OPEN ACCESS

University of BRISTOL

Sgardoni, V., Bull, D., & Nix, A. R. (2012). ARQ-Aware Scheduling and Link Adaptation for Video Transmission over Mobile Broadband Networks. Journal of Communications and Networks, 2012, [369803]. 10.1155/2012/369803

Peer reviewed version

Link to published version (if available):
10.1155/2012/369803

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms.html

Take down policy

Explore Bristol Research is a digital archive and the intention is that deposited content should not be removed. However, if you believe that this version of the work breaches copyright law please contact open-access@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline of the nature of the complaint

On receipt of your message the Open Access Team will immediately investigate your claim, make an initial judgement of the validity of the claim and, where appropriate, withdraw the item in question from public view.

Research Article

ARQ-Aware Scheduling and Link Adaptation for Video Transmission over Mobile Broadband Networks

Victoria Sgardoni,[1,2] David R. Bull,[1] and Andrew R. Nix[1]

[1] *Department of Electrical and Electronic Engineering, University of Bristol, Bristol BS8 1UB, UK*
[2] *Department of Aircraft Technology, Technological Educational Institute (ATEI) of Chalkis, Psachna, 34400 Evia, Greece*

Correspondence should be addressed to Victoria Sgardoni, victoria.sgardoni@bristol.ac.uk

This paper studies the effect of ARQ retransmissions on packet error rate, delay, and jitter at the application layer for a real-time video transmission at 1.03 Mbps over a mobile broadband network. The effect of time-correlated channel errors for various Mobile Station (MS) velocities is evaluated. In the context of mobile WiMAX, the role of the ARQ Retry Timeout parameter and the maximum number of ARQ retransmissions is taken into account. ARQ-aware and channel-aware scheduling is assumed in order to allocate adequate resources according to the level of packet error rate and the number of ARQ retransmissions required. A novel metric, namely, *goodput per frame*, is proposed as a measure of transmission efficiency. Results show that to attain quasi error free transmission and low jitter (for real-time video QoS), only QPSK 1/2 can be used at mean channel SNR values between 12 dB and 16 dB, while 16QAM 1/2 can be used below 20 dB at walking speeds. However, these modes are shown to result in low transmission efficiency, attaining, for example, a total goodput of 3 Mbps at an SNR of 14 dB, for a block lifetime of 90 ms. It is shown that ARQ retransmissions are more effective at higher MS speeds.

1. Introduction

Mobile WiMAX (IEEE 802.16e) [1] and 3GPP LTE (Long-Term Evolution) [2] represent mobile broadband standards that offer high user data rates and support for bandwidth hungry video applications. Both standards use very similar PHY and MAC layer techniques, especially for downlink (DL) transmission. In order to provide strong QoS, cross-layer adaptive strategies must be implemented in the wireless network [3, 4]. Video applications demand a low Packet Error Rate (PER), which may be achieved via the use of MAC layer Automatic Repeat ReQuest (ARQ) and the choice of suitable Modulation and Coding Schemes (MCS). However, ARQ consumes additional bandwidth and causes increased end-to-end latency and jitter. ARQ is controlled in the MAC layer by the block lifetime and ARQ Retry Timer parameters, which define how many and how frequently retransmissions may occur. Link adaptation is used in mobile broadband networks to improve the PER by matching the QAM constellation and forward error correction coding rate to the

time varying channel quality. The impact of specific ARQ parameters and mechanisms has been extensively studied in the literature, for example, [5–9].

In [8], the authors analyze delay and throughput using probabilistic PHY layer error modelling. In [9], packet errors were modelled as an uncorrelated process in time. Often packet errors are modelled using statistical channel models, such as Markov chains, for example, [3, 10, 11], based on statistical measurements that have limited scalability and adaptability to a variety of fading, shadowing, or mobility circumstances. However, this type of modelling fails to represent the bursty nature of errors in a fading channel and the impact it has on ARQ retransmission performance.

To deliver video QoS the mobile WiMAX and LTE standards specify a number of scheduling mechanisms, such as Unsolicited Grant Service (UGS), rtPS (real-time Polling Service), and BE (Best Effort). As scheduling of resources is not specified in the standards, but instead left open for vendor implementation, this is an area of considerable research interest. In [12], a survey of several scheduling

algorithms showed that, due to the nature of the wireless medium and user mobility, the scheduler should take into account the PER and the Carrier-to-Interference-plus-Noise Ratio (CINR) reported by the channel quality indicator (CQI) per connection. These schedulers are denoted as *"channel aware."* A channel-aware scheduler must take into consideration the MCS mode selected through link adaptation. The scheduling of resources must also take into account ARQ retransmissions, as discussed in [6, 8, 12, 13]. The authors of [8, 13] propose an *ARQ-aware* scheduler, where ARQ retransmissions have priority over new data. For applications that are very sensitive to jitter and delay, such as video, the QoS guarantees a maximum delay and error rate for a given bitrate. If ARQ is enabled on these connections, the BS scheduler must allocate sufficient resources in each frame to accommodate new data and ARQ retransmissions. The resources required per connection vary also according to the MCS mode selected by the Link Adaptation (LA) process.

Many recent publications have studied video streaming over WiMAX, for example, [10, 11, 14], but very few investigate unicast video with ARQ retransmission [4, 9]. In [4], the authors proposed cross-layer parameter optimization to achieve the required QoS, using queuing theory to minimize the required bandwidth while assuming stop-and-wait ARQ retransmission. None of the ARQ mechanisms specified in the 802.16e standard were considered in [4]. In [3, 14], the issue of "bandwidth hungry" video applications was highlighted. Nevertheless, the only video transmissions considered in recent publications are based on low resolution video (CIF, QVGA) with bitrates up to 400 kbps [8, 11, 14].

This work focuses on the transmission of high resolution real-time video, at a bitrate of 1.03 Mbps, over the downlink (DL) of a mobile broadband connection. The simulated transmission of a flow of UDP packets corresponds to the flow of video packets. Simulations are performed for a UDP unicast DL transmission, with Selective ACK (S-ACK) ARQ enabled. Moreover, multicast transmission without ARQ enabled is also included in the analysis. The transmission efficiency of ARQ enabled mobile WiMAX networks is computed by proposing a novel efficiency metric, the *goodput per frame*, which takes into account the amount of radio resources required per DL subframe and the PER attained. Channel-aware and ARQ-aware scheduling at the MAC layer is assumed. Very importantly, block errors are time-correlated, based on the use of the accurate time-correlated 3GPP SCM fading channel model [15]. Results are based on the WiMAX Forum recommendations [16, 17] for the ARQ Retry Timeout parameter and the maximum number of ARQ retransmissions. The study shows for the first time how PER and delay/jitter are affected by scheduling sufficient (or insufficient) channel resources per frame, to cater for ARQ retransmissions, according to the MCS mode selected. The work identifies which MCS modes are suitable to deliver QoS for real-time video, by maintaining quasi-zero PER and low jitter at the application layer. Our previous work [18] focused on received video quality (based on PSNR), for a 7.63 Mbps HD video sequence, when no limitations were applied to the ARQ Retry Timeout parameter and the maximum number of ARQ retransmissions, as assumed in

[8, 9] (since the frequency of ACK is not specified in the IEEE 802.16e standard [1]). The effect of ARQ-aware scheduling was not investigated in our previous work.

In this paper, the effect of MS velocity on ARQ retransmissions is explored. This is made possible by the use of an accurate time-correlated fading channel model. MS velocities of 1 and 10 km/h are considered.

Mobile WiMAX [1], together with 3GPP LTE [2], is key technology for next-generation broadband wireless access (BWA) networks [19]. Both technologies have very similar DL PHY layers and strong similarities in their MAC layers. For both technologies, radio resource management techniques, such as scheduling and resource allocation, are pivotal in research work on QoS support for multimedia services [19].

In the next section, key aspects of the mobile WiMAX PHY and MAC layers are described along with the time-varying channel model. In Section 3, the MAC-PHY simulator is described, detailing the assumptions made. Sections 4, 5, and 6 present an analysis of the simulation results. Conclusions are presented in Section 7.

2. Overview of Mobile WiMAX and the SCM Channel Model

Medium Access Control (MAC) Layer. The 802.16e MAC layer [20] includes a number of adjustable features, such as adaptive MCS, ARQ, packet fragmentation and aggregation, variable size MAC Protocol Data Units (PDU), and application-specific service flows and PDU scheduling based on QoS. Packets from the higher layers arrive in the convergence sublayer (CS) of the MAC as MAC Service Data Units (SDUs). Based on their QoS requirements, MAC SDUs are classified into service flows. There is the option for SDU fragmentation into PDUs, and this feature is assumed here. SDUs are partitioned into ARQ blocks of fixed size when ARQ is enabled. The MAC PDU is the data unit exchanged between the BS and MS MAC layers. Once a PDU has been constructed, it is placed in the appropriate service flow queue and managed by the scheduler, which determines the PHY resource allocation (i.e., bandwidth and OFDMA symbol allocation) on a frame-by-frame basis.

Each transmitted PDU is either received correctly or in error, depending on the channel response at the time of transmission. The time-varying PHY layer PER is accurately calculated based on the channel model for each ARQ block in the transmitted PDU. The standard specifies a number of ARQ feedback mechanisms, such as Cumulative ACK, Cumulative and Selective ACK, and Selective ACK (S-ACK) [1]. Here S-ACK feedback is used. An S-ACK feedback message is generated for each transmission burst and any PDUs containing errors are placed in the retransmission queue [6]. No block rearrangement is enabled. The retransmission of PDUs in error continues until they are received correctly or their ARQ block lifetimes expire. The number of retransmissions is determined by the block lifetime and the ARQ Retry Timeout Timer, as shown in (1). The ARQ Retry Timeout represents the minimum number of OFDMA frames a transmitter will wait to retransmit an

Table 1: OFDMA PHY profile parameters in 802.16e.

Parameters	Values			
Channel bandwidth (MHz)	1.25	5	*10*	20
FFT size	128	512	*1024*	2048
Sampling frequency (MHz)	1.4	5.6	*11.2*	22.4
Subcarrier frequency spacing (kHz)	10.94			
Useful symbol time T_b (μs)	91.4			
Guard time T_g (μs)	11.4			
OFDMA symbol duration $(T_s = T_b + T_g)$ (μs)	102.9			
Number of OFDMA symbols (5 ms frame)	47			

Table 2: 802.16e link speeds.

No.	Link speed	Bits per slot	Total data rate (Mbps)
0	QPSK 1/2	48	6.14
1	QPSK 3/4	72	9.21
2	16QAM 1/2	96	12.29
3	16QAM 3/4	144	18.43
4	64QAM 1/2	144	18.43
5	64QAM 2/3	192	24.58
6	64QAM 3/4	216	27.64

unacknowledged block [1]. This retry period begins from the frame when the ARQ block was last transmitted. If the block lifetime expires before it is received correctly then the block is discarded.

Physical Layer (PHY). The mobile WiMAX standard has adopted Scalable-OFDMA (S-OFDMA) [1]. Table 1 shows the relevant parameters for the S-OFDMA PHY. Simulations for this paper were performed for the 10 MHz channel bandwidth profile (highlighted in italics in Table 1). The payload data is modulated using the full range of link speeds (MCS modes) as defined in the standard [1] and shown in Table 2. Assuming a PUSC DL [1], the modulation symbols allocated to a sequence of slots in each DL OFDMA frame are assigned to a number of logical subchannels. An OFDMA slot is the minimum possible data allocation unit. For PUSC DL, it is defined as one subchannel by two OFDMA symbols. For the 10 MHz channel, an OFDMA symbol consists of 30 subchannels for PUSC DL, each containing 24 data subcarriers [21]. Hence, a slot contains 48 data subcarriers. Based on this, the slot payload capacity P_{sl} for each MCS mode is computed for PUSC DL. It is shown in Table 3, where m represents the MCS modulation order and r the coding rate. The channel resources (in terms of slots) required for data transmission over a mobile WiMAX network are evaluated based on the slot payload capacity for each MCS mode.

PHY Layer Abstraction. To simplify the interface between the link and system level simulators, whilst still modelling dynamic system behaviour, a technique known as Effective

SINR Mapping (ESM) is used. This method, which can also be used to model the LTE PHY layer, compresses the SINR (per subcarrier) vector into a single effective SINR (ESINR). The technique is described in detail in [22]. The PHY abstraction model is described and validated in [23]. This PHY abstraction model allows the instantaneous Block Error Rate (BLER) to be computed for each channel realization, based on the instantaneous fading channel and the length of the ARQ block. Although many commercial network simulators exist, such as OpNet and QualNet, these tend to provide simplified physical layer support. For example, QualNet uses bit error rate look up tables that average the effects of time-varying fast fading. Video analysis requires the use of time-varying instantaneous BLER in a fading channel (not averaged BLER), since the bursty nature of the errors has a detrimental effect on video quality, as shown in [24].

Wideband Channel Model. The channel model follows the ETSI 3GPP spatial channel model (SCM), as described in [15]. A time varying "urban micro" tapped delay line (TDL) was generated for each channel snapshot. The TDL consists of 6 time-correlated fading taps with nonuniform delays. The carrier frequency is 2.3 GHz and the FFT size is 1024. Each radio channel is made up of a number of channel samples (sampled every 2.5 ms) corresponding to a duration of 85 seconds.

3. WiMAX MAC-PHY Simulator

Unicast and multicast transmission of high resolution video is simulated over the mobile WiMAX system. This work is based on a MAC-PHY simulator developed according to the standard [1] and presented in [18]. The mobile WiMAX PHY layer simulator is described in [25]. As discussed in Section 2, the PHY layer PER is generated from the ESM PHY layer abstraction method developed in [23].

3.1. Simulator Assumptions. The following key assumptions were made for the design of the mobile WiMAX MAC-PHY simulator. MAC SDU fragmentation (not packing) is assumed, according to the 802.16e standard [1]. The MAC PDU size is fixed for all MCS modes. It is small (less than 200 Bytes) to improve the error rate, according to [5]. The simulated ARQ mechanism is Selective ACK (S-ACK) [1]. It is assumed that no errors occur in the ARQ feedback messages. When errors occur during PDU transmission, no block rearrangement is performed within the PDUs. PDUs that are not acknowledged are placed in a separate queue within the user flow, known as the retransmission queue [6, 26]. The scheduler gives priority to PDUs from the retransmissions queue. SDUs are delivered in order at the receiver, since the ARQ-DELIVER_IN_ORDER MAC parameter is enabled.

Video is assumed to be sent at a constant bit rate (CBR), with fixed size packets; therefore, the use of UGS scheduling is assumed. The scheduler allocates a fixed amount of resources per MAC frame for each DL burst, according to the operation of UGS scheduling. As the standard [1] does not specify how the scheduling of resources is performed,

Table 3: Slot payload capacity per MCS for PUSC DL.

MCS	m	r	Slot payload P_{sl} (bits)
QPSK 1/2	4	1/2	48
QPSK 3/4	4	3/4	72
16QAM 1/2	16	1/2	96
16QAM 3/4	16	3/4	144
64QAM 1/2	64	1/2	144
64QAM 2/3	64	2/3	192
64QAM 3/4	64	3/4	216

but instead leaves the issue open for vendor implementation, it is assumed that a "channel-aware scheduler" is used that allocates resources according to the MCS mode selected with an overallocation for ARQ retransmissions [13]. If additional resources (i.e., overallocation) are not provided, when retransmissions occur they will take up resources from the new arriving data, since retransmissions have priority. This would result in a queuing delay that is unacceptable for real-time video applications. It is assumed that adequate additional resources for ARQ are available per frame, as required to cater for the expected number of retransmissions depending on the BLER.

The data payload that each slot can carry for each MCS mode is given in Table 3. The number of PDUs that can fit within the allocated resource is calculated according to the MCS mode and the size of retransmission PDU queue. For each DL burst the retransmission PDUs are included in the allocated resources, taking priority over new PDUs.

3.2. Simulator Functionality. The 802.16e MAC-PHY simulator provides an error modelling tool that predicts the loss patterns for a sequence of RTP/UDP packets and thus the losses in the sequence of video packets at the receiver. The PHY abstraction model allows the instantaneous BLER to be computed for each channel realization, based on the instantaneous fading channel and the length of the ARQ block. Thus, the computed BLER is time-correlated. A flow of fixed size RTP/UDP packets arrive at the MAC layer and are passed into the simulator. It is assumed that UDP packets arrive at a constant rate. Each UDP packet corresponds one-to-one to a MAC SDU. More details on the MAC-PHY simulator are given in [18].

At the receiver, SDUs are reassembled from the appropriate ARQ blocks. Since the MAC parameter ARQ_DELIVER_IN_ORDER is enabled, SDUs are delivered in sequential order to the transport layer, as UDP packets. This means that an SDU cannot be delivered to the higher layers unless all the SDUs preceding it in the flow have been received correctly, or have been discarded. If an ARQ block is finally discarded, despite retransmissions, the IEEE 802.16e standard mandates that the SDU to which it belongs cannot be delivered to the higher layers [1]. A block diagram of the simulator is shown in Figure 1.

The MAC-PHY simulator provides an accurate way to determine the MAC BLER and SDU error rate (SER) for the SDUs that are discarded, taking into account the MAC

layer parameters, data encapsulation, and the ARQ process. Importantly, the BLER for contiguous blocks is not independent due to the time-correlated nature of the fading channel, and this is enabled by modelling the instantaneous PHY PER $P_e(t)$. The simulator computes the following as a function of mean channel SNR, MCS mode, ARQ block lifetime, and MS velocity, taking into account MAC layer parameters such as packet size, the ARQ Retry timer, and the ARQ feedback time:

 (i) block error rate (BLER),

 (ii) SDU error rate (SER), equivalent to UDP PER (one-to-one mapping of SDUs to UDP packets),

 (iii) the time pattern of the ARQ block and SDU losses,

 (iv) end-to-end delay and jitter for blocks and SDUs,

 (v) channel capacity consumed during each DL sub-frame, measured in physical slots,

 (vi) transmission throughput and goodput.

The simulator records the transmission times for each ARQ block and SDU for a flow of N SDUs. The number of ARQ retransmissions k_{re} is estimated from the total ARQ block transmission time T_{ARQ}, computed as

$$T_{ARQ} = k_{re} \cdot (ARQRetry \cdot T_{OFDMA}) + T_{OFDMA}, \quad (1)$$

where T_{OFDMA} is the duration of an OFDMA frame (i.e., 5 ms) and *ARQRetry* is the *ARQRetry* Timeout parameter (number of frames).

The number of ARQ retransmissions k_{re} is estimated when a bound is imposed on T_{ARQ} by the block lifetime l_{bl}, as

$$T_{ARQ} \leq l_{bl}, \quad (2)$$

$$\max(k_{re}) = \left\lfloor \frac{l_{bl} - T_{OFDMA}}{ARQRetry \cdot T_{OFDMA}} \right\rfloor. \quad (3)$$

In general, if insufficient resources are allocated for all retransmissions in the frame where the PDU is to be resent, according to the ARQ Retry Timeout timer, an additional queuing time, Tq, will be added in (1). This represents the block queuing time in the retransmission queue. In this case, the number of retransmissions that will take place will be less than $\max(k_{re})$, as T_{ARQ} is limited by (2). Hence, k_{re} is not deterministic for all blocks.

The simulator calculates the end-to-end SDU latency and jitter for a flow of N SDUs $\in \{1, 2, 3, \ldots, j, j+1, j+2, \ldots, N\}$. The SDU end-to-end latency for SDU j, D_j is calculated as the time difference between the arrival of the MAC SDU j at the MAC layer transmitter, $T_{arr,j}$, and the delivery of the SDU j to the transport layer at the receiver, $T_{rec,j}$, as shown in (4)

$$D_j = T_{rec,j} - T_{arr,j}, \quad (4)$$

$$D_j = D_{tx} + D_{del}. \quad (5)$$

The end-to-end latency of an SDU j, D_j, consists of the transmission time D_{tx} and the delivery time D_{del}.
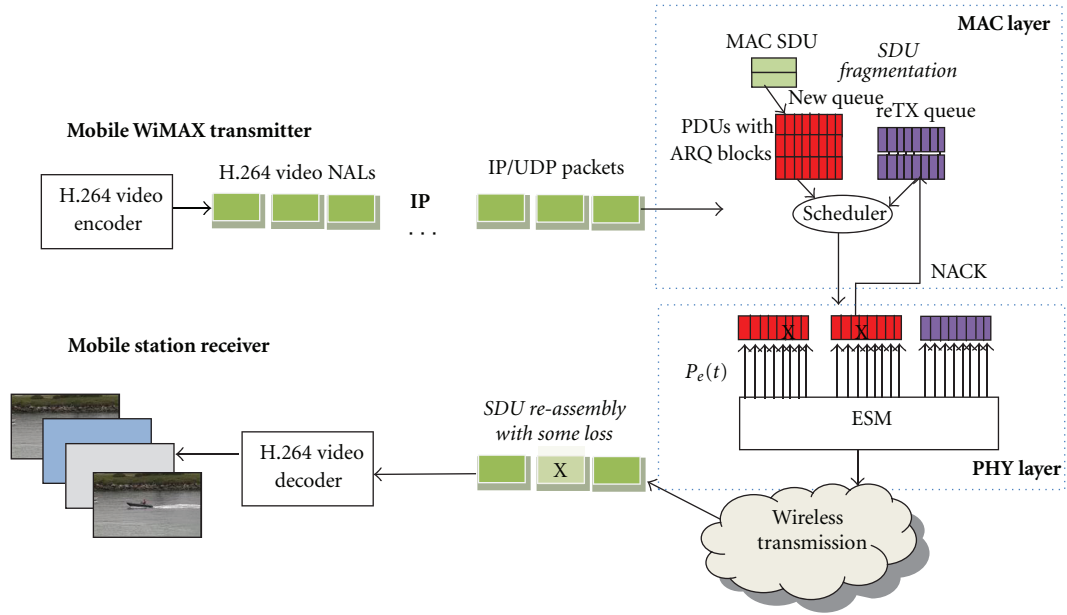
FIGURE 1: WiMAX MAC-PHY simulator.

The transmission time D_{tx} includes the retransmission time for the PDUs containing blocks of SDU j and the waiting time in the retransmission buffer, if any of the ARQ blocks contained in the SDU were retransmitted. An SDU is delivered to the receiver (at the transport layer) when all the ARQ blocks it consists of have been correctly received, after retransmission. Also, SDU packets are delivered in order. This is because the receiver must first receive correctly and reassemble all the ARQ blocks of SDU j and then deliver the SDUs following j. This means that if the ARQ blocks in SDU j have undergone retransmission, the SDUs $j + 1, j + 2, \ldots$ which follow SDU j will be delayed as well, even if no errors and retransmissions occurred for them. Therefore, retransmissions can cause a build up of delay, not only for the SDU which suffered the transmission errors, but also for SDUs following it. If the channel is poor and errors occur frequently, the delay build-up can be significant.

Figure 2 depicts the delay build-up for a number of SDUs, when some of the ARQ blocks of SDU S3 are lost and later retransmitted. Although SDUs S4, S5, and S6, succeeding S3, are received correctly, they are not delivered to the higher layers until all the ARQ blocks from SDU S3 have been received correctly (after ARQ retransmission). So SDUs S4 to S6 are delayed.

The variation in the end-to-end delay (latency) of the SDUs, for a flow of SDUs, is referred to as jitter. Another term commonly used is Packet Delay Variation (PDV), defined in ITU-T Recommendation Y.1540. Jitter is calculated as the variance of the SDU delay as follows

$$\text{jitter} = \text{var}(D_{\text{SDU}}) = \frac{1}{N}\sum_{i=1}^{N}(D_i - \mu_D)^2, \qquad (6)$$

where D_{SDU} is the discrete function of the SDU latency. D_i is the end-to-end delay of SDU i, μ_D is the mean SDU delay, and N is the total number of SDUs transmitted.
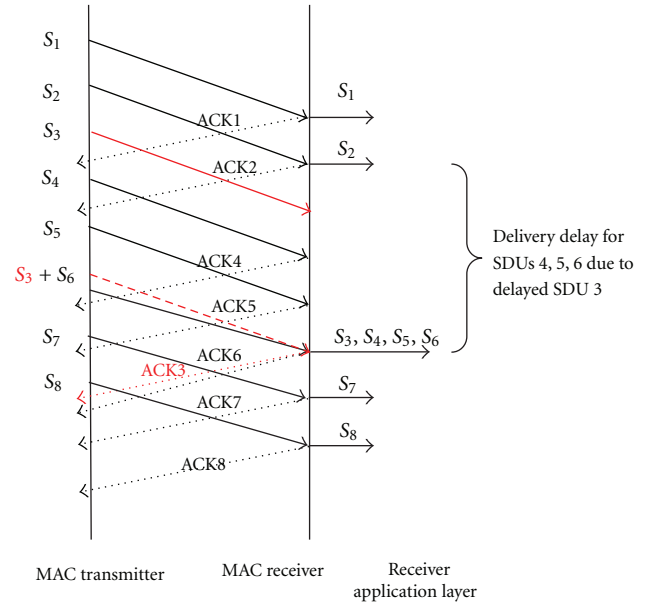

FIGURE 2: SDU delays due to retransmissions and ARQ_Deliver_In_Order.

In accordance with 802.16e recommendations [16, 17], here it is assumed that retransmissions cannot occur in the next frame but are sent at the earliest on the 4th DL subframe after transmission. Furthermore, the maximum number of retransmissions is limited to 4. These values result from processing time at the receiver and transmission delays in the radio network. The MAC parameter ARQ Retry Timeout is set to 4 and performance is simulated for 1–4 retransmissions, corresponding to 30–90 ms block lifetimes, according to (3).

Table 4: Mobile WiMAX simulator parameters.

Parameter	Value
OFDMA	
Carrier frequency	2.3 GHz
Channel bandwidth	10 MHz
FFT length	1024
Subcarrier frequency spacing	10.94 kHz
Frame length	5 ms (48 OFDMA symbols)
Guard interval	1/8
DL subcarrier permutation scheme	PUSC DL
Number of active DL subcarriers	840
Number of subchannels	30 DL/35 UL
Data subcarriers per subchannel	24
OFDMA data symbols	22 DL/15 UL
DL capacity	330 slots
DL/UL ratio	60/40
MAC	
MAC SDU size	815 bytes
ARQ block size	32 bytes
MAC PDU size	max 200 bytes
PDU packing	No
SDU fragmentation	Yes
QoS scheduling	UGS
Block rearrangement	No
ARQ feedback type	Selective-ACK
ARQ retry timeout	4 frames
ARQ block lifetime	0–90 ms
Max no. retransmissions	4
ARQ_DELIVER_IN_ORDER	Yes

3.3. Simulation Parameters. Table 4 summarises the PHY and MAC layer parameters used in the MAC-PHY mobile WiMAX simulator. Simulation results are presented in the next sections for a flow of UDP packets corresponding to the "ICE" video sequence [27] (704 × 576 pixels at 30 fps) encoded using H.264 at 1.03 Mbps, with I, P, and B frames.

The simulator captures the transmission of 2000 UDP packets, 815 Bytes each, transmitted through the 802.16e PHY and MAC layers. UDP packets correspond 1 : 1 to MAC SDUs, also 815 Bytes long. Each MAC SDU was fragmented into ARQ blocks, 32 bytes long. Simulations were carried out for MS velocities $v = 1$ km/h and 10 km/h for mean channel SNR values varying from 8 dB–22 dB.

4. Analysis of BLER and UDP PER

In Figures 3 and 4, the BLER and SDU error rate are compared for an ARQ lifetime of 65 ms (i.e., up to 3 retransmissions may have occurred). The simulated MS speed is 1 km/h. It is observed that the BLER seen at the MAC layer after the retransmissions is projected to a much higher SER and UDP PER at the higher layers. For example, at SNR = 16 dB 16QAM 3/4 results in 0.063 BLER, which corresponds to 0.089 SER. This is because, according to the standard [1],
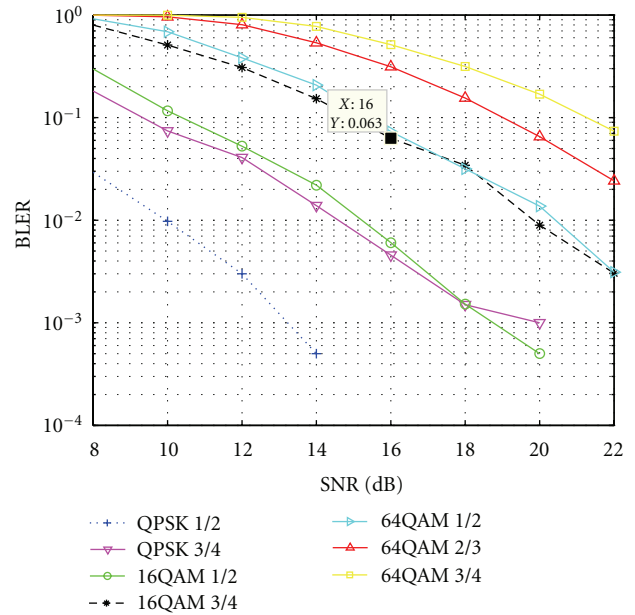

Figure 3: BLER versus SNR when block lifetime is 65 ms and MS $v = 1$ km/h.

even if just one ARQ block in an SDU is discarded it will result in the whole SDU being discarded. The SDU error rate is accentuated more for higher BLER, for example, at 14 dB 64QAM 2/3 gives 0.53 BLER and 0.65 SER. This shows that in order to achieve high video quality and quasi-zero PER at the video receiver (i.e., SER), the ARQ retransmissions must achieve quasi-zero BLER. However, with the number of retransmissions limited to 3, only lower modes can deliver error free data in a slowly time-varying channel. From Figure 3, it is obvious that only the QPSK modes and 16QAM 1/2 can deliver SER $\leq 10^{-2}$ for SNR ≤ 18 dB, when the MS speed is 1 km/h and the block lifetime is 65 ms.

In Figure 5, SER versus channel SNR is shown for an MS speed of 1 km/h, when up to 4 retransmissions are allowed, for a block lifetime of 90 ms. It is shown that the SER is lower with a longer block lifetime. QPSK 1/2 delivers error free data for SNR ≥ 12 dB and 16QAM 1/2 attains an SER < 0.02 for SNR ≥ 14 dB.

Our previous work in [18] focused on the BLER attained (rather than SER) when ARQ retransmissions occurred in the next DL subframe, without limitation on the maximum number of retransmissions (since the ARQ retransmission frequency is not specified in the 802.16e standard [1]). In [18], it was shown that the BLER achieved was below 10^{-2} at a mean channel SNR of 8 dB, for a block lifetime of 100 ms, with MCS modes 16QAM 1/2 or lower. All MCS modes up to 64QAM 1/2 attained quasi error free transmission when the mean channel SNR was 12 dB and block lifetimes were greater than 70 ms. This was possible because, in that scenario, the maximum number of permitted retransmissions was 7 for a block lifetime of 70 ms; this resulted in more favorable ARQ performance. In this work, it is shown that imposing practical limitations on the ARQ Retry Timeout parameter and the maximum number of retransmissions (as
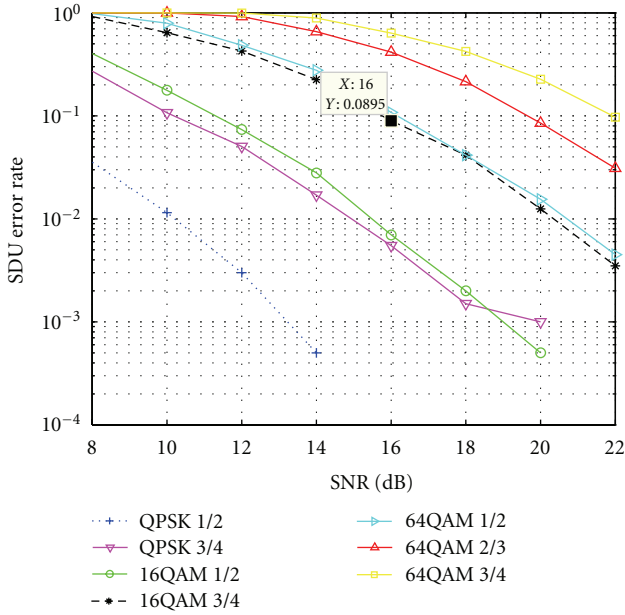
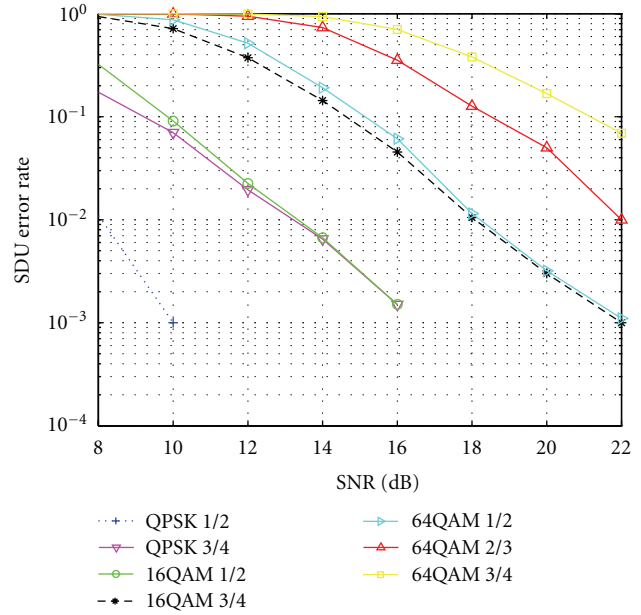FIGURE 4: SER (or UDP PER) versus SNR when block lifetime is 65 ms and MS $v = 1$ km/h.

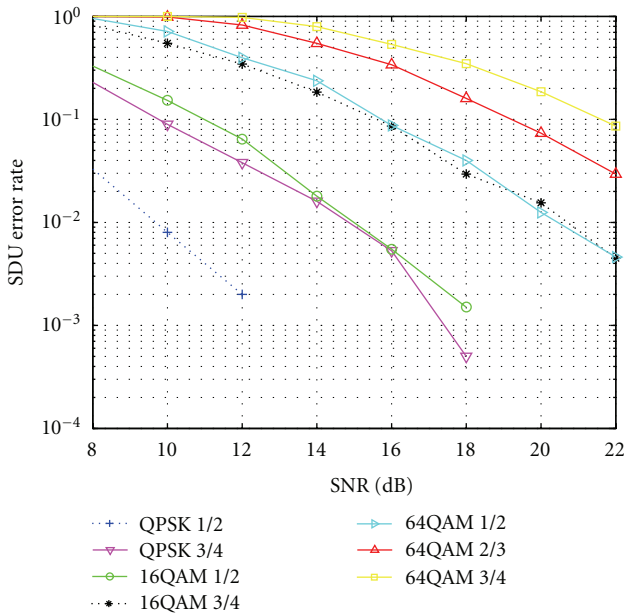FIGURE 6: SER versus SNR when block lifetime is 65 ms and MS $v = 10$ km/h.

FIGURE 5: SER versus SNR for a block lifetime of 90 ms and MS $v = 1$ km/h.

recommended in [16, 17]) results in a residual BLER. This residual BLER is further accentuated as UDP PER at the higher layers.

In Figure 6, the SER versus SNR across all MCS modes for an MS speed of 10 km/h is compared with the SER versus SNR for an MS speed of 1 km/h, as shown in Figure 4. The maximum number of retransmissions in both cases is determined by a 65 ms block lifetime. It is clear that the SER attained for the 1 km/h channel is higher than the SER for

the 10 km/h channel at all channel SNR values. For example, at 16 dB 16QAM 3/4 attains SER = 0.089 at 1 km/h while this drops to SER = 0.045 at 10 km/h. QPSK 1/2 at 12 dB attains SER = 0.03 for 1 km/h speed, whereas it gives SER = 0 (below the simulation accuracy level) for a 10 km/h MS speed. The low MS speed means that the channel coherence time is longer and the channel decorrelates slower in time. Hence, more ARQ retransmissions are required at the slower speed to achieve a quasi-zero level of SER. The effect of channel coherence time on ARQ retransmissions was also studied in [24] for 802.11 a/g networks, where ARQ retransmission was implemented according to a stop-and-wait mechanism and was governed by the CSMA/CA access protocol.

Next, the BLER and SER attained when multicasting is studied. In Figures 7 and 8 ARQ is not enabled, as is the case for multicasting. The MS speed is 1 km/h. It can be seen that the SER in Figure 8 is much higher than the BLER for the same channel SNR and MCS mode. For example, at 14 dB QPSK 3/4 delivers BLER = 0.039, while the SER is 0.15. The lack of ARQ error correction limits the video broadcast performance over mobile WiMAX, as explained in [11]. Without additional error correction, real-time video multicasting could not be offered for an SNR range below 16 dB (as SER $> 10^{-2}$ [28]), and even then only with the lowest throughput mode QPSK 1/2, which consumes considerable channel capacity.

5. ARQ-Aware Scheduling and Latency/Jitter

Another very important aspect of video transmission with ARQ is the latency and jitter that occurs. As discussed in [29–31], for video applications the playback buffer that masks network jitter can take values in the order of 250 ms, while latency is acceptable up to 100–150 ms, depending on
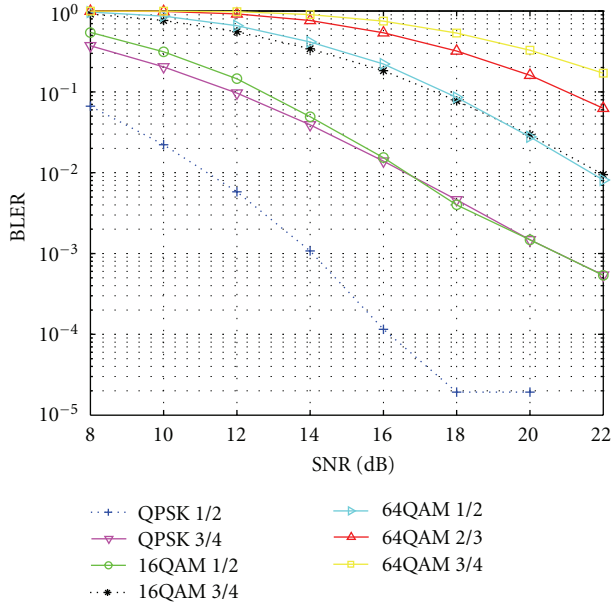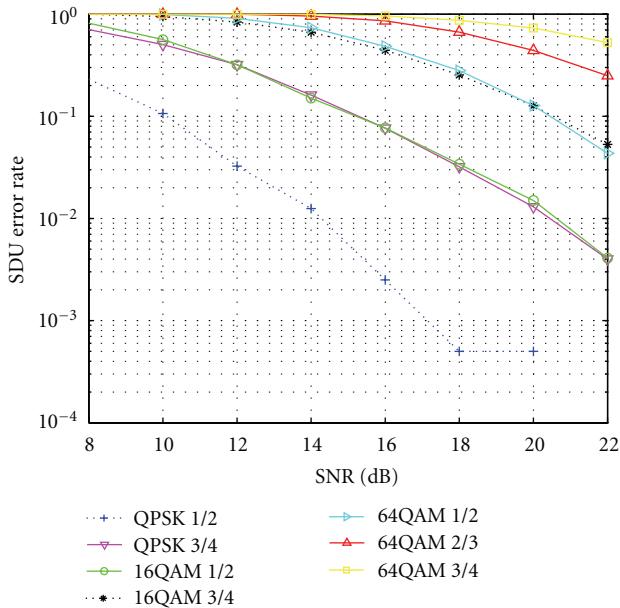
FIGURE 7: BLER versus SNR without ARQ, MS $v = 1$ km/h.

FIGURE 9: PDF of end-to-end SDU delay, SNR = 12 dB, mode = 1, $v = 1$ km/h.

FIGURE 8: SER versus SNR without ARQ, MS $v = 1$ km/h.

FIGURE 10: Maximum delay versus block lifetime, for MCS modes 0 and 1 when channel SNR = 12 dB and MS $v = 1$ km/h.

the specific video characteristics and applications. Assuming ARQ-aware scheduling, here the latency and jitter associated with ARQ retransmissions is studied over an 802.16e network.

For the 2000 transmitted SDUs, the simulator calculates the total latency for each SDU. Figure 9 shows the PDF of the end-to-end SDU delay computed during the transmission of 2000 SDUs at a channel SNR = 12 dB, MCS mode 1 (QPSK 3/4), and MS speed 1 km/h, for block lifetimes of 30 ms (up to 1 retransmission), 65 ms (up to 3 retransmissions), and 90 ms (up to 4 retransmissions).
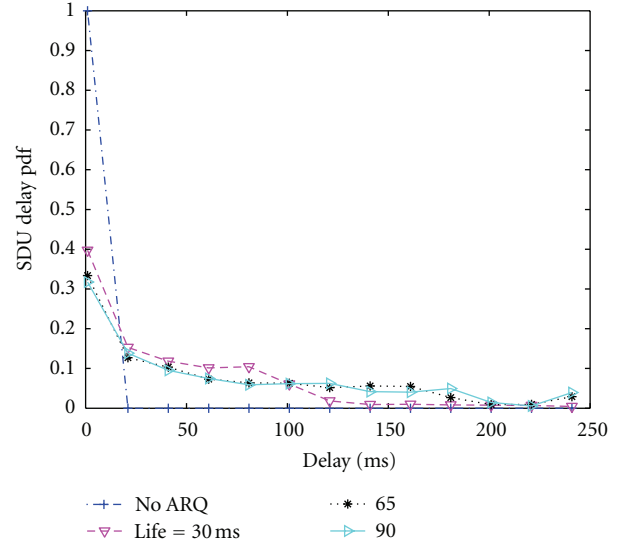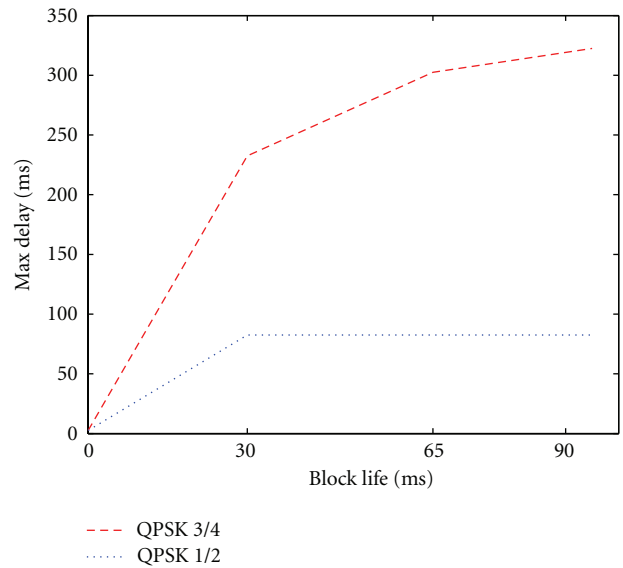
The simulator computes the latency during transmission and the jitter as the variance of the SDU latency across the 2000 SDUs, for each SNR, MCS mode, and block lifetime. Figure 10 shows the maximum latency attained versus block lifetime, when the mean channel SNR is 12 dB and the MS speed is 1 km/h. It is observed that at SNR = 12 dB (when the channel is poor) the maximum latency is fixed to approximately 82 ms for block lifetimes of 30–90 ms when QPSK 1/2 is used. This MCS mode delivers SER = 0.003 (see Figure 4). If QPSK 3/4 is used, the maximum delay increases for each block lifetime, reaching over 300 ms for a block lifetime of 90 ms. This occurs because a very large number of ARQ blocks are in error and many retransmissions occur.

TABLE 5: Overallocation of resources for lifetime 65 ms.

| MCS	SNR = 10 dB	SNR = 22 dB
0	0.06	0
1	0.42	0
2	0.68	0
3	2.48	0.03
4	3.05	0.02
5	3.98	0.15
6	3.99	0.36

This MCS mode attains SER = 0.05 when the block lifetime is 65 ms, despite the retransmissions that take place (see Figure 4). This mode would not be selected for transmission at SNR = 12 dB by the link adaptation algorithm because the BLER is very high for the amount of retransmissions allowed. The amount of resources allocated by the scheduler in this case does not cater for the very large number of retransmissions that occur, resulting in a buildup of queuing delay and also a large increase in jitter. The amount of "overallocation" of resources the ARQ-aware scheduler predicts is related to the level of PER attained by the selected MCS mode and the ARQ block lifetime.

In Table 5, the overallocation γ required for different MCS modes is given for mean channel SNR values of 10 dB and 22 dB, when the block lifetime is 65 ms. If S is the number of slots required per DL frame, for a given bitrate and MCS mode, the scheduler needs to allocate $(1 + \gamma) \cdot S$ slots per DL frame. The overallocation γ is calculated by dividing the number of slots required per DL frame for the desired number of ARQ retransmissions (according to the MCS mode), by S. From this table, it is obvious that, for example, if mode 1 was selected at SNR = 10 dB, the scheduler would need to allocate $\gamma = 0.42$ more resources than that required if ARQ was not enabled. A smaller allocation than this results in queuing delays. It is also obvious that the overallocation required for the higher modes at low SNR values is unacceptable (i.e., three times the amount of resources required for the video bitrate).

Figures 11 and 12 show jitter versus block lifetime for SNR = 22 dB at an MS speed of 1 km/h and for SNR = 12 dB at an MS speed of 10 km/h, correspondingly. Jitter is studied for the MCS modes that attain SER ≤ 0.02 and that can deliver quality video. In Figure 11, where the mean channel SNR = 22 dB, the jitter is below 100 ms when the block lifetime is 65 ms, for MCS modes 0 to 4. All of these modes deliver quasi error free data.

Figure 12 shows the jitter when the MS speed is 10 km/h at SNR = 12 dB. The jitter is approximately 100 ms for mode 1 when the block lifetime is 65 ms. Both QPSK 3/4 and 16QAM 1/2 attain an SER of approximately 0.02, but for 16QAM 1/2 more retransmissions occur and therefore higher jitter ensues.

The study of the simulation results on SER, latency and jitter, when the recommendations from [17, 32] regarding ARQ parameters are applied, leads to the conclusion that when the channel SNR is poor (SNR ≤ 14 dB) only the lower
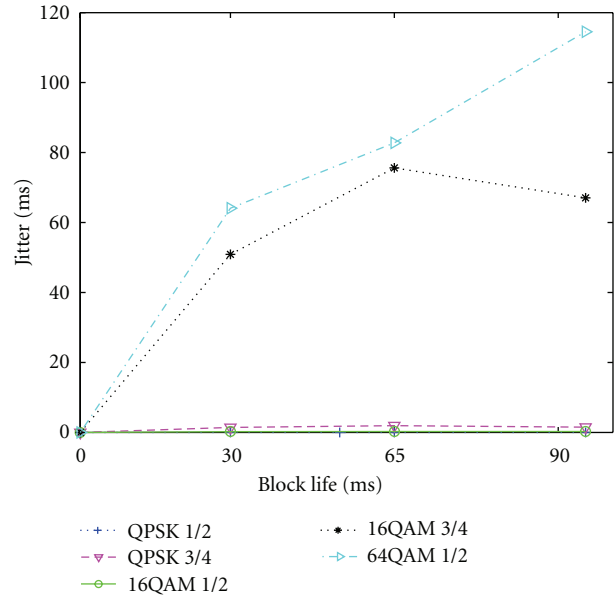
FIGURE 11: Jitter versus block lifetime, for MCS modes 0 to 4 when channel SNR = 22 dB and MS v = 1 km/h.
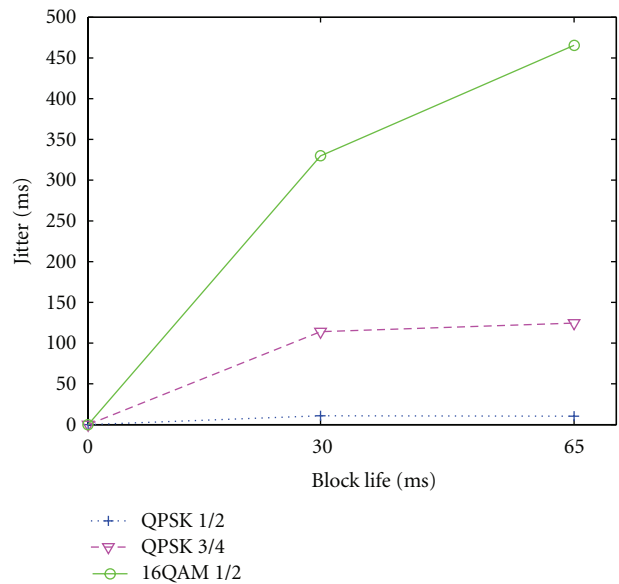
FIGURE 12: Jitter versus block lifetime, for MCS modes 0 to 3 when channel SNR = 12 dB and MS v = 10 km/h.

QPSK modes can deliver quality video with acceptable jitter. For a UDP unicast video transmission, in order to attain quasi error free SER, and at the same time limit jitter to below 100 ms, QPSK modes should be used with up to 3 retransmissions. When the channel has a long coherence time more retransmissions are required in order to deliver the ARQ blocks error free, and only QPSK 1/2 can be used. However, even then the SER attained (0.03) is not quasi error free.
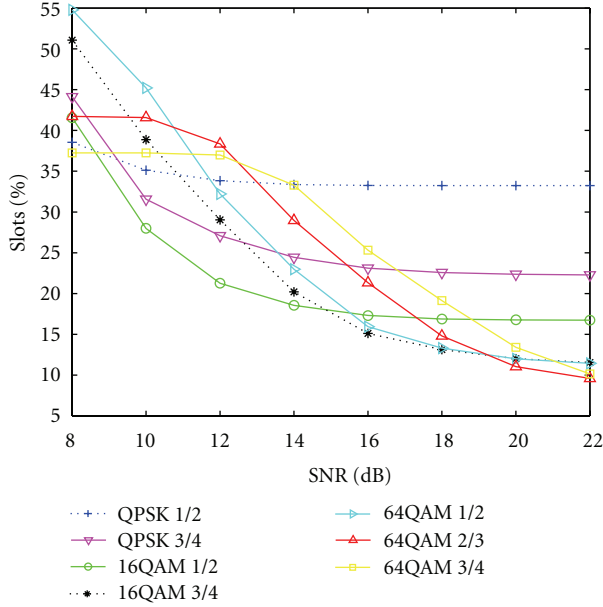
FIGURE 13: Percentage of physical DL slots required versus SNR, block lifetime = 65 ms, v = 1 km/h.

6. Transmission Efficiency

Having discussed the mobile WiMAX performance in terms of SER and jitter when ARQ is enabled, the channel resources required during unicast video transmission are now studied. The simulator estimates the total number of physical slots required for the transmission of 2000 SDUs, at each mean channel SNR and MCS mode, including all ARQ retransmissions. Then the channel capacity required is calculated as a percentage of the total number of DL slots available, for the duration of the transmission. This work focuses on the goodput attained. A novel transmission efficiency metric is proposed, namely the goodput per frame, which takes into account the goodput achieved for the amount of channel resources required per DL subframe.

Figure 13 shows the channel capacity required, as a percentage of the total physical DL slots available per DL subframe, versus the mean channel SNR, for all MCS modes when up to 3 retransmissions are allowed. It is shown that for QPSK 1/2 there is a very small differentiation in the channel resources required for SNR $\geq$ 14 dB, as resource requirements drop from 38.5% of slots at SNR = 8 dB to 33%. This is because very few retransmissions occur for higher channel SNR values with QPSK 1/2, therefore the resources required are constant, corresponding to the new data arrivals. In other cases, the differentiation of resources required across SNR values is much greater. For example, for 64QAM 1/2 at SNR = 8 dB approximately 55% of the total slots are required, whereas for SNR = 22 dB only 11.5% of the total slots are required. This is because more ARQ retransmissions are required at low SNR values. It is also obvious that the higher throughput modes require less resources than the lower modes, even at low SNR values when retransmissions occur. For example, 16QAM

3/4 requires 29% of the total resources at SNR = 12 dB, while QPSK 1/2 requires 34% of the resources. This is because the lower modes pack less data bits per slot, as shown in Table 3.

It is clear that greater bandwidth efficiency can be achieved when higher MCS modes are used. However, as shown by the SDU loss rate, in order to support QoS for real time video transmission, only lower MCS modes can be successfully used at lower SNR values.

The goodput delivered per OFDMA frame for one flow of data, $G_{\text{flow}}(s, l, m)$, is computed as the correct number of bits received, $CorrectBits$, divided by transmission duration, F_T, in number of OFDMA frames, required for the transmission of N UDP packets. The goodput delivered per OFDMA frame is calculated for each mean channel SNR s, MCS mode m, and block lifetime l, as

$$G_{\text{flow}}(s, l, m) = \frac{CorrectBits(s, l, m)}{F_T(s, l, m) \cdot T_{\text{OFDMA}}}, \qquad (7)$$

where T_{OFDMA} is the duration of an OFDMA frame (i.e., 5 ms).

The average channel capacity, in slots, required for the data flow per OFDMA frame, $\theta_{\text{flow,fr}}$, is calculated as the total number of required physical slots C_T divided by the transmission duration, F_T, for each mean channel SNR s, MCS mode m, and block lifetime l

$$\theta_{\text{flow,fr}}(s, l, m) = \frac{C_T(s, l, m)}{F_T(s, l, m)}. \qquad (8)$$

Therefore, a $\theta_{\text{flow,fr}}$ capacity per DL frame delivers G_{flow} goodput per frame, for each channel SNR, MCS mode, and block lifetime. The goodput-per-frame efficiency metric measures the goodput G_{flow} delivered per frame taking into account the average capacity required per frame (in slots), $\theta_{\text{flow,fr}}$. Hence goodput per frame, g_{fr}, is defined as

$$g_{\text{fr}}(s, l, m) = \frac{G_{\text{flow}}(s, l, m)}{\theta_{\text{flow,fr}}(s, l, m)}. \qquad (9)$$

If the total capacity of the DL frame, S_{fr}, is used, the system can support a total goodput per frame Gdp_{fr}, calculated using (10). The total capacity of a DL frame, S_{fr}, in slots, is a system dependent parameter and for the mobile WiMAX system simulated, assuming a PUSC enabled DL, S_{fr} = 330 slots (Table 1). The system therefore can support a total goodput per frame Gdp_{fr} given by

$$Gdp_{\text{fr}}(s, l, m) = \frac{G_{\text{flow}}(s, l, m) \cdot S_{\text{fr}}}{\theta_{\text{flow,fr}}(s, l, m)}. \qquad (10)$$

Figure 14 shows the total goodput per frame versus channel SNR, estimated for all MCS modes when the block lifetime is 90 ms and the MS speed is 1 km/h. It can be seen that QPSK 1/2 offers the highest transmission efficiency for SNR $\leq$ 10 dB and 16QAM 1/2 for SNR values in the range 10 dB to 16 dB. For higher SNR values 16QAM 3/4 is more efficient. 16QAM 3/4 can support 7-8 Mbps when the mean channel SNR is 18–20 dB. The goodput per frame metric helps identify the most efficient mode per channel SNR for
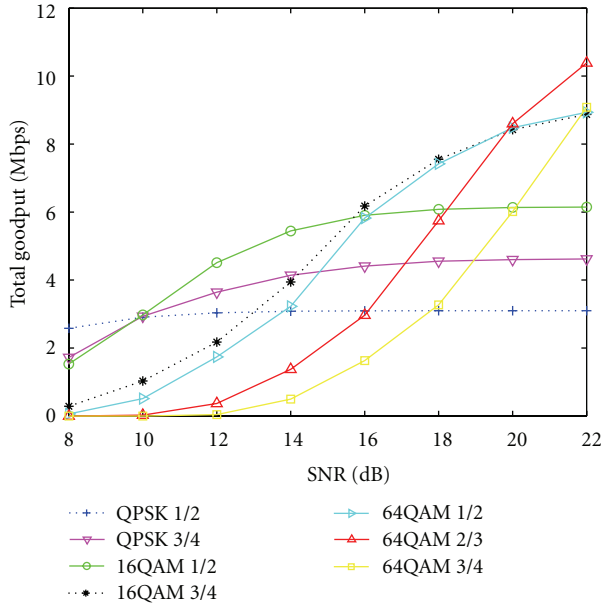
FIGURE 14: Total goodput for lifetime = 90 ms and MS v = 1 km/h.

FIGURE 15: Total goodput versus SNR, PER $\leq 10^{-2}$, block lifetime = 90 ms, and no ARQ, MS v = 1 km/h.

a particular block lifetime. It can also be seen that only lower throughput modes (up to mode 3) are selected for SNR values below 20 dB.

In order to select the most efficient MCS mode that also delivers zero PER, or the PER below a minimum acceptable value (such that video QoS can be guaranteed), a constraint should be applied based on the PER attained at the application layer per channel SNR. Loss of packets can seriously degrade the quality of received video [28, 33]. In the following, UDP PER is constrained to less than 1%. Figure 15 shows the total goodput per frame versus channel SNR, for all MCS modes with UDP PER $\leq 10^{-2}$, when the block lifetime is 90 ms and the MS speed is 1 km/h. If the PER attained for a particular MCS mode and SNR value is higher than 10^{-2}, then the total goodput is set to zero. It is observed that below an SNR value of 12 dB no MCS mode can achieve quasi error free transmission. Therefore, quality unicast video cannot be offered for channel SNR below 12 dB. For SNR values in the range 12 dB to 16 dB only QPSK 1/2 can achieve the desired PER. For SNR values in the range 16 dB to 20 dB the most efficient MCS mode that offers quasi error free transmission is 16QAM 1/2, whereas for SNR = 22 dB mode 64QAM 1/2 offers the highest transmission efficiency, for a PER below 10^{-2}. In Figure 15, the total goodput attained for multicast transmission (no ARQ enabled) is also shown. The multicast transmission also observes the PER $\leq 10^{-2}$ constraint. It can be seen that the multicast video service cannot be offered quasi error free for SNR values below about 16 dB. For SNR values in the range 16 dB to 18 dB only QPSK 1/2 can achieve the desired PER. For SNR values in the range 20 dB to 22 dB mode 16QAM 1/2 offers the highest transmission efficiency while keeping the PER below 10^{-2}. This comparison shows that using ARQ with a block lifetime of 90 ms offers a significant gain of approximately 3 Mbps (i.e., double the goodput) for
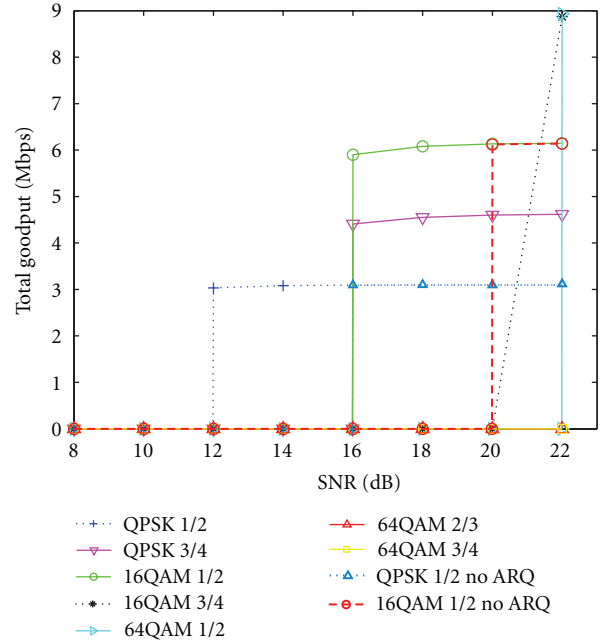
16 dB $\leq$ SNR $\leq$ 18 dB and a gain of 2.8 Mbps at a channel SNR of 22 dB. Also the range of the service is extended by approximately 4 dB when ARQ is enabled with a block lifetime of 90 ms.

7. Conclusions

From the simulation results presented in this work, it has been possible to study the performance of different MCS modes and block lifetimes for various channel SNR values in a mobile broadband network. The performance of ARQ retransmissions was shown to depend on the MS velocity. It was shown that at walking speeds (e.g., 1 km/h) ARQ retransmissions are less efficient. Only mode 0 succeeds in delivering data with PER $\leq 10^{-2}$ for 12 $\leq$ SNR < 16 dB, and only modes 1 and 2 for SNR $\leq$ 20 dB, with up to 4 retransmissions. The SER attained is lower for the same MCS mode, block lifetime and channel SNR, when the channel coherence time is short.

It was demonstrated that a channel-aware and ARQ-aware scheduler should be used in order to predict and provide sufficient resources per frame for delay sensitive services, such as real-time video. The novel efficiency metric, goodput per frame, has enabled a performance comparison in terms of PER achieved and radio resources required when S-ACK ARQ is enabled, for various MCS modes. This efficiency metric showed the total goodput that each MCS mode could support for each channel SNR for a given block lifetime. The goodput per frame metric was found to be a valuable tool for radio resource management in broadband wireless networks.

Insight has been gained on the importance of the ARQ Retry Timeout MAC parameter and how it affects system performance, not only in terms of SDU delay, but PER as well. When practical considerations are taken into account regarding the frequency and number of possible retransmissions based on the WiMAX Forum recommendations, only modes 0, 1, and 2 can deliver quasi error free data. Jitter in these cases is maintained within acceptable limits for video QoS. However, the lower modes, 0 and 2, that attain quasi error free transmission, require 35% and 18% of the total channel resources, correspondingly, to successfully deliver video at 1.03 Mbps quasi error free. The total goodput attained is 3Mbps for SNR $\leq$ 16 dB (with MCS mode 0) and 6 Mbps for channel SNR values in the range 16 dB to 20 dB (with MCS mode 2). A quality unicast video service cannot be offered for channel SNR values below 12 dB.

Multicasting real-time video, while observing a QoS without ARQ retransmissions, is ineffective below about 16 dB SNR. Additional error correction mechanisms are necessary in order to support high quality multicast video, such as the Application Layer Forward Error Correction (FEC) mechanism based on Raptor codes, endorsed by the 3GPP MBMS [34].

References

[1] *IEEE Standard for Local and metropolitan area networks part 16: 802.16*, IEEE Std., 2006.

[2] LTE, *Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (3GPP TS 36.300 version 10.5.0 Release 10)*, 3GPP Std., http://www.3gpp.org/ftp/Specs/html-info/36300.htm.

[3] O. Oyman, J. Foerster, Y. J. Tcha, and S. C. Lee, "Toward enhanced mobile video services over WiMAX and LTE," *IEEE Communications Magazine*, vol. 48, no. 8, pp. 68–76, 2010.

[4] I. Cerutti, F. Meucci, P. Castoldi, and L. Pierucci, "An adaptive cross-layer strategy for QoS-guaranteed links in 4G networks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '08)*, pp. 1–5, New Orleans, La, USA, December 2008.

[5] H. Martikainen, A. Sayenko, O. Alanen, and V. Tykhomyrov, "Optimal MAC PDU size in IEEE 802.16," in *Proceedings of the 4th International Telecommunication Networking Workshop on QoS in Multiservice IP Networks (IT-NEWS '08)*, pp. 66–71, Venice, Italy, February 2008.

[6] A. Sayenko, V. Tykhomyrov, H. Martikainen, and O. Alanen, "Performance analysis of the IEEE 802.16 ARQ mechanism," in *Proceedings of the 10th ACM Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems (MSWiM '07)*, pp. 314–322, ACM, October 2007.

[7] F. Hou, J. She, P. H. Ho, and X. Shen, "Performance analysis of ARQ with opportunistic scheduling in IEEE 802.16 networks," in *Proceedings of the 50th Annual IEEE Global Telecommunications Conference (GLOBECOM '07)*, pp. 4759–4763, Washington, DC, USA, November 2007.

[8] W. Wang, Z. Guo, X. Shen, and C. Chen, "Performance analysis of ARQ scheme in IEEE 802.16," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '06)*, pp. 1–5, San Francisco, Calif, USA, December 2006.

[9] W. Wang, Z. Guo, X. Shen, C. Chen, and J. Cai, "Dynamic bandwidth allocation for QoS provisioning in IEEE 802.16

[10] M. Chatterjee, S. Sengupta, and S. Ganguly, "Feedback-based real-time streaming over WiMax," *IEEE Wireless Communications*, vol. 14, no. 1, pp. 64–71, 2007.

[11] W. Jianfeng, M. Venkatachalam, and F. Yuguang, "System architecture and cross-layer optimization of video broadcast over WiMAX," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 4, pp. 712–721, 2007.

[12] C. So-In, R. Jain, and A. K. Tamimi, "Scheduling in IEEE 802.16e mobile WiMAX networks: key issues and a survey," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 2, pp. 156–171, 2009.

[13] A. Sayenko, O. Alanen, and T. Hämäläinen, "ARQ aware scheduling for the IEEE 802.16 base station," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 2667–2673, Beijing, China, May 2008.

[14] S. Deb, S. Jaiswal, and K. Nagaraj, "Real-time video multicast in WiMAX networks," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 1579–1587, Phoenix, Ariz, USA, April 2008.

[15] *3GPP TR 25.996 V6.1.0 Spatial channel model for Multiple Input Multiple Output (MIMO) simulations (Release 6)*, 3GPP Std., 2009.

[16] Mobile System Profile, *WMF-T23-001-R010v09*, WiMAX Forum Std., 2010.

[17] *IEEE 802.16m Evaluation Methodology Document (EMD) IEEE 802.16m-08/004r5*, IEEE 802.16 Broadband Wireless Access Working Group Std., 2009, http://ieee802.org/16.

[18] V. Sgardoni, D. Halls, S. M. M. Bokhari, D. Bull, and A. Nix, "Mobile WiMAX video quality and transmission efficiency," in *Proceedings of the IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC '11)*, pp. 1077–1082, September 2011.

[19] A. Esmailpour and N. Nasser, "Dynamic QoS-based bandwidth allocation framework for broadband wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 6, pp. 2690–2700, 2011.

[20] *IEEE Standard for Local and metropolitan area networks part 16: Air Interface for Broadband WirelessAccess Systems 802.16-2009*, IEEE Std., 2009.

[21] *Mobile WiMAX—part I: A Technical Overview and Performance Evaluation*, WiMAX Forum Std., 2006.

[22] L. Wan, S. Tsai, and M. Almgren, "A fading-insensitive performance metric for a unified link quality model," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '06)*, pp. 2110–2114, April 2006.

[23] D. Halls, A. Nix, and M. Beach, "System level evaluation of UL and DL interference in OFDMA mobile broadband networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '11)*, pp. 1271–1276, Quintana Roo, Mexico, March 2011.

[24] V. Sgardoni, P. Ferré, A. Doufexi, A. Nix, and D. Bull, "Frame delay and loss analysis for video transmission over time-correlated 802.11A/G channels," in *Proceedings of the 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '07)*, pp. 1–5, Athens, Greece, September 2007.

[25] M. Tran, D. Halls, A. Nix, A. Doufexi, and M. Beach, "Mobile WiMAX: MIMO performance analysis from a quality of service (QoS) viewpoint," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '09)*, pp. 1–6, Budapest, Hungary, April 2009.

[26] C. Cicconetti, A. Erta, L. Lenzini, and E. Mingozzi, "Performance evaluation of the IEEE 802.16 MAC for QoS support," *IEEE Transactions on Mobile Computing*, vol. 6, no. 1, pp. 26–38, 2007.

[27] "video test sequences," http://media.xiph.org/video/derf/.

[28] H. Jenkač, T. Stockhammer, and G. Liebl, "H.264/AVC video transmission over MBMS in GERAN," in *Proceedings of the IEEE 6th Workshop on Multimedia Signal Processing*, pp. 191–194, October 2004.

[29] T. Stockhammer and M. M. Hannuksela, "H.264/AVC video for wireless transmission," *IEEE Wireless Communications*, vol. 12, no. 4, pp. 6–13, 2005.

[30] M. Baldi and Y. Ofek, "End-to-end delay analysis of videoconferencing over packet-switched networks," *IEEE/ACM Transactions on Networking*, vol. 8, no. 4, pp. 479–492, 2000.

[31] S. Liang and D. Cheriton, "TCP-RTM: using TCP for real time multimedia applications," in *International Conference on Network Protocols*, 2002.

[32] WiMAX Forum Air Interface Specifications, *WiMAX Forum Mobile System Profile, Release 1.0, WMF-T23- 001-R010v09*, WiMAX Forum Std., 2007.

[33] D. Wu, Y. T. Hou, and Y. Q. Zhang, "Transporting real-time video over the Internet: challenges and approaches," *Proceedings of the IEEE*, vol. 88, no. 12, pp. 1855–1877, 2000.

[34] *3GPP TS 26.346 V8.0.0 (2008-10) Universal Mobile Telecommunications System (UMTS); Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs*, ETSI Std., 2008.