm} = \frac{DR_{agg}}{DR_{max}}.
\]

The mean interarrival time \(I\) (in s) is left open, so as, in conjunction with the value of \(DR_{req}\), to be used as the regulator of the traffic intensity.

As far the exact implementation is concerned, the simulation is executed on per timeframe basis, in a time-driven manner. In particular, at each timeframe instant \(t \in T\), \(T\) being the set of timeframes throughout the simulation duration, two sequential procedures take place:

i. CAC. In case of a new connection arrival, the CAC routine checks the availability of resources and accepts/rejects the admission request, according to the predefined criteria of the simulated CAC scheme (proposed or basic).

ii. RRM. The RRM process at simulation level is comprised of two distinct tasks:

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Physical layer parameters of the 802.16e uplink subsystem.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters of 802.16e uplink</td>
<td>Value</td>
</tr>
<tr>
<td>Total transmission power per MS ((T_{agg}))</td>
<td>200 mW</td>
</tr>
<tr>
<td>Total bandwidth ((BW_{agg}))</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Total number of OFDMA subcarriers ((C_{to}))</td>
<td>1024</td>
</tr>
<tr>
<td>Number of subcarriers per subchannel ((C_{sch}))</td>
<td>14</td>
</tr>
<tr>
<td>Number of timeslots per timeframe ((T_{ts}))</td>
<td>42</td>
</tr>
<tr>
<td>Number of timeslots per subchannel ((T_{s}))</td>
<td>2</td>
</tr>
<tr>
<td>Duration of a timeframe ((T_{frm}))</td>
<td>0.005 s</td>
</tr>
<tr>
<td>Gaussian noise power ((N_{0}))</td>
<td>(10^5 \cdot 17.4) mW/Hz</td>
</tr>
</tbody>
</table>
6.2. RRM evaluation

In Section 4.2 it was proved analytically that the proposed RRM algorithm guarantees spectrum saving strategy through increasing the average amount of user-data that is being forwarded per modulation symbol. However, beyond the qualitative study, in order to obtain a thorough assessment of the improvement achieved by the implementation of RRMUL, it is necessary that the decrease in bandwidth demands is also quantitatively estimated. For this purpose, the performance of RRMUL is tested against the basic RRM mechanism that was described in Section 3.

As a matter of fact, we introduce three new metrics that reflect the efficiency of RRMUL in the fields of spectrum exploitation as well as QoS provision:

- Bandwidth Gain Ratio (BGR). Let $G$ be the set of flows that have been serviced for the whole duration of the system’s study and $N_g$ be the set of timeframes that the connection $g \in G$ spans. Moreover, let $S_{pro,g}$ and $S_{bas,g}$ be the number of symbols that are assigned to connection $g$ for the $n^g$ timeframe, according to the proposed and the basic RRM technique respectively. Then, $\forall n_g \in N_g$, the percentage bandwidth gain between RRMUL and the RRM scheme of reference is equal to $\frac{S_{pro,g} - S_{bas,g}}{S_{bas,g}}$, where $S_{bas,g} = S_{g,t}$. The mean value of all the recorded values of $G_{n_g}$, $\forall g \in G$ and $\forall n_g \in N_g$ is defined as the Bandwidth Gain Ratio $\text{BGR} = \frac{1}{G} \sum_{g=1}^{G} \left[ \frac{1}{N_g} \sum_{n_g=1}^{N_g} \text{GS}_{n_g} \right]$. (49)

Obviously, the value of $\text{BGR}$ is indicative of the amount of spectrum being saved.

- Datarate Variation Ratio (DVR). Let $D_{pro,g}$ and $D_{bas,g}$ be the number of data-bits that are forwarded during the $n^g$ timeframe of connection $g$, in the case of the proposed and the basic RRM routine, respectively. Then, the percentage variation of $g$’s datarate between the two compared schemes is equal to $\frac{D_{pro,g} - D_{bas,g}}{D_{bas,g}}$, (50)

where $D_{bas,g} = D_{g,t}$. Similarly to $\text{BGR}$, DVR is calculated as the mean value of all the $V_{D_{n_g}}$ samples $\forall g \in G$ and $\forall n_g \in N_g$ $\text{DVR} = \frac{1}{G} \sum_{h=1}^{G} \left[ \frac{1}{N_g} \sum_{n_g=1}^{N_g} \text{VD}_{n_g} \right]$. (51)

The DVR represents the difference between the offered and the agreed QoS, as far as the flows’ sustainable datarate is concerned. In detail, the requested QoS is realized as long as $\text{DVR} > 0$, while, on the contrary, to avoid any overallocation phenomena, we should have $\text{DVR} < 0$. Hence, ideally, DVR must be kept as close as possible to zero ($\text{DVR} \rightarrow 0$).

- $k_{max}$. According to the proposed RRM scheme, initially each flow $g$ is allocated slightly degraded resources while for the subsequent $k$ frames extra bandwidth is assigned to $g$ until its nominal datarate is restored. Thus, the duration of the $k_{max} + 1$ frames should not exceed the connection’s delay boundaries. In this respect, for each flow $g \in G$, we define $k_{g,\text{max}}$ as the maximum number of $k$ for the whole forwarding process of $g$’s data. Statistically, the distribution of $k_{g,\text{max}} \forall g \in G$ is indicative of the algorithm’s ability to respect the connections’ QoS limitations from the latency point of view. Needless to say that $k_{max}$ performance has to be taken into account only for the transfer of real-time (delay sensitive) applications.

Based on the analysis of Section 4.2, the effectiveness of the proposed RRM algorithm is expected to be dependent upon the datarate requirements of the serviced connections. In this respect, in order to have an as objective as possible assessment of RRMUL’s efficiency, the three metrics ($\text{BGR}$, DVR and $k_{max}$) are tested against a great variety of $D_{g,t}$ values, ranging from 128 kbps to 512 kbps, i.e. $\text{D}_{g,t} = \{128, … , 512\} \text{ kbps}$.

The simulation output for $\text{BGR}$ and $\text{DVR}$ is displayed in Figs. 3 and 4 respectively. Moreover, Fig. 5 depicts the $k_{max}$ distribution. In detail, based on Fig. 3, $\text{BGR}$ values follow a characteristic jagged pattern with periodical picks. These periodical picks gradually degrade as the marginal values of $\text{D}_{g,t}$ are approached (128, 512 kbps), while, the algorithm’s efficiency ($\text{BGR}$) also decreases progressively in-between every two consecutive picks. Such a behavior is in complete compliance with the theoretical study of Section

- $\text{RRM evaluation}$

where $S_{bas,g} = S_{g,t}$. The mean value of all the recorded values of $G_{n_g}$, $\forall g \in G$ and $\forall n_g \in N_g$ is defined as the Bandwidth Gain Ratio $\text{BGR} = \frac{1}{G} \sum_{g=1}^{G} \left[ \frac{1}{N_g} \sum_{n_g=1}^{N_g} \text{GS}_{n_g} \right]$. (49)

Obviously, the value of $\text{BGR}$ is indicative of the amount of spectrum being saved.

- $\text{Datarate Variation Ratio (DVR)}$. Let $D_{pro,g}$ and $D_{bas,g}$ be the number of data-bits that are forwarded during the $n^g$ timeframe of connection $g$, in the case of the proposed and the basic RRM routine, respectively. Then, the percentage variation of $g$’s datarate between the two compared schemes is equal to $\frac{D_{pro,g} - D_{bas,g}}{D_{bas,g}}$, (50)

where $D_{bas,g} = D_{g,t}$. Similarly to $\text{BGR}$, DVR is calculated as the mean value of all the $V_{D_{n_g}}$ samples $\forall g \in G$ and $\forall n_g \in N_g$ $\text{DVR} = \frac{1}{G} \sum_{h=1}^{G} \left[ \frac{1}{N_g} \sum_{n_g=1}^{N_g} \text{VD}_{n_g} \right]$. (51)

The DVR represents the difference between the offered and the agreed QoS, as far as the flows’ sustainable datarate is concerned. In detail, the requested QoS is realized as long as $\text{DVR} > 0$, while, on the contrary, to avoid any overallocation phenomena, we should have $\text{DVR} < 0$. Hence, ideally, DVR must be kept as close as possible to zero ($\text{DVR} \rightarrow 0$).

- $k_{max}$. According to the proposed RRM scheme, initially each flow $g$ is allocated slightly degraded resources while for the subsequent $k$ frames extra bandwidth is assigned to $g$ until its nominal datarate is restored. Thus, the duration of the $k_{max} + 1$ frames should not exceed the connection’s delay boundaries. In this respect, for each flow $g \in G$, we define $k_{g,\text{max}}$ as the maximum number of $k$ for the whole forwarding process of $g$’s data. Statistically, the distribution of $k_{g,\text{max}} \forall g \in G$ is indicative of the algorithm’s ability to respect the connections’ QoS limitations from the latency point of view. Needless to say that $k_{max}$ performance has to be taken into account only for the transfer of real-time (delay sensitive) applications.

Based on the analysis of Section 4.2, the effectiveness of the proposed RRM algorithm is expected to be dependent upon the datarate requirements of the serviced connections. In this respect, in order to have an as objective as possible assessment of RRMUL’s efficiency, the three metrics ($\text{BGR}$, DVR and $k_{max}$) are tested against a great variety of $D_{g,t}$ values, ranging from 128 kbps to 512 kbps, i.e. $\text{D}_{g,t} = \{128, … , 512\} \text{ kbps}$.

The simulation output for $\text{BGR}$ and $\text{DVR}$ is displayed in Figs. 3 and 4 respectively. Moreover, Fig. 5 depicts the $k_{max}$ distribution. In detail, based on Fig. 3, $\text{BGR}$ values follow a characteristic jagged pattern with periodical picks. These periodical picks gradually degrade as the marginal values of $\text{D}_{g,t}$ are approached (128, 512 kbps), while, the algorithm’s efficiency ($\text{BGR}$) also decreases progressively in-between every two consecutive picks. Such a behavior is in complete compliance with the theoretical study of Section

- $\text{Bandwidth Gain Ratio (BGR)}$. Let $G$ be the set of flows...
4.2, according to which, $\forall h \in H$, the execution of the advanced RRM scheme is triggered under specific combinations of $D_{Rh}$ and channel quality. In particular, as it has been previously analyzed, when $D_{Rh}$ follows the directive of (40), then the set of $S_{Nh}$ values that satisfy both (23) and (11) at the same time is maximized. As a result, under such a configuration, the proposed algorithm is applicable to a larger percentage of data flows. This is exactly the behavior that causes the iterative saw-like shape of the BGR figure that fades out towards the edges.

On the other hand, as far as the impact on QoS is concerned, the DVR figure proves the excellent performance of the proposed RRM mechanism regarding the average provided datarate; each flow is allocated exactly the amount of bits per timeframe that satisfies its datarate needs ($DVR \rightarrow 0^*$). Moreover, judging from the $k_{max}$ distribution, the maximum time interval necessary for restoring the datarate balance is kept below $(k + 1) = 11$ timeframes for the vast majority of the cases, while under no circumstances does it exceed the $(k + 1) = 15$ timeframes. The corresponding time duration is equal to $15 \cdot T_{frm} = 75$ ms, which is far lower than the maximum tolerable latency even for the most sensitive real-time applications, e.g. VoIP – 150 ms [4,14].

6.3. CAC evaluation

In this section, the proposed CAC algorithm is tested against a conventional CAC mechanism $CAC_{UL}$, which follows exactly the same sampling procedure with $CAC_{UL}$ with the only difference that $CAC_{RUL}$ does not take into account the positive impact of $RRM_{UL}$ upon the RRM efficiency, i.e. for $CAC_{RUL}$ a connection can only be in state $ST2$ or $ST3$. This way, not only the enhanced performance of the novel CAC scheme is ascertained, but we can also evaluate the overall improvement that is introduced by $RRM_{UL}$ to an operational M-WiMAX system.

A well established criterion for assessing the effectiveness of a CAC algorithm is Blocking Probability ($BP$), which is defined as the sum of the blocked connections divided by the aggregate number of connections’ arrivals [20]. As it becomes apparent, the lower $BP$ is, the higher the percentage of serviced flows as well as the bulk of concurrently forwarded user-data rises. Therefore, for a given traffic load, the decrease of $BP$ practically corresponds to an increase in the number of supported subscribers, which in turn maximizes the network revenue.

Nevertheless, before proceeding with the comparative study, firstly the optimum configuration of $CAC_{UL}$ must be pinpointed. Specifically, according to Section 5, the proposed CAC algorithm is alleged to require a minimum sampling history ($N_{imp} \rightarrow 0$) due to its hybrid methodology that incorporates real-time measurements for the states $ST2$ and $ST3$ along with off-line calculations of the $RRM_{UL}$ spectrum allocations during $ST1$. In this respect, in order to verify the aforementioned analysis of the $CAC_{UL}$’s behavior and to determine the lowest possible value of $N_{imp}$, Fig. 6 displays the alterations of $BP$ as a function of $N_{imp}$ when $T_{imp} = 10 \cdot T_{frm}$. Based on these results, for the case of $CAC_{UL}$, $BP$ is rather independent from $N_{imp}$, hence the minimum value of $N_{imp} = 3$ can be chosen, so as to avoid any unnecessary signaling and processing overhead.

Moreover, Fig. 7 presents a graphical juxtaposition of $CAC_{UL}$ and $CAC_{RUL}$ for a wide range of traffic loads, while
Table 4 summarizes the exact percentage decrease of \( BP \) between the two schemes. It must be mentioned that the numerical analysis is performed for two different \( Drq \) values, since the bandwidth gain achieved by \( RRML \) has been found to vary upon the connections’ datarate requirements. Based on this output, it is proven that the ratio of serviced flows can be incremented by a factor of ca. 30%. As a matter of fact, even under the worst simulated scenario, the proposed combined RRM-CAC scheme succeeds in decreasing the system’s Blocking Probability at functional levels, i.e. for \( Drq = 384 \text{ kbps} \) and \( DR_{ag, rim} = 0.49 \), Blocking Probability is reduced from 21% (\( CAC_{UL} \)) to the operational value of 12% (\( CAC_{UL} \)).

7. Conclusions

In the present paper a combined resource management and admission control mechanism for maximizing the capacity of M-WiMAX uplink is introduced. Initially, the proposed scheme implements a novel RRM routine which takes advantage of the fact that a temporal, low scale compromise regarding each connection’s datarate can be proven to be highly resource efficient. As a result, the bandwidth (average number of modulation symbols per timeframe) that is necessary for servicing the datarate requirements of the admitted connections is significantly degraded, without however violating their contractual QoS demands. In parallel, a suitable CAC algorithm is developed so as to exploit the enhanced spectrum efficiency of the RRM mechanism. Despite the constant fluctuations of each connection’s resource requirements that are further intensified by the functionality of the advanced RRM routine, the proposed CAC scheme manages to perform a precise estimation of every flow’s bandwidth demands. In consequence, the network’s Blocking Probability is significantly narrowed down and the resource utilization is analogously increased.

Regarding the complexity of the algorithms’ implementation, it must be mentioned that the proposed dual scheme mainly makes use of the info that is already available to the system from the conventional RRM procedure. Only two extra variables are defined for each flow \( h \in H \), i.e. \( m_h \) and \( M_h \), which can be considered as an utterly non-substantial overhead, especially under the contemporary advanced capabilities of digital processing. Moreover, as far as the execution of the CAC mechanism is concerned, it has been shown that the off-line calculation of the bandwidth gain achieved by the advanced RRM scheme allows the CAC algorithm to perform optimally while diminishing the necessary sampling window.

Finally, it must be underlined the fact that, although the aforementioned RRM and CAC algorithms have been studied and optimized within the framework of the M-WiMAX architecture, no special features of this specific system have been taken into account other than the functionality of the OFDMA and AMC techniques. Hence, the basic principles of the proposed combined RRM-CAC scheme can be easily extended and applied to the whole set of mobile access networks that utilize OFDMA and AMC at their Physical layer.

Table 4
Percentage improvement of \( BP \) under the proposed CAC algorithm.

<table>
<thead>
<tr>
<th>( DR_{ag, rim} )</th>
<th>( BP ) Decrease (%) for ( DR_{ag} = 192 \text{ kbps} )</th>
<th>( BP ) Decrease (%) for ( DR_{ag} = 384 \text{ kbps} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.40</td>
<td>55.65</td>
<td>48.87</td>
</tr>
<tr>
<td>0.41</td>
<td>37.75</td>
<td>50.71</td>
</tr>
<tr>
<td>0.43</td>
<td>39.60</td>
<td>38.62</td>
</tr>
<tr>
<td>0.44</td>
<td>37.27</td>
<td>30.73</td>
</tr>
<tr>
<td>0.46</td>
<td>41.66</td>
<td>33.36</td>
</tr>
<tr>
<td>0.47</td>
<td>30.21</td>
<td>36.39</td>
</tr>
<tr>
<td>0.49</td>
<td>29.42</td>
<td>34.92</td>
</tr>
</tbody>
</table>

References


Georgios Theodoridis (getheod@auth.gr) received a Ph.D. degree in electrical and computer engineering from Aristotle University of Thessaloniki, Greece, in 2010 and a diploma in electrical and computer engineering in 2004 from the same institution. His research interests are in the field of call admission control and radio resource management in wireless terrestrial and Satellite/HAP networks. He is involved in Greek and European projects in these fields. He is a member of the Technical Chamber of Greece.

Fotini-Niovi Pavlidou (niovi@auth.gr) received a Ph.D. degree in electrical engineering from Aristotle University of Thessaloniki, Greece, in 1988 and a diploma in mechanical–electrical engineering in 1979 from the same institution. She is currently a full professor at the Department of Electrical and Computer Engineering at Aristotle University, teaching in the undergraduate and postgraduate programs in the areas of mobile communications and telecommunication networks. Her research interests are in the field of mobile and personal communications, satellite and HAP communications, multiple access systems, routing and traffic flow in networks, and QoS studies for multimedia applications over the Internet. She is involved in many national and international projects in these areas, and chairs the European COST262 Action on Spread Spectrum Techniques. She has served as a member of the TPC of many IEEE/IEE conferences. She is a permanent reviewer for many international journals. She has published about 80 papers in refereed journals and conferences. She is currently chairing the joint IEEE VTS & AESS Chapter in Greece.