Video Quality Assessment Based on Data Hiding for IEEE 802.11 Wireless Networks

Mylène C.Q. Farias[⋆] and Marcelo M. Carvalho[†]
[⋆] Department of Computer Science
[†] Department of Electrical Engineering
University of Brasília (UnB), Brasília, Brazil
{mylene, mmcarvalho}@unb.br

***Abstract*— This paper investigates the quality evaluation of H.264-encoded digital videos when transmitted over IEEE 802.11 wireless networks. To accomplish this task, we use a no-reference video quality metric based on a data hiding technique. The impact of IEEE 802.11 DCF parameters on the quality of H.264-encoded videos is studied through a detailed analytical model for saturated single-hop networks under perfect channel conditions. The numerical results obtained indicate that, in spite of the fact that the minimum contention window size has a significant impact on other network performance metrics (e.g., throughput, delay, and jitter), it is not as relevant as the retransmission limit parameter as far as perceptual video quality is concerned. Such observation leads to the fact that, depending on the number of nodes in the network, one must trade faster video streaming for buffering space by using the retransmission limit parameter as a key parameter control in the design of adaptive multimedia-oriented IEEE 802.11 wireless networks.**

1. INTRODUCTION

As high-speed Internet access has become significantly more affordable to a larger number of users worldwide, and with the ever-increasing interest for on-demand web applications, in particular, video streaming applications, the need for guarantees on the quality of service provided to end users has become a vital concern among network designers and administrators alike. In particular, compared to their wired counterparts, the provisioning of quality of service guarantees for video streaming over wireless networks poses significant challenges due to the intrinsic random nature of wireless channels and the intricate mechanisms of medium access control among users.

In the specific case of digital video streaming, a critical aspect is the feasibility of real-time quality evaluation of the streamed video at the receiver side. Unfortunately, the most accurate way to determine the quality of a video is by measuring it using psychophysical experiments with human subjects [1]. Such experiments are very expensive, time-consuming and hard to incorporate into a design process or any automatic quality of service mechanism. With this in mind, the development of fast algorithms that give a physical measure of the video quality (objective quality metrics) has become an active area of research in the past few years [**?**].

Customarily, quality measurements have been largely limited to a few objective measures, such as the *peak signal-to-noise ratio* (PSNR) and the *total squared error* (TSE). Because the outputs of these metrics do not always correspond well with human judgements of quality, there is an ongoing effort to develop video quality metrics that are able to detect impairments and estimate their annoyance as perceived by human viewers. Most of the quality metrics proposed in the literature are *full-reference* (FR) metrics [2, 3, 4], i.e., metrics that need the original video to compute an estimate of the quality. Requiring the reference video or even limited information about it becomes a serious impediment in real-time video transmission applications. To measure video quality in such applications, it is essential to use a *no-reference* (NR) video quality metric, i.e., a metric that blindly estimates the quality of the video. To date, most of the proposed NR metrics estimate annoyance by detecting and estimating the strength of commonly found artifact signals [5, 6].

Recently, we have proposed an unconventional approach to blindly estimate the quality of a video by making use of data hiding techniques [7]. In this approach, a digital watermark is embedded into the original video at the transmitter side. At the receiver side, an estimate of the quality of the 'host' (received) video is obtained by measuring the degradation of the extracted mark. This degradation is computed by simply taking the *total squared error* (TSE) between the extracted and inserted marks. The less degraded the received video, the smaller the value of the computed TSE. In this case, the proposed system is based on the assumption that both the embedded image mark and the host video degrade at similar rates.

In this paper, we use our NR video quality metric to study the impact of IEEE 802.11 [8] parameters on the quality of H.264-encoded digital videos. For this purpose, we consider a detailed analytical model of the ditributed coordination function of the IEEE 802.11 standard to assess the impact of some the MAC parameters on the quality of H.264-encoded video transmissions. In particular, we investigate the impact of the minimum contention window size, the retransmission limit, and the network size on video quality degradation, for the case of saturated single-hop IEEE 802.11 ad hoc networks.

The paper is organized as follows. In Section 2, we present the no-reference video quality metric based on data hiding techniques. In Section 3, we describe the analytical model for the IEEE 802.11 DCF MAC we use for performance

evaluation, and in Section 4 we present the numerical results regarding the impact of IEEE 802.11 DCF parameters on the quality of H.264-encoded transmitted videos.

2. Video Quality Metric Based on Data Hiding

A. The Embedding Algorithm

Figure 1 depicts the block diagram of the embedding process used by the proposed quality assessment method. The image mark ($\mathbf{m}$), which is a binary image, is embedded in each frame of the video using a spread-spectrum technique [9]. A pseudo-random algorithm is used to generate zero-mean pseudo-noise images ($\mathbf{p}$) whose individual pixel values $p(i,j,k)$ assume values -1 or 1 (the indices i and j correspond to the horizontal and vertical positions, while k stands for the video frame index). A different pseudo-noise image is generated for each frame of the video to avoid temporal summation.

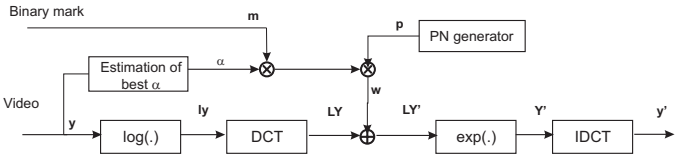

Figure 1. Block diagram of the embedding stage of the video quality assessment system with an automatic estimation of best α.

The final mark $\mathbf{w}$ to be embedded is obtained by multiplying each element $m(i,j)$ of the binary mark image $\mathbf{m}$ by the corresponding elements $p(i,j,k)$ of the pseudo-noise image:

$$w(i,j,k) = m(i,j) \cdot p(i,j,k). \tag{1}$$

Before being added to the mid-frequency DCT coefficients of the frame, the final mark $\mathbf{w}$ is multiplied by a scaling factor α. Then, the logarithm of the luminance $\mathbf{y}$ of the video frame is computed, followed by the DCT transform (denoted by $\mathbf{LY}$):

$$\mathbf{LY} = \mathrm{DCT}(\log \mathbf{y}). \tag{2}$$

The logarithm is used for scaling purposes, since this allows the use of smaller values of α, leading to smaller distortions. The value of α is chosen by an automated system, described in the next section. After the embedding, the DCT coefficients are given by:

$$LY'(i,j,k) = \\ = \begin{cases} LY(i,j,k) + \alpha \cdot w(i,j,k), & 120 \le i,j \le 240, \\ LY(i,j,k), & \text{elsewhere.} \end{cases} \tag{3}$$

After the mark is inserted, the exponential of the video is taken followed by the inverse DCT (IDCT). The video is then coded (compressed) and sent over the communication channel.

B. Automated System for Estimating α

The scaling factor α is used to vary the strength of the mark. An increase on its value increases the robustness of the mark, but also decreases the quality of the video. The appropriate value for α depends on the type of application and video format. The design of an embedding system requires that appropriate values of α be chosen for each video or set of frames.

In Table I, column 5, we present the 'best empirical' α values according to the results of a psychophysical experimental that took into consideration both the visibility and degradation of the extracted mark [10, 7]. As can be noticed from Table I, the α values corresponding to the visibility threshold (α_T, in column 4) and the 'best empirical' α values are not correlated.

In columns 2 and 3 of Table I, we also present the data hiding capacity and the standard deviation (σ) of the set of videos. The data hiding capacity determines how many bits can be hidden in the host video and is given by the following expression [11]:

$$C = 0.5 \cdot \log\left(1 + \frac{\sigma^2_{mark}}{\sigma^2_{video}}\right) \tag{4}$$

where σ^2_{mark} is the variance of the mark embedded and σ^2_{video} is the variance of the (host) video.

TABLE I

Data hiding capacity, standard deviation, α_T intervals for the visibility threshold, and 'best empirical' α values.

Test Seq.	Capacity	σ_{video}	α_T Interval	Best α
Flower	0.009	11.896	$0.1 < \alpha_T < 0.2$	0.100
Bus	0.021	7.671	$0.2 < \alpha_T < 0.3$	0.050
Cheerleader	0.025	7.019	$0.2 < \alpha_T < 0.3$	0.050
Football	0.072	1.269	$0.2 < \alpha_T < 0.3$	0.025
Hockey	0.155	0.827	$0.0 < \alpha_T < 0.1$	0.013

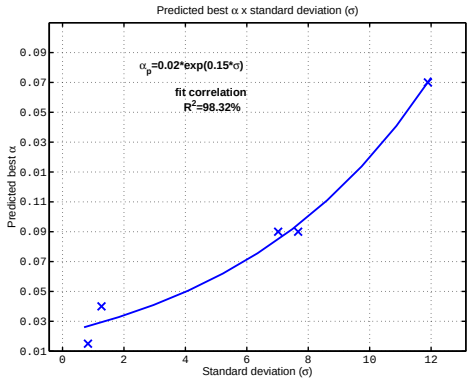

Figure 2. Predicted α versus standard deviation of the host video.

Figure 2 depicts the 'best empirical' α versus the standard deviation for all videos, according to psychophysical data [7]. The following exponential curve was fitted to the data:

$$\alpha_p(\sigma) = a \cdot \exp(b \cdot \sigma), \tag{5}$$

where α_p is the predicted value for α, and σ is the standard deviation of the host video ($a = 0.0162$ and $b = 0.1530$). We can notice from columns 3 and 5 of Table I that the standard deviation and the 'best empirical' α values are correlated. Therefore, the automated system for estimating the value of α

for each video can be implemented by simply measuring the standard deviation of the video frames and using Eq. (5).

C. The Extraction Algorithm

Figure 3 shows the block diagram of the extraction stage of the video quality assessment system. If no errors are introduced by compression or transmission, the input to the extraction stage (Y'') is equal to the output of the embedding stage (Y'). On the other hand, if errors are added $Y'' = Y' + \eta$, where η represents the error signal. In order to explain the mark extraction process, we will assume $Y'' = Y'$. First, the

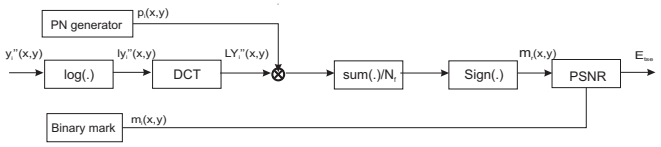

Figure 3. Block diagram of extraction stage of the video quality assessment system.

logarithm of the luminance of the received video (denoted by $\mathbf{y}''$) is computed followed by the computation of its DCT (see Eq. (2)). Then, the mid-frequency DCT coefficients (where the mark was inserted) is multiplied by the corresponding pseudo-noise image, leading to

$$LY''(i,j,k) \cdot p(i,j,k) = LY(i,j,k) \cdot p(i,j,k) + \alpha \cdot m(i,j), \quad (6)$$

for $120 \leq i \leq 240$ and $120 \leq j \leq 240$. Notice that $p(i,j,k) \cdot p(i,j,k) = 1$ because $p(i,j,k)$ is either -1 or +1.

Synchronization is crucial at this step because the image mark can only be extracted if the same pseudo-noise matrix used in the embedding process is used in Eq. (6). Some bits of synchronization information can be easily embedded in the video to assure recovery. The result of Eq. (6) is then averaged over a chosen number of frames N_f. This step is necessary to eliminate the noise (pseudo-noise signal) introduced by the spread-spectrum embedding algorithm. The binary mark is extracted by taking the sign of its average, as follows:

$$m_r(i,j) = \\ = \operatorname{sgn}\left(\frac{1}{N_f}\sum_{k=1}^{N_f}\left(LY(i,j,k)\cdot p(i,j,k) + \alpha \cdot m(i,j,k)\right)\right) \quad (7)$$

Because the pseudo-noise matrix has zero mean, the sum $\sum_{k=1}^{N_f} LY(i,j,k) \cdot p(i,j,k)$ approaches zero for large values of N_f. When errors are added by the compression or transmission systems, $Y'' = Y' + \eta$, and the extracted mark $\mathbf{m}_r$ is an approximation of $\mathbf{m}$.

D. A NR quality metric

A measure of the degradation of the mark is given by the *total squared error* (TSE) of the extracted mark $\mathbf{m}_r$:

$$\mathrm{E_{tse}} = \sum_i \sum_j \left[m(i,j) - m_r(i,j)\right]^2. \quad (8)$$

The less the amount of errors introduced by either processing, compression or transmission, the smaller is the value of $\mathrm{E_{tse}}$. On the other hand, the more degraded the video, the higher the value of $\mathrm{E_{tse}}$. In other words, $\mathrm{E_{tse}}$ gives an estimate of the degradation of the video.

3. The IEEE 802.11 Analytical Model

In this paper, we are interested in undertanding how the quality of H.264-encoded videos degrade when they are transmitted over an IEEE 802.11 wireless network. In particular, we aim at investigating the impact (and, hence, the relative importance) that some of the IEEE 802.11 parameters have on the perceived quality of H.264-encoded videos. By doing so, the design and deployment of IEEE 802.11 wireless netwoks targetted at multimedia content delivery can be better optimized and fine-tuned. For this work, we focus on saturated (i.e., all nodes have a packet to send at any time) single-hop networks under perfect channel conditions, and we consider the ad hoc mode of operation for the IEEE 802.11 DCF MAC [8]. This is because we are mainly interested in understanding the actual impact of specific IEEE 802.11 parameters on the quality of transmitted videos.

A number of analytical models for the IEEE 802.11 DCF MAC have been proposed in the past few years [12, 13]. Considering that we are dealing with saturated single-hop networks under perfect channel conditions, we can make use of a simplified discrete-time Markov model for the IEEE 802.11 DCF MAC based on our previous work [14]. Based on this Markov model, we develop the probability that a node drops a data frame after a certain number of retransmission attempts. This probability is a fundamental figure of merit for our analysis.

Let $b_j(t)$ be the stochastic process representing the backoff time counter for a node j at a time t, and $s_j(t)$ be the stochastic process representing node j's backoff stage $[0,m]$ at time t, for which the maximum window size is $W_i = 2^i W_{\min}$, $i \in [0,m]$. If we assume that each handshake *fails* with a constant and independent probability p_j, regardless of the number of retransmissions experienced, and that a node detects the channel busy with a constant and independent probability g_j, then the process $\{s_j(t), b_j(t)\}$ can be modeled with the discrete-time Markov chain depicted in Fig. 4. In the Markov chain, the only non-null one-step transition probabilities are

$$\begin{array}{ll}
P\{i,k|i,k+1\} = 1-g_j, & k \in [0, W_i - 2],\ i \in [0, M] \\
P\{i,k|i,k\} = g_j, & k \in [1, W_i - 1],\ i \in [0, M] \\
P\{i,k|i-1,0\} = p_j/W_i, & k \in [0, W_i - 1],\ i \in [1, M] \\
P\{0,k|i,0\} = (1-p_j)/W_0, & k \in [0, W_0 - 1],\ i \in [0, M-1] \\
P\{0,k|M,0\} = 1/W_0, & k \in [0, W_0 - 1].
\end{array}$$

The first and second equations indicate that the backoff counter is decremented if the channel is sensed idle (with probability $1 - g_j$), and frozen if the channel is sensed busy (with probability g_j). The third equation indicates that, after an unsuccessful handshake at stage $i-1$, a backoff interval is chosen within the interval $[0, W_i - 1]$ for stage i. The fourth equation indicates that a packet has experienced a successful

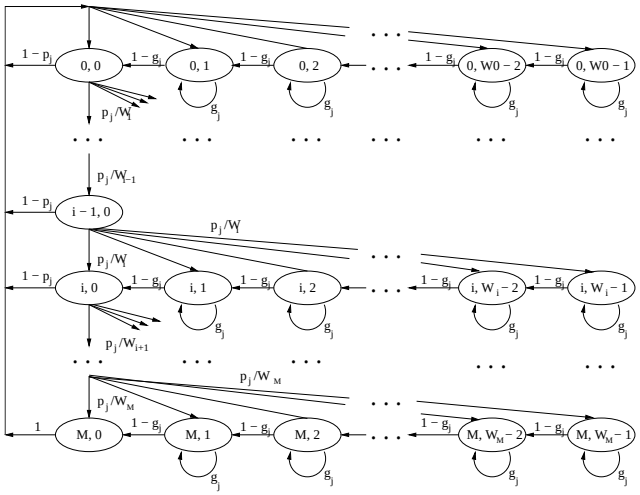

Figure 4. Markov chain model representing the binary exponential backoff algorithm of the IEEE 802.11 DCF MAC.

handshake and a new packet starts at backoff stage 0 with a backoff window size chosen in $[0, W_0 - 1]$. The last equation describes that a new packet starts at backoff stage 0 after either a successful handshake or an unsuccessful handshake (the packet was dropped). Let $b_{i,k} = \lim_{t\to\infty} P\{s(t) = i, b(t) = k\}, i \in [0, M], k \in [0, W_i - 1]$ be the stationary distribution of the Markov chain. We note that

$$b_{i,0} = p_j\, b_{i-1,0} \Longrightarrow b_{i,0} = p_j^i b_{0,0}, \quad 1 \le i \le M. \tag{9}$$

For $k \in [1, W_i - 1]$, we have

$$b_{i,k} = \frac{W_i - k}{(1 - g_j)W_i} \cdot \begin{cases} \sum_{l=0}^{M-1}(1 - p_j)b_{l,0} + b_{M,0}, & i = 0 \\ p_j\, b_{i-1,0}, & i \in [1, M]. \end{cases} \tag{10}$$

From Eq. (9), and noting that $\sum_{l=0}^{M-1}(1-p_j)b_{l,0}+b_{M,0} = b_{0,0}$, Eq. (10) can be rewritten as

$$b_{i,k} = \frac{W_i - k}{(1 - g_j)W_i} b_{i,0}, \quad i \in [0, M],\ k \in [1, W_i - 1]. \tag{11}$$

From Eqs. (9) and (11), all values of $b_{i,k}$ can be expressed as functions of $b_{0,0}$, which can be found from the normalization condition $\sum_{i=0}^{M}\sum_{k=0}^{W_i-1} b_{i,k} = 1$, yielding

$$b_{0,0} = \frac{2(1 - g_j)(1 - p_j)(1 - 2p_j)}{(1 - p_j^{M+1})(1 - 2p_j)(1 - 2g_j) + \kappa W_{\min}}, \tag{12}$$

with $\kappa = (1 - p_j)\left[1 - (2p_j)^{M+1}\right]$ if $m = M$, and $\kappa = 1 - p_j\{1 + (2p_j)^m[1 + p_j^{M-m}(1 - 2p_j)]\}$ if $m < M$. Finally, by taking $\tau_j = \sum_{i=0}^{M} b_{i,0}$, we obtain

$$\tau_j = \frac{2(1 - g_j)(1 - p_j^{M+1})(1 - 2p_j)}{(1 - p_j^{M+1})(1 - 2p_j)(1 - 2g_j) + \kappa W_{\min}}, \tag{13}$$

with κ assuming the previous values depending on whether $m \le M$. Given that we are dealing with a single-hop network under perfect channel conditions, we do not need to treat each node individually. Therefore, we can drop the subscript j from now on.

To find the probability p that a handshake fails, it is sufficient to note that, because of the perfect channel assumption and the fact that all nodes are within range of each other, the probability that a handshake fails is simply the probability that at least one of the $n - 1$ remaining nodes transmit at the same time as the node under consideration. Following the same reasoning, the probability g that the channel is perceived busy by a node during its backoff is the probability that at least one of the $n - 1$ remaining nodes transmit at a given time slot. By the independence assumption, and because each node transmits an RTS frame with probability τ at any time, we have that

$$p = g = 1 - (1 - \tau)^{n-1} \tag{14}$$

It is important to mention that if physical layer aspects are considered, the computation of the probabilities p and g become totally different: detecting that the channel is busy demands the decision whether the energy level perceived by a node is above some target threshold. On the other hand, a successful handshake demands the ability of a node to correctly deliver a frame to its destination, which relies on a number of PHY-layer aspects, such as modulation/demodulation scheme, receiver design, etc [14].

Equations (13) and (14) form a nonlinear system in the unknowns τ, p, and g. Following the approach used in previous works [13, 14], we find an approximate solution to this system by linearizing all equations according to a Taylor series expansion. Considering $q = 1 - p$, we obtain

$$\tau \approx \frac{2}{(W_{\min} + 1)^2} + \frac{2W_{\min}q}{(W_{\min} + 1)^2} - \frac{2(W_{\min} - 1)g}{(W_{\min} + 1)^2}, \tag{15}$$

$$q \approx 1 - (n - 1)\tau, \tag{16}$$

$$g = 1 - (1 - \tau)^{n-1} \approx (n - 1)\tau. \tag{17}$$

Solving the system in Eqs. (15), (16) and (17), we obtain

$$p = \frac{2(n - 1)(W_{\min} + 1)}{(W_{\min} + 1)^2 + 2(n - 1)(2W_{\min} - 1)}. \tag{18}$$

Because each data frame is going to be either successfully transmitted or discarded from the output queue, a key information to network performance is the probability P_{drop} that a data frame is discarded. From the independence assumption regarding the successful handshake across different backoff stages, and considering that each data frame is entitled to a maximum number M of retransmissions, the probability that a data frame is dropped is simply given by

$$P_{drop} = \left[\frac{2(n - 1)(W_{\min} + 1)}{(W_{\min} + 1)^2 + 2(n - 1)(2W_{\min} - 1)}\right]^{M+1}, \tag{19}$$

where the power of $M+1$ stands for the $(M+1)$-th attempt to transmit the data frame at the end of the M-th backoff stage (the counting of backoff stages starts at zero, and we need to count the first transmission attempt before the node starts making retransmissions (see Figure 4)). The probability P_{drop} is key for determining the successful delivery of individual packets of a digital video stream. In the next section, we

investigate the impact of IEEE 802.11 parameters and network size on the perceptual quality of the H.264-encoded videos.

4. Numerical Results

In this Section, we investigate the impact of some of the IEEE 802.11 parameters on the perceptual quality of H.264-encoded digital videos according to the data hiding technique presented in Section 2.

Given that we are dealing with single-hop networks under perfect channel conditions, packet losses will occur only as a result of packet drops due to the backoff mechanism of the IEEE 802.11 DCF MAC. Also, we assume that consecutive packets are dropped independently of one another (with probability P_{drop}), and that no packet is discarded if it is received beyond a certain time constraint (as some decoders would do). Instead, packets are buffered long enough before being displayed.

In simulations, we used 300 frames of 9 publicly-available videos in YUV 4:2:0 qcif format (176×144). The chosen videos contain different amounts of motion, color, and varied content. The results we present next correspond to single transmissions of the "foreman" video sequence over different network scenarios. In Figure 5, we show samples of received video frames when transmitted over different network sizes, for the scenarios under study. It can be noticed that, as network size increases, the visual quality of the frames degrade considerably.

Figure 6 contains the results of $\mathrm{E_{tse}}$ as a function of the maximum number M of retransmission attempts. Because P_{drop} decays exponentially with M, the video quality increases significantly as M increases. This improvement is particularly significant in larger networks: when $n = 40$, the perceptual quality improves by 160% as M increases from 3 to 7, whereas it improves by 50% in the case of 20 nodes (for same variations in M).

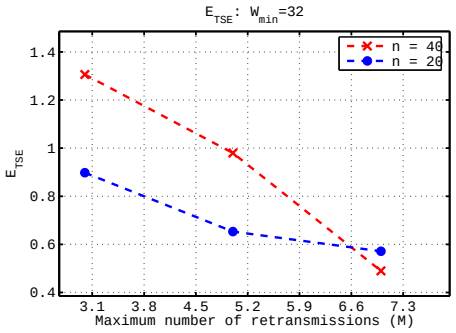

Figure 6. $\mathrm{E_{tse}}$ as a function of the maximum number M of retransmissions ($W_{\min} = 32$).

Figure 7 contains the results for the impact of $W_{\min}$ on the values of $\mathrm{E_{tse}}$ when $M = 5$. As we can see, the perceptual quality of the video is practically unchanged for $W_{\min} > 64$. Such behavior is also expected in larger networks, as Figure 8 illustrates: P_{drop} does not decrease significantly when $W_{\min} > 64$ for other network sizes. Nevertheless, some performance improvement should be expected if $W_{\min}$ increases from 32 to 64, especially in large networks.

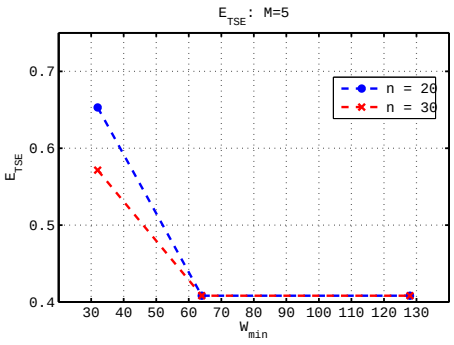

Figure 7. $\mathrm{E_{tse}}$ as a function of the minimum contention window size $W_{\min}$ ($M = 7$).

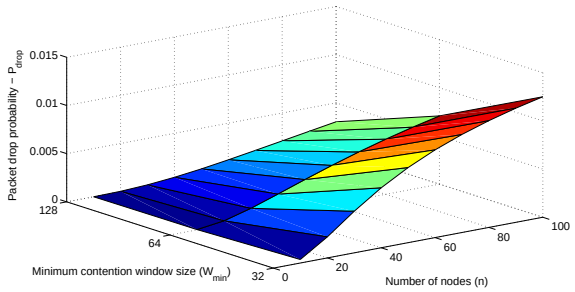

Figure 8. Packet drop probability as a function of $W_{\min}$ and the number of nodes n when $M = 5$.

Regarding the impact of the number of nodes n on the values of $\mathrm{E_{tse}}$, Figures 9 and 10 depict the results for $M = 3$ and $M = 7$, respectively (with the default value of $W_{\min} = 32$). From the fitted curves, the perceived quality degrades by 200% for $M = 3$, and by 50% for $M = 7$, as n increases from 10 to 70. It is interesting to notice that the degradation in quality appears to have an exponential behavior for $M = 3$, whereas it shows an approximate linear behavior for $M = 7$. Such degradation is particularly strong when $M = 3$, in which case the values of $\mathrm{E_{tse}}$ saturate for $n \geq 80$. In those cases, the video cannot simply be played due to significant packet drops.

In Figure 10, we can notice that there is an outlier value for $n = 50$. This is an example of the case when packet losses affect key header information that prevents a significant portion of the video from being correctly decoded. It is important to point out that the appearance of such outliers on numerical results is due to the fact that, as mentioned before, each of the measured $\mathrm{E_{tse}}$ values corresponds to an *instance* of a video transmission over the network (not average values).

Given the previous results, we observe that, although the minimum contention window size $W_{\min}$ has a significant impact on network throughput, delay, and jitter [13], as far as perceptual video quality is concerned, its impact is not as relevant as the one dictated by the retransmission limit parameter

(a) (b) (c) (d) (e)

Figure 5. Sample frames of "foreman" video showing quality degradation when $W_{\min} = 32$, $M = 7$ and network sizes: (a) 20 nodes (b) 40 nodes (c) 60 nodes (d) 80 nodes (e) 100 nodes.

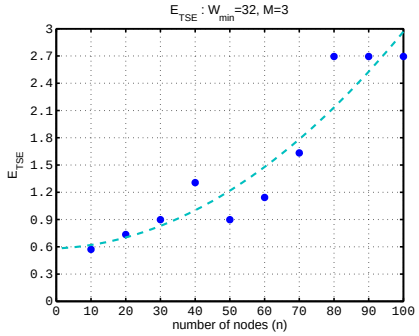

Figure 9. $\mathrm{E_{tse}}$ as a function of number of nodes n ($W_{\min} = 32$, $M = 3$).

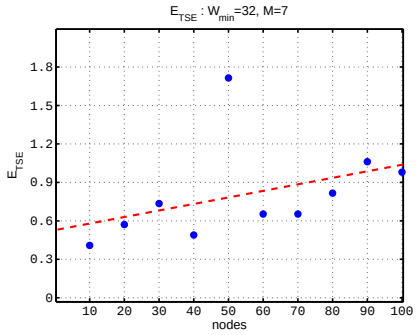

Figure 10. $\mathrm{E_{tse}}$ as a function of number of nodes n ($W_{\min} = 32$, $M = 7$).

of the IEEE 802.11 DCF MAC. Therefore, depending on the number of nodes in the network, one must trade faster video streaming for buffering space at the receiver side. In other words, in order to maintain the same level of video quality, the higher the contention on the network, the higher the value of M and the larger the buffering space needed. Given that, the retransmission limit parameter M appears as a natural candidate for acting as an effective quality control "knob" in multimedia-oriented IEEE 802.11 wireless networks.

5. Conclusions

In this paper, we have investigated the quality evaluation of H.264-encoded digital videos when transmitted over IEEE 802.11 wireless networks. To accomplish this task, we have used a no-reference video quality metric based on a data hiding technique. The impact of IEEE 802.11 DCF MAC parameters on the quality of H.264-encoded videos were studied through a detailed analytical model of saturated single-hop networks under perfect channel conditions.

From the numerical results, we have observed that, in spite of the fact that the minimum contention window size has a significant impact on other network performance metrics (e.g., throughput, delay, and jitter), it is not as relevant as the retransmission limit parameter as far as perceptual video quality is concerned. Such observation leads to the conclusion that, depending on the number of nodes in the network, one must trade faster video streaming for buffering space by using the retransmission limit parameter as a key parameter control in the design of adaptive multimedia-oriented IEEE 802.11 networks.

References

[1] ITU Recommendation BT.500-8, "Methodology for the subjective assessment of the quality of television pictures," 1998.

[2] A. B. Watson, Hu James, and J. F. McGowan, "Digital video quality metric based on human vision," *Journal of Electronic Imaging*, vol. 10, no. 1, pp. 20–9, 2001, Publisher: SPIE-Int. Soc. Opt. Eng, USA.

[3] S. Wolf, M. H. Pinson, S. D. Voran, and A. A. Webster, "Objective quality assessment of digitally transmitted video," in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, Victoria, BC, Canada, 1991, pp. 477–82 vol.

[4] S. Winkler, *Vision Models and Quality Metrics for Image Processing Applications*, Ph.d. dissertation, Ecole Polytechnique Federale de Lausanne, 2000.

[5] H.R. Wu and M.; Yuen, "A generalized block-edge impairment metric for video coding," *IEEE Signal Processing Letters*, vol. 4, no. 11, pp. 317–320, 1997.

[6] J. Caviedes and J. Jung, "No-reference metric for a video quality control loop," in *Int. Conf. on Information Systems, Analysis and Synthesis*, 2001, vol. 13.

[7] MCQ Farias, M. Carli, and SK Mitra, "Objective video quality metric based on data hiding," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 3, pp. 983–992, 2005.

[8] *IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Nov 1997, P802.11.

[9] I. Cox, J. Kilian, F. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. on Image Processing*, vol. 6, no. 12, 1997.

[10] M.C.Q. Farias, M. Carli, J.M. Foley, and S.K. Mitra, "Detectability and annoyance of artifacts in watermarked digital videos," in *Proceedings of XI European Signal Processing Conference*, France, September 2002.

[11] M. Barni, F. Bartolini, A. De Rosa, and A. Piva, "Capacity of full frame dct image watermarks," *IEEE Transactions of Image Processing*, vol. 9, no. 8, pp. 1450–5, 2000.

[12] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, March 2000.

[13] M. M. Carvalho and J. J. Garcia-Luna-Aceves, "Delay analysis of IEEE 802.11 in single-hop networks," in *Proc. 11th IEEE International Conference on Network Protocols (ICNP)*, Atlanta, USA, November 2003, pp. 146–155.

[14] M. M. Carvalho and J. J. Garcia-Luna-Aceves, "Modeling wireless ad hoc networks with directional antennas," in *Proc. IEEE 25th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, Barcelona, Spain, April 2006.