An Experimental Investigation of Real-time Video Transmission over IEEE 802.11e WLAN Networks using TXOP

Nicola Cranley
*Technological University Dublin*, nicola.cranley@tudublin.ie

Mark Davis
*Technological University Dublin*, mark.davis@tudublin.ie

Follow this and additional works at: [https://arrow.tudublin.ie/commcon](https://arrow.tudublin.ie/commcon)

Part of the [Electrical and Computer Engineering Commons](https://arrow.tudublin.ie/commcon)

**Recommended Citation**

This Conference Paper is brought to you for free and open access by the Communications Network Research Institute at ARROW@TU Dublin. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact yvonne.desmond@tudublin.ie, arrow.admin@tudublin.ie, brian.widdis@tudublin.ie.

This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/)
An Experimental Investigation of IEEE 802.11e TXOP facility for Real-Time Video Streaming

Nicola Cranley  
CNRI, Focas Institute,  
DIT Kevin Street, Dublin 8, Ireland

Mark Davis  
CNRI, Focas Institute,  
DIT Kevin Street, Dublin 8, Ireland

Abstract—Real-time multimedia streaming applications require a strict bounded end-to-end delay and are considered to be bursty as each video frame is typically transmitted as a burst of packets. In this paper we show how the distribution of video frame sizes can be used to efficiently dimension the IEEE 802.11e TXOP limit parameter to efficiently deal with this burstiness in order to enhance the transmission of real-time video streaming services. Through experimental investigation, we show that by using the mean video frame size to dimension the TXOP limit parameter, the transmission delay for the video frame is reduced by 67% under heavily loaded conditions. Other techniques investigated in this paper include applying the TXOP facility separately to each of the constituent I, P, and B video frame types.

Index Terms—Video Streaming, Performance Evaluation, Quality of Service, WLAN.

I. INTRODUCTION

Streaming multimedia over wireless networks is becoming an increasingly important service. This trend includes the deployment of WLANs that enable users to access various services including those that distribute rich media content anywhere, anytime, and from any device e.g. in-home wireless entertainment systems. There are many performance-related issues associated with the delivery of time-sensitive multimedia content using current IEEE 802.11 WLAN standards. Among the most significant are low delivery rates, high error rates, contention between stations for access to the medium, back-off mechanisms, collisions, signal attenuation with distance, signal interference, etc. Multimedia applications, in particular, impose onerous resource requirements on bandwidth constrained WLAN networks. Moreover, it is difficult to provide QoS in WLAN networks as the capacity of the network also varies with the offered load [1][2].

For real-time multimedia applications such as IPTV, video conferencing, and video telephony, packet loss and packets dropped due to excessive delay are the primary factors affecting the user-perceived quality. Real-time multimedia is particularly sensitive to delay as it has a strict bounded end-to-end delay constraint. Every multimedia packet must arrive at the client before its playout time with enough time to decode and display the contents of the packet. For video streams the delay incurred transmitting the entire video frame from the sender to the client is of particular importance. The loss rates incurred due to packets being delayed past their playout time is heavily dependent on the delay constraint imposed on the video stream. Video streaming applications typically impose an upper limit on the tolerable packet loss. Specifically, the packet loss ratio is required to be kept below a threshold to achieve acceptable visual quality. Although WLAN networks allow for packet retransmissions in the event of an unsuccessful transmission attempt, the retransmitted packet must arrive before its playout time or within a specified delay constraint. If the packet arrives too late for its playout time, the packet is effectively lost.

In IEEE 802.11b WLANs, the access point (AP) is a critical component that determines the performance of the network since it carries all of the downlink transmissions to wireless clients and is usually where congestion is most likely to occur. The AP can become saturated due to a heavy downlink load which results in packets being dropped from its transmission buffer and this manifests itself as bursty losses and increased delays [3]. Such losses and delays have a significant impact on multimedia streaming applications. This situation however need no longer apply following the approval of the IEEE 802.11e QoS MAC Enhancement standard which allows for up to four different transmit queues, known as Access Categories (ACs), with different access priorities [4] allowing the QoS enabled AP (QAP) to provide differentiated service to different applications and enable them to meet their target QoS requirements. The Enhanced Distributed Channel Access (EDCA) mechanism of the IEEE 802.11e standard also defines a transmission opportunity (TXOP) as the interval of time during which a particular QoS enabled station (QSTA) has the right to initiate transmissions without having to re-contend for access. During an EDCA TXOP, a QSTA is allowed to transmit multiple MPDUs from the same AC with a SIFS time gap between an ACK and the subsequent frame transmission [5]. The duration of the TXOP is determined by the value of the TXOP limit parameter.

This TXOP mechanism is particularly suited to video streaming applications. Video streaming is often described as “bursty” and this can be attributed to the frame-based nature of video. Video frames are transmitted with a particular frame rate. For example, video with a frame rate of 25fps will result in a frame being transmitted every 40ms. In general, video frames are large, often exceeding the MTU of the network and results in several packets being transmitted in a burst for each video frame.
where the frequency of these bursts corresponds to the frame rate of the video. A video frame cannot be decoded or played out at the client until all or most of the constituent video packets for the frame are received correctly and on time. The TXOP feature can be used to transmit a burst of video packets corresponding to a single video frame during the allocated TXOP interval.

The TXOP has been investigated in a number of previous works primarily through simulation. Suzuki et. al. [6] have investigated the IEEE 802.11e QoS capabilities through simulation using the default values for the TXOP but do not optimise its value. Kim et al. [7] have used the TXOP limit parameter as a means to provide bandwidth fairness among contending stations. However not all applications exhibit a bursty nature and consequently stations may not need to avail of the TXOP facility to transmit a burst of packets in a transmission opportunity. In [8] the authors describe a cross-layer adaptive video streaming system that adapts the TXOP limit parameter for layered encoded video streaming applications. Such a scheme is dependent on the adaptive capabilities of the end-to-end video streaming system. However, multicast video streaming applications have limited adaptive functionality.

In this paper we show through experimental investigation how the statistical characteristics of the video stream can be used to efficiently dimension the TXOP limit parameter in order to minimise the delay required to transmit a video frame under heavily loaded conditions. We focus on video streaming applications with strict real-time delay constraints and investigate the effects of varying the TXOP limit parameter for a number of different video encoding configurations. We do not assume any client server interaction to dynamically adapt and adjust the video streams since most commercially video streaming applications have limited real-time adaptation capabilities. We show that over-dimensioning the TXOP limit parameter has a negative effect on the competing access categories whilst under-dimensioning the TXOP limit parameter yields little benefit.

The remainder of this paper is structured as follows: Section 2 describes the experimental test bed. Section 3 discusses video streaming and provides an analysis of the video content and encoding configurations used during the experiments. From the distribution of video frame sizes we show how to efficiently dimension the TXOP limit parameter as described in Section 4. In Section 5 we describe the different test cases and configurations used in our experiments. We present experimental results showing the loss rate, packet delay, and frame transmission delay for the different test cases.

II. EXPERIMENTAL TEST BED

To investigate the use of the 802.11e TXOP mechanism for video frame transmission, the video server was set up on the wired network and streamed video to a wireless client via the QAP (Figure 1). The QAP used was the Cisco Aironet 1200 using the firmware version IOS 12.3(8)JA which allowed us to access the 802.11e/WMM capability of the device [9]. The QAP was configured with a QoS policy where the Differentiated Services Code Point (DSCP) values in the IP header are used to apply a particular Class of Service (CoS) to the incoming packets. Each CoS is then mapped to a particular AC where the CWmin, CWmax, AIFSN and TXOP limit parameters can be configured. In the experiments reported here only the TXOP limit parameter is varied and the parameters CWmin, CWmax, and AIFSN were fixed with the original IEEE 802.11b settings.

The video streaming server consists of a modified version of RTPSender [10]. RTPSender reads from an encoded video file and identifies the different video frame types, i.e. I, P, or B frames. The frame type indicator is used to set the IP DSCP value of the packets for this video frame. By modifying the IP DSCP value of video packets for the different frame types the QAP can identify the different video frame types and assign them to the appropriate AC so that they can receive differentiated service as defined by the QAP QoS policy. Both the video client and server used the packet monitoring tool WinDump [11] to log all packets transmitted and received and the clocks of both the client and server are synchronised before each test using Nett ime [12]. However, in spite of the initial clock synchronisation, there was a noticeable clock skew observed in the delay measurements and this was subsequently removed using Paxson’s algorithm as described in [13]. The delay measured here is the difference between the time at which the packet was received at the link-layer of the client and the time it was transmitted at the link-layer of the sender. The background traffic was generated using Distributed Internet Traffic Generator (D-ITG) [14]. The background traffic load had an exponentially distributed inter-packet time with a mean offered load of 5Mbps and an exponentially distributed packet size with a mean packet size of 1024B. The background traffic was transmitted from a wired source station via the QAP to a wireless sink station.

III. VIDEO ENCODING ANALYSIS

In the experiments reported here, the video content was encoded with a number of different encoding configurations using the commercially available X4Live MPEG-4 encoder from Dicas. The video clips were prepared for streaming by creating an associated hint track using MP4Creator from MPEG4IP. The hint track tells the server how to optimally packetise a specific amount of media data. The hint track MTU setting means that the packet size will not exceed in the MTU size. In the experiments reported here the hint track MTU is 1024B for all video content types and encoding configurations. Although the mean packet size is less than the hint track MTU setting since if the video
This video content is approximately 10 minutes in duration and was encoded as MPEG-4 ASP (i.e. I, P, and B frames) with a flexible frame rate (Fr), a specified refresh rate (Rr) of 10 frames indicating the I-frame frequency, CIF resolution and a target CBR bitrate using 2-pass encoding. When it is not possible to achieve the specified target bitrate, a flexible frame rate allows for frames to be preferentially dropped so that the target bit rate can be achieved. The encoder drops video frames in order of their relative priority, i.e., B-frames followed by P-frames and finally I-frames.

Table 1 summarises the characteristics of the encoded video clips used during the experiments for a number of different encoding configurations labeled V1 to V4. The top rows indicate the target encoding bitrate of the stream, the frame rate (Fr) and mean packet size of the video content. The next row shows the mean and standard deviation of the frame size over all frames followed by the Peak-to-Mean Ratio (PMR) of the video frame sizes averaged over all frames. The following rows show the same analysis for the I, P, and B frames respectively.

![CDF of Number Packets per Video Frame](image)

**Figure 2. CDF of Number Packets per Video Frame**

**TABLE I. VIDEO STREAM ANALYSIS FOR DIFFERENT ENCODING CONFIGURATIONS**

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bitrate (kbps)</td>
<td>128</td>
<td>384</td>
<td>512</td>
<td>1000</td>
</tr>
<tr>
<td>Frame Rate (fps)</td>
<td>17</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Mean Packet Size (B)</td>
<td>495</td>
<td>814</td>
<td>861</td>
<td>931</td>
</tr>
<tr>
<td>Mean Frame Size (kb)</td>
<td>7.52</td>
<td>15.44</td>
<td>20.57</td>
<td>39.99</td>
</tr>
<tr>
<td>PMR over All Frames</td>
<td>8.4</td>
<td>7.9</td>
<td>7.4</td>
<td>6.9</td>
</tr>
<tr>
<td>Mean I-Frame Size (kb)</td>
<td>21.90</td>
<td>54.01</td>
<td>68.55</td>
<td>109.86</td>
</tr>
<tr>
<td>PMR over I-Frames only</td>
<td>2.00</td>
<td>2.30</td>
<td>2.20</td>
<td>2.50</td>
</tr>
<tr>
<td>Mean P-Frame Size (kb)</td>
<td>7.58</td>
<td>17.32</td>
<td>23.42</td>
<td>47.80</td>
</tr>
<tr>
<td>StdDev</td>
<td>±4.71</td>
<td>±9.43</td>
<td>±11.75</td>
<td>±17.96</td>
</tr>
<tr>
<td>PMR over P-Frames only</td>
<td>3.40</td>
<td>3.60</td>
<td>3.50</td>
<td>3.80</td>
</tr>
<tr>
<td>Mean B-Frame Size (kb)</td>
<td>4.90</td>
<td>9.85</td>
<td>13.41</td>
<td>28.20</td>
</tr>
<tr>
<td>StdDev</td>
<td>±3.00</td>
<td>±6.80</td>
<td>±9.02</td>
<td>±15.90</td>
</tr>
<tr>
<td>PMR over B-Frames only</td>
<td>5.20</td>
<td>5.20</td>
<td>4.50</td>
<td>4.40</td>
</tr>
</tbody>
</table>

and standard deviation of the frame sizes is used to dimension the TXOP limit parameter.

**IV. DETERMINING THE TXOP LIMIT PARAMETER**

The distribution of the frame size is used to correctly dimension the TXOP limit parameter as it statistically describes the encoding characteristics of the video stream and the time required to transmit the video frame. The time it takes to transmit a single video packet (Tp) during a TXOP interval is related to the packet size in bytes (PSz) and the physical line rate (Rate) which for 802.11b has a maximum value of 11Mbps [15].

\[
T_p = \left( \frac{8 \cdot PSz}{Rate} \right) + (2 \cdot SIFS) + Ack \tag{1}
\]

Np is the number of packets required to transmit the video frame of size FSz and is given by,

\[
N_p = \left( \frac{FSz}{PSz} \right) \tag{2}
\]

Usage of the TXOP is not wasteful since when the AC_VI queue has won a TXOP and has no more packets to send during the TXOP interval, the TXOP limit parameter is an integer value in the range (0,255) and gives the duration of the TXOP interval in units of 32μs. If the calculated TXOP duration requested is not a factor of 32μs, that value is rounded up to the next higher integer that is a factor of 32μs. The maximum allowable TXOP limit is 8160μs with a default value of 3008μs [5].

\[
TXOP_N = \left[ N_p \cdot T_p \right] \tag{3}
\]

Figure 2 shows the CDF of the number of packets required to transmit video frames for the video stream V4 encoded at 1000kbps. It can be seen that the number of packets required to transmit I-frames is significantly higher than for B- or P-frames. However since I-frames have a lower frequency they pull the CDF averaged over all frames only slightly to the right. In contrast B-frames have the highest frequency and pull the CDF of the frame sizes to the left. The solid vertical line shows the mean number of packets required to transmit the video frames while the dashed line shows the mean plus one standard deviation of the number of packets required to transmit the video frames. By dimensioning the TXOP limit parameter based upon the mean number of packets/video frame 60% of video frames can be delivered in a single TXOP which translates to 3%, 26%, and 74% of I, P, and B-frames respectively. However if the mean plus one standard deviation of the frame size is used 92% of video frames can be delivered in a single TXOP which translates into 13%, 81%, and 98% of I, P, and B-frames.
V. RESULTS

A. Experimental Design

In all cases the AC queues were configured with IEEE 802.11b settings for CWmin, CWmax, and AIFSN while the value for TXOP limit parameter is varied. Before video streaming can be optimized using multiple IEEE 802.11e parameters it is important that the behaviour of a single parameter is known under a diverse range of test conditions. The purpose of this is so that the effects of varying the TXOP limit parameter can be observed in isolation. The 802.11e standard defines a number of AC queues into which different traffic streams can be directed: Voice (AC_VO), Video (AC_VI), Best-Effort (AC_BE), and Background (AC_BK). In this work we investigate a number of different scenarios and methods of setting the TXOP limit parameter.

- **Case A**: Only the video stream is being transmitted through an IEEE 802.11b AP. This represents the best case scenario.
- **Case B**: The video stream and 5Mbps of background traffic is being transmitted through an IEEE 802.11b AP. This represents the worst case scenario as both the video and background traffic packets are put into the same queue and must wait for their turn in accessing the medium.
- **Cases C, D, and E**: The video stream is transmitted through the AC_VI queue and 5Mbps of background traffic is transmitted through the AC_BK queue. In Case C both AC queues have IEEE 802.11b settings and a TXOP limit = 0. In Case D the AC_VI queues has TXOP limit parameter value that is related to the mean number of packets required to transmit the video frame ($\bar{N}$) averaged over all frames (ALL) irrespective of frame type i.e. $TXOP_{\bar{N}}^{ALL}$. In Case E the AC_VI queue has a TXOP limit that is related to the mean number of packets plus one standard deviation ($\bar{N} + \sigma$) averaged over all frames (ALL) irrespective of frame type i.e. $TXOP_{\bar{N} + \sigma}^{ALL}$.
- **Case F and G**: The I, P, and B frames of the video stream are transmitted through the AC_VO, AC_VI and AC_BE queues and the background traffic is transmitted through the AC_BK queue with a TXOP limit = 0. The AC queues used for the video frames are configured with a TXOP limit parameter that is related to the number of packets for each frame type where the subscripts I, P, and B refer to the I, P, and B video frames respectively. In Case F the TXOP limit parameter that is related to the mean number of packets ($\bar{N}$) for each frame type i.e. $TXOP_{\bar{N}}^{X}$, $TXOP_{\bar{N}}^{P}$, and $TXOP_{\bar{N}}^{B}$. In Case G the TXOP limit parameter that is related to the mean plus one standard deviation of the number of packets ($\bar{N} + \sigma$) for the different frame types.

For the purposes of comparison Cases A and B represent the best and worst case scenarios respectively. Cases C, D and E use just two AC queues namely the AC_VI and AC_BK queues. Cases F and G utilise the full availability of the four AC queues: AC_VO, AC_VI, AC_BE, and AC_BK under the 802.11e standard.

B. Analysis

For video streaming applications, not only is the end-to-end packet delay important, but also the delay incurred when transmitting the entire video frame from the sender to the client. Video streaming is often described as “bursty” and this can be attributed to the frame based nature of video. Video frames are transmitted with a particular frame rate and are generally large, often exceeding the MTU of the network which results in a number of packets being transmitted in a burst for each video frame. In a WLAN environment, the bursty behaviour of video traffic has been shown to result in a sawtooth-like delay characteristic [16]. Since a video frame cannot be decoded or played out at the client until all or most of the constituent video packets for the frame are received correctly and on time, we consider the end-to-end delay required to transmit the entire video frame, Queuing Frame Transmission Delay (QFTD).

Figure 3 shows the mean QFTD for each video encoded bit rate for each of different test cases averaged over all frames. As expected Case A and B provide the best and worst-case values for QFTD. It can be clearly seen that in Cases D-G that by appropriately tuning the TXOP limit parameter, the QFTD for the video frames can be significantly reduced. Over-dimensioning the TXOP limit parameter causes the AC queue to seize too much bandwidth which results in a deterioration in performance for the other competing traffic streams. Table 2 summarises the mean loss rate, packet delay, and QFTD for the different test cases averaged over all video bit rates for the different test cases. The reduction in QFTD (RQFTD) from the worst case scenario, Case B, is quantified as follows:

$$RQFTD = \left( \frac{QFTD_{CaseB} - QFTD_{Case}}{QFTD_{CaseB}} \right)$$ (4)
differentiating between the individual frame types the mean QFTD for I-frames is reduced by 5ms while the mean QFTD for B-frame is increased by 6ms. A similar effect can be seen in Cases E and G, the I-frame QFTD is reduced by 2ms while the QFTD for B-frames is increased by 5ms. By differentiating between to the constituent frame types the end-to-end video frame transmission delay for I- or P-frames can be reduced. I- and P-frames have a higher priority and a greater impact on the end-user perceived QoS over B frames.

VI. CONCLUSIONS

In this paper, we have experimentally investigated the use of the TXOP facility for streaming real-time video over IEEE 802.11e WLAN networks. Video is a frame-based media where video frames are transmitted from the server to the client at regular intervals that is related to the frame rate of the video. In general, several packets are required to transmit a single video frame. The video frame cannot be decoded at the client until all the packets for the video frame have been received. In this paper we exploit the periodic packet bursts that characterize video streaming applications to reduce the end-to-end transmission delay for video frames.

We have shown that the distribution of video frame sizes can be used to efficiently dimension the TXOP limit parameter such that the 60% of video frames are capable of being transmitted within a single TXOP interval in order to transmit the complete burst of packets corresponding to a single video frame. We showed that by using the mean video frame size to dimension the TXOP limit parameter, the transmission delay for the video frame is reduced by 67% under heavily loaded conditions. By differentiating between the constituent video frame types through transmitting the I and P frames through the VI AC queue and the B-frames through the BE AC queue, there is a performance improvement in terms of reducing the frame transmission delay for the I-frames at the cost of increasing the frame transmission delay for the B-frames. Furthermore by providing prioritized access to the different frame types we can reduce the likelihood of packets relating to I or P frames being lost since these frames have a higher priority and a greater impact on the end-user QoS over B frames.

Work is underway to enable the QAP to dynamically adapt the TXOP limit parameter by monitoring the size of the incoming video packet bursts. We are also investigating the provision of prioritized access to the different frame types through the AIFSN, CWmin, CWmax settings in conjunction with the TXOP limit parameter in order to realize significant performance improvements.

ACKNOWLEDGEMENT

The support of the Science Foundation Ireland, grant 03/IN3/1396, under the National Development Plan is gratefully acknowledged.

<table>
<thead>
<tr>
<th>Case</th>
<th>TXOP Limit</th>
<th>Loss Rate</th>
<th>Mean Pkt Delay (ms)</th>
<th>Mean StdDev</th>
<th>RQFTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>--</td>
<td>0.45</td>
<td>3.43</td>
<td>8.42</td>
<td>3.61</td>
</tr>
<tr>
<td>B</td>
<td>--</td>
<td>3.67</td>
<td>26.38</td>
<td>52.85</td>
<td>25.42</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0.93</td>
<td>19.15</td>
<td>38.14</td>
<td>32.97</td>
</tr>
<tr>
<td>D</td>
<td>TXOP_{N}</td>
<td>0.40</td>
<td>6.70</td>
<td>17.28</td>
<td>2.25</td>
</tr>
<tr>
<td>E</td>
<td>TXOP_{N+\sigma}</td>
<td>0.45</td>
<td>6.52</td>
<td>14.98</td>
<td>3.20</td>
</tr>
<tr>
<td>F</td>
<td>TXOP_{N}</td>
<td>1.10</td>
<td>7.97</td>
<td>17.21</td>
<td>3.33</td>
</tr>
<tr>
<td>G</td>
<td>TXOP_{N+\sigma}</td>
<td>1.09</td>
<td>8.16</td>
<td>16.85</td>
<td>5.11</td>
</tr>
</tbody>
</table>

Figure 4. Mean QFTD for I-, P- and B-Frames
REFERENCES


