Application of Traffic Shaping for Quality-Preserving in Video Streaming over IP

CARLOS MARCELO PEDROSO
Federal University of Parana
Department of Electrical Engineering
Centro Politcnico, Curitiba PR
BRAZIL
pedroso@eletrica.ufpr.br

VALTER KLEIN
Department of Electrical Engineering
Pontifical University of Parana
Rua Imaculada Conceicao 1155, Curitiba PR
BRAZIL
valter.klein@pucpr.br

Abstract: In this paper, we propose the use of traffic shaping to mitigate the effects of network impairments in viewer's perceived quality of video streaming over IP. Traffic shaping is used to change the burstiness of video considering the characteristics of MPEG-4 encoding. In MPEG-4, the bursts of traffic are caused by variable length of frames. I-frames are very important in image reconstruction and produce the biggest bursts, so the packets carrying the I-frames are more likely to experience higher delay or be discarded. We propose a method of shaping the video traffic to distribute the bursts over time as much as possible. This procedure reduces the negative effects of bursts on the viewer's perceived quality. We choose to use the token bucket algorithm due to its low computational complexity and wide availability in servers and routers. The main contribution of the paper is a method to configure the token bucket parameters using the video characteristics. The efficiency of the proposed method is demonstrated through computer simulations, and the results indicate that the proposed method effectively mitigates the effects of network impairments in viewer's perceived quality.

Key–Words: Multimedia Communication, Video Streaming, Viewer-perceived Quality, Traffic Shaping.

1 Introduction

The Internet Protocol (IP) has proven to be an interesting technological solution for transporting data of multimedia applications, supporting different types of real-time transmission. Video streaming plays a major role among those applications, and the use of IPTV (television over IP) is consistently growing over the past years. A network to support IPTV systems must provide appropriate levels of QoS (Quality of Service) and QoE (Quality of Experience) [5] and is required to transfer a large volume of information.

The MPEG (Moving Picture Experts Group) was established to set standards for audio/video compression and transmission [26]. The MPEG-4 part 10, also known as H.264 or Advanced Video Coding (AVC), is specified in ISO/IEC 14496 [1] and was developed by ITU-T Video Coding Experts Group together with ISO/IEC Moving Picture Experts Group. The H.264 can achieve good video quality at lower bit rate than previous standards, including MPEG-4 part 2 [24], and is widely used in IPTV systems [18].

The structure of MPEG-4 contains three different types of frames, namely I-, P-, and B-frames. A specific arrangement of these frames is called group of pictures (GOP). A GOP always starts with an I-frame, followed by B- and P-frames. I-frames are only spatially compressed, P-frames depend on previous I-frame to be decoded, and the B-frames depend on past and future I- and P-frames, resulting in a variable bit rate behavior.

The effects of packet loss in viewer's perceived quality were studied by Greengrass et al. [14]. The authors showed that the discard of packets carrying I-frames could result in impairments propagated over all frames of a GOP. As a consequence of packet loss, degradation can last for a long period of time, typically $1/2$ to 1 second, and video quality would be retrieved only when the decoder receives a new unimpaired I-frame. Depending on which packet is lost, the distortion could result in different levels of image degradation. For example, the loss of a packet at the start of an I-frame, containing part of the frame header, can have the same effect as losing the I-frame entirely. The packet loss can occur due to lack of space in the buffer or because of expiration of the packet maximum delay. According to [6], network delay has become a determining factor in the quality of video streaming and voice over IP, as the bandwidth seems to be growing to meet the requirements of video transmission. Delay and jitter must be lower than 200 milliseconds and 50 milliseconds, respectively, to ensure a good quality of experience in video streaming encoded with MPEG-2 and MPEG-4 [2].

Hong et al. in [15] proposed a packet scheduler algorithm to improve the viewer's perceived quality, adjusting the time intervals between packets based on their significance. The significance is defined as the importance of a packet to the image reconstruction considering the GOP structure, and it is obtained through analyses of the consequence of loss for each pixel. The evaluation of significance requires payload processing, with high computational complexity. The scheduler is called Significance-Aware Packet Scheduler (SAPS). With SAPS, the packets with higher significance will take a longer inter-packet time interval than the less significant packets. This allows network elements to free up some space on their buffers before the next packet arrival. The most significant packets wait a longer time to be transmitted and are likely to be preserved in case of network congestion. As a result, the video quality perceived by the viewers is improved. The implementation of SAPS is done at the streaming server.

Huang et al. in [16] presented an algorithm that combines packet scheduling and queue management to improve video quality. The authors developed an algorithm called Active Drop Queue (ADQ) to improve the quality of video stream over networks with bandwidth constraints. ADQ implements separate queues for the traffic within the bandwidth limit, for the traffic exceeding the bandwidth limit, and for the best-effort traffic. Each packet in queue is associated with a time stamp. This allows the evaluation of an excess delay beyond a specified deadline where the transmission of packet is useless. In case of network congestion, the ADQ removes from traffic exceeding the bandwidth limit the packets with expired deadline, freeing the queue to receive new packets.

An active queue management algorithm based on priority dropping (PD) and a proportional-integral-derivative (PID) controller is proposed in [27] using the control theory, called PID_PD, which first drops the least important packets when network congestion arises. The packet is marked by the application layer using the priority field of the IP packet. The B-frames receive the least priority, and I-frames receive the higher priority. The results show that the schema can prevent the high-priority layer or frame from dropping, thus preserving viewer perceived quality in case of network congestion.

Schier et al. presented in [20] a strategy to preserve video quality in situations of network congestion. The method performs a classification of packets based on macroblock distortion estimation and requires superficial decoding of the video stream. The results showed an increase of perceived video quality.

The problem of shaping video traffic was explored previously in by Alam et al. [9], but the objective of the authors was to adapt the video traffic to the policing strategy of the Integrated Services model [10].

A priority packet discard strategy to avoid discarding packets that carry the most relevant information for image reconstruction was proposed by Maffini et al. [19]. The strategy uses a real-time packet payload classifier implemented with artificial neural networks. The results showed an increase in the viewer's perceived quality in cases of network congestion. This approach does not require changes at the video source to classify outgoing packets or video streaming decoding.

In this paper, we explore the application of traffic shaping implemented in the video streaming server, with the aim of mitigate the effects of network impairments in quality of video as perceived by the final user. The token bucket algorithm [23] was chosen to shape video traffic because of its low computational complexity and wide availability. The main contribution of the paper is a method for setting the token bucket parameters, considering the video characteristics. The proposed method has lower computational complexity than SAPS.

This article is organized as follows. Section 2 presents a method to use the token bucket shaper to improve the viewer's perceived quality of video streaming over IP, including an automatic method for parameter setting. Section 3 presents the evaluation of proposed method using objective video quality metrics and computer simulations. Section 4 presents the conclusions and future work.

2 Traffic Shaping to Improve Perceived Video Quality

The traffic shaper can change the temporal characteristics of outgoing packets [13]. Thus, the sent traffic can be adapted to traffic policing used by the network elements to check if the user has not violated the service-level agreement. However, our proposal is to use traffic shaping to change the temporal characteristic of video stream packets to improve user-perceived quality.

The token bucket algorithm has two parameters: the flow average rate r and the associated burstiness b. In the token bucket algorithm, a token is added to the bucket every $1/r$ seconds. The bucket can hold b tokens at the most. If a token arrives when the bucket is full, it is discarded. When a packet of n bytes arrives, n tokens are removed from the bucket, and the packet is sent. In the token bucket shaper, if fewer than n tokens are available, packet transmission is postponed until n tokens are available in the bucket.

2.1 A Method to Set Token Bucket Parameters

The aim is to distribute traffic bursts on time, thereby reducing the problems caused by queuing bursts of packets in the buffers of network elements. As a result, we expect to improve the viewer's perceived quality. The main problem is how to set the parameters of the algorithm, namely the bucket rate r and bucket size b. The algorithm assures that, in a given time interval δ, the maximum amount of information sent does not exceed $b+\delta r$. Thus, the r parameter can be set to the average bit rate required by the video. This can be done by taking the ratio between the amount of information to be transmitted and the total length of the video. For the bucket size, some additional considerations are necessary. If bucket size is small, the bursts would be completely eliminated, and the delay added to the frames could impair the video quality. If bucket size is too large, the packet flow would remain unchanged. Thus, to preserve video quality, it is necessary to ensure that the transmission of the I-frames, which is the largest frame, can be done on the average before transmission of the next image, namely

$$t_{\tilde{\phi}_I} \leq \frac{1}{f}, \tag{1}$$

where $t_{\tilde{\phi}_I}$ is the average time required to send an I-frame and f is the frame rate.

The average time to send an I-frame, shaped by token bucket algorithm, can be written as

$$t_{\tilde{\phi}_I} = \frac{b}{C} + \frac{\tilde{\phi}_I - b}{r}, \quad r < C, \tag{2}$$

where b is the bucket size, C is the channel rate, $\tilde{\phi}_I$ is the average size of an I-frame, and r is the token bucket rate.

Then we substitute Eq. (1) and (2) to yield

$$b \leq \frac{1}{r/C-1}\left[\frac{r}{f} - \tilde{\phi}_I\right], \quad r < C. \tag{3}$$

Eq. (3) indicates the range for the bucket size so as not to overlap the transmission of packets belonging to subsequent frames. Considering that I-frames are the larger frames, then the use of a bucket size greater than the size of the biggest I-frame of video would not shape traffic. So that an important rule to be established is given by

$$b < \mathrm{MAX}_{\phi_I}, \tag{4}$$

where MAX_{ϕ_I} represents the larger I-frame of the video.

Based on the above considerations we propose the use of the following heuristic in choosing the value of bucket size:

$$b = \min\left(\frac{1}{r/C-1}\left[\frac{r}{f} - \tilde{\phi}_I\right], \tilde{\phi}_I\right). \tag{5}$$

The method proposed in Eq. (5) complies with the limits established by Eq. (3) and (4), as the average size of I-frames is always lower or equal than the maximum size of I-frames. This heuristic is based on the following observations: (a) the average size of I-frames is a simple measure to be achieved in practice; (b) as there is a variation in the size of I-frames, limiting the burst size to the average size of I-frames is an interesting measure to reduce network congestion without excessively penalizing the transmission of I-frames; and (c) shaped traffic also allows bursts for majority of video frames, in a way that we expect that this configuration does not lead to degradation of perceived quality, as only the larger frames will be affected.

2.2 An automatic method for parameter setting

Observing the behavior of traffic at a certain scale, it is possible to infer on the parameters for evaluation of b using Eq. (5). The parameters are the frame rate f, the average video rate r, the average size of I-frames $\tilde{\phi}_I$, and the largest I-frame MAX_{ϕ_I}. This facilitates the operation of the proposed method for the end user.

Assume Y_t, $t = 1, 2, ..., N$, denotes a temporal series representing arrived traffic, with the interval between Y_t and Y_{t+1} being fixed and constant time given by τ. Define X_t as the total traffic observed in an interval given by $X_t = Y_t - Y_{t-1}$. Thus, we propose to use X_t to infer the parameters for evaluation of b.

2.2.1 Frame rate

We propose using a trend detector to identify the frame rate, which is called the Moving Average Convergence Divergence (MACD) [12]. The MACD is a method originally developed in forecasting time series in financial markets. The trend indicator uses an Exponential Moving Average (EMA), which is actually a low pass filter, given as follows:

$$S_t = (1-\alpha)S_{t-1} + \alpha X_t \quad (6)$$

where α is the value of the filter constant.

The MACD uses two EMAs with values of α chosen to get short- and long-term EMAs. A trend to ascent or descent in the series is indicated by the intersection of the short- and long-term moving averages. When the MACD is used in the financial stock market, trend detection is a trigger used to buy and sell stocks. In the case of intended application, only upward trends are of interest. In the proposed method, if an upward trend is detected, it is considered indicative of the presence of a frame. Figure 1 illustrates the application of the method using video SW1 with $\tau = 10$ milliseconds. The figure shows a "$\times$" symbol indicating the detection of th up trend. Continuous line represents a short-term EMA with $\alpha = 0.5$ and dashed line indicates a long-term EMA with $\alpha = 0.2$. In this case, 12 up trends were detected in a total sample time of 400 milliseconds, leading to a frame rate of $f = 12/0.4 = 30$ frames per second.

2.2.2 $\tilde{\phi}_I$ and MAX_{ϕ_I}

By appropriately choosing the values of α, identifying all frames or only the larger frames using MACD is possible. I-frames are the larger frames of GOP [19]. Thus, using the MACD can possibly identify only the I-frames. In the tests, we used the values of $\alpha = 0.1$ for short-term EMA and $\alpha = 0.6$ for long-term EMA. As there is only one I-frame in GOP, it is possible to estimate the approximate size of the I-frame. If an I-frame was detected in a time t, the approximate size of the I-frame can be found using $y(t - 1/(2f)) - y(t + 1/(2f))$. From this, it is possible to evaluate $\tilde{\phi}_I$ and MAX_{ϕ_I}.

2.2.3 Video rate

The r parameter can be obtained through a long-term EMA. After a warm-up period, typically of two seconds, the value of long-term EMA converges to a stationary value representing the video rate. Figure 2 represents a long-term EMA with $\alpha = 0.01$ for video Coastguard with $\tau = 10$ milliseconds. It can be seen that the series converges in approximately 2.5 seconds. After a warm-up period, it is possible to evaluate video bit rate to $r = 8(1150/0.01) = 920$ Kbps.

3 Simulation

The simulation was configured using a dumbbell topology, as illustrated in Figure 3. The link between routers is the bottleneck of the network with transmission rate given by C. The link between server-router and client-router was configured with a transmission rate much greater than C, with negligible propagation delay. Delay of the link between routers was defined to mimic the typical latency found in real systems for access networks, with average value of 50ms [6]. The simulator used was the *Network Simulator* (NS-2), version 2.38 [11]. In all simulations, we used two scenarios for background traffic: first consisting of a file transfer over TCP, and second consisting of a video streaming.

In the tests, the bottleneck rate C was changed to produce several levels of utilization. The utilization is a dimensionless value with a maximum size of 1.0, given by the ratio between the link usage by video transmission and C. We simulated several scenarios of utilization, contemplating possibilities from 10% to 90%. The buffer size of routers, client, and server was large enough to avoid packet discard at queues. In each simulation, the bucket size was recalculated using Eq. (5). The receiver uses a playout buffer with size of 200ms.

3.1 Video test set

Two sets of videos were used for the tests. The first were Akiyo, News, Highway, Football, Foreman, and Coastguard, which are all publicly available in [7]. These videos consisted of a single scene, each with a different spatial detail and movement patterns. Figure 4(a)-(f) illustrates the first scene of each of the

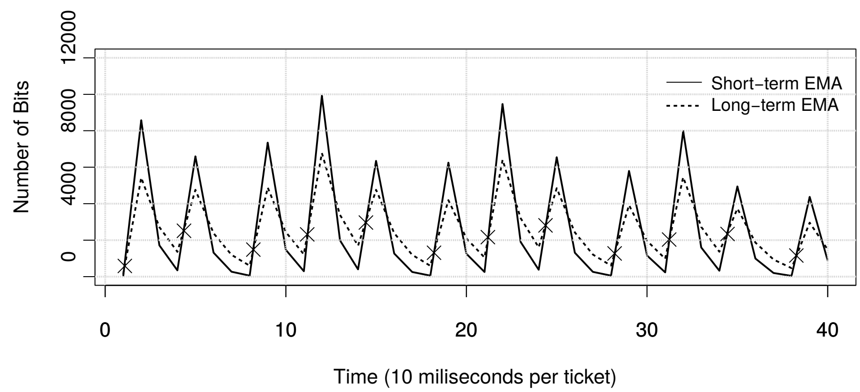

Figure 1: Frame rate detection for video SW1.

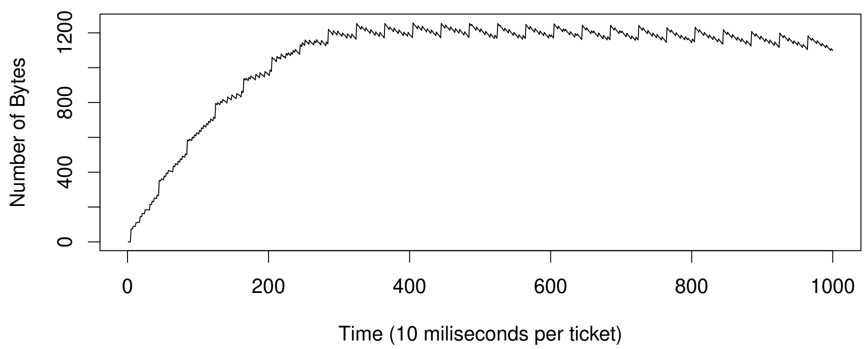

Figure 2: Video rate detection - video Coastguard.

videos. These videos were encoded with resolution of 352x288 pixels.

To allow better generalization of the results, snippets of 120 seconds of films Jurassic Park, Silence of the Lambs, and Star Wars Episode IV were also used. We used one snippet from Jurassic Park (JP) and Silence of the Lambs (SL), and three snippets from Star Wars Episode IV (SW1, SW2, and SW3). Those videos present different movement patterns and scene changes. Particularly, SW1, SW2, and SW3 can be classified by visual inspection as presenting a medium, high, and low movement pattern, respectively. JP and SL present medium movement pattern. All film snippets were encoded with resolution of 640x480 pixels. Figure 5(a)-(e) illustrates the first scene of each of the videos.

The videos were encoded with H.264/AVC and MPEG-4 part 2, with GOP lengths of 8, 12, 18, and 30 frames. In all the videos, the frame rate was configured to 25 fps.

3.2 Evaluation of automatic method for parameter setting

The automatic method for parameter setting proposed in section 2.2 was tested using the videos under study. The parameters f, r, $\tilde{\phi}_I$, and MAX_{ϕ_I} were estimated for each video. The error was evaluated using the difference between real and estimated values.

Table 1 shows the error of each parameter after the automatic parameter detection. The results indicate a good estimation for videos Akiyo, Coastguard, Football, Foreman, Highway, and News. These videos consisted of a single scene, and the motion pattern is

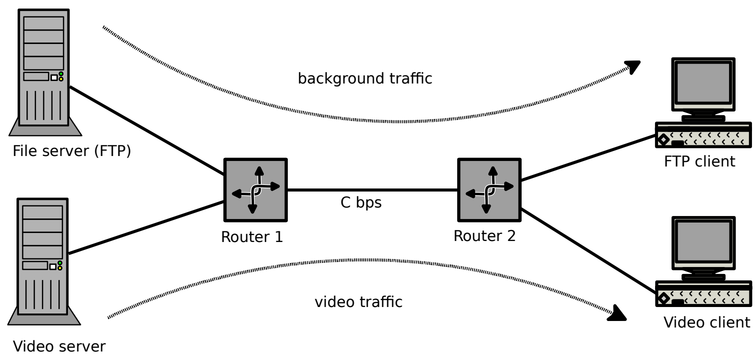

Figure 3: Simulation topology.

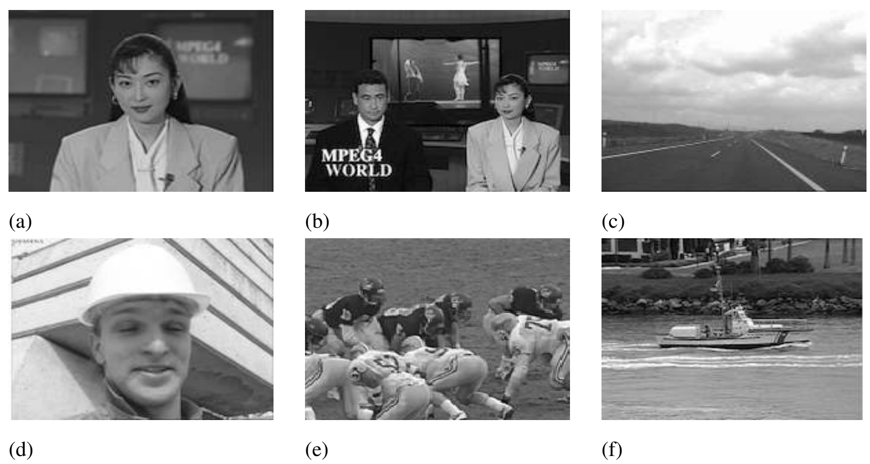

Figure 4: First frame of single-scene videos used in the tests.

the same over all video length. The parameter estimation for film snippets JP, SL, and SW3 also presents good results. The parameters estimated for SW1 and SW2 present a greater error, particularly in the estimation of $\tilde{\phi}_I$ and MAX_{ϕ_I}. This can be explained by the movement pattern of the SW1 and SW2, with a lot of scene changes. The error presented in Table 1 was evaluated using the complete video. Thus, the grater error in this case does not indicates a poor performance of proposed method because this video actually experienced variation of its characteristics over time. The automatic method for parameter setting can actually adapt to this behavior.

Table 1: Average error after automatic parameter estimation.

Video	r	fps	$\tilde{\phi}_I$	MAX_{ϕ_I}
Akiyo	8.68%	0.59%	0.00%	0.00%
Coastguard	5.33%	0.25%	9.59%	17.51%
Football	22.52%	0.58%	49.52%	0.00%
Foreman	5.37%	0.59%	0.00%	0.00%
Highway	0.32%	0.04%	47.33%	27.03%
News	7.51%	0.25%	0.00%	0.00%
JP	0.97%	1.67%	32.91%	6.74%
SL	0.08%	0.83%	9.74%	21.54%
SW1	0.33%	2.60%	81.93%	199.19%
SW2	87.96%	1.28%	99.01%	132.96%
SW3	1.45%	1.44%	0.15%	26.23%

3.3 Evaluation of Video Quality

Quality of experience is defined by the International Telecommunication Union as a subjective measure using the user's perception on the quality of the video

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

Figure 5: First frame of snippets from films used in the tests.

image reproduced. Subjective evaluation of QoE is standardized by [3] and [4]. The test result is called Mean Opinion Score (MOS), which is given by a number in the range of 1-5 that indicates the perceived quality, with 1 being the worst and 5 being the best quality.

In contrast, the objective evaluation of the quality of the received video is performed using computational tools. The available methods for objective evaluation of quality can be classified as No-Reference (NR), Reduced-Reference (RR), or Full-Reference (FR) approaches. The FR method accesses the original and received content in the evaluation process. The NR approach assesses the content quality level without any knowledge of the original video, while RR uses some information of the original video to evaluate the quality of the received video.

Peak-Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) are the most frequently used FR metrics. According to [21], PSNR is a basic, yet, important metric that assesses the similarity between two different images. PSNR computes the mean square error (MSE) of each pixel between the original and the received images, represented in dB. Images with more similarity will result in higher PSNR values. The PSNR is calculated frame by frame using the MSE given by

$$\mathrm{MSE} = \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c}\left[X_o(i,j) - X_r(i,j)\right]^2 \tag{7}$$

where r and c represent the number of rows and columns of image, respectively, and $X_o(i,j)$ and $X_r(i,j)$ represent the luminance of pixel (i,j) of original and received frames, respectively. The PSNR is evaluated using

$$\mathrm{PSNR} = 20\log_{10}\left(\frac{\mathrm{MAX}_I}{\sqrt{\mathrm{MSE}}}\right) \tag{8}$$

where MAX_I represents the maximum value of pixel intensity. For the videos in consideration, $\mathrm{MAX}_I = 255$.

The SSIM was developed based on the human psycho-visual system to detect some human perceptible video disruptions [25]. To address this problem, SSIM combines luminance, contrast, and structural similarity of the images to compare the correlation between the original image and the received one [21]. The test result is a number between 0 and 1 that represents the video quality. Quality evaluation of SSIM has proved superior to PSNR in some ways [22].

The assessment of quality in the experiments used both methods, PSNR and SSIM. The evaluation was performed with EvalVid tool [17], with modifications to adapt to the problem. Table 2 shows the mapping between MOS, PSNR, and SSIM, and it is a good reference to understand how the variation of PSNR and SSIM impairs the quality perceived by users.

3.4 Simulation Results

Figure 6 shows the quality evaluation using SSIM of videos Highway, Coastguard, News, SW1, SW2, SW3, SL, and JP. The results using the proposed method are marked as “shaped”. Tables 3 and 4 present the values of quality evaluation using PSNR

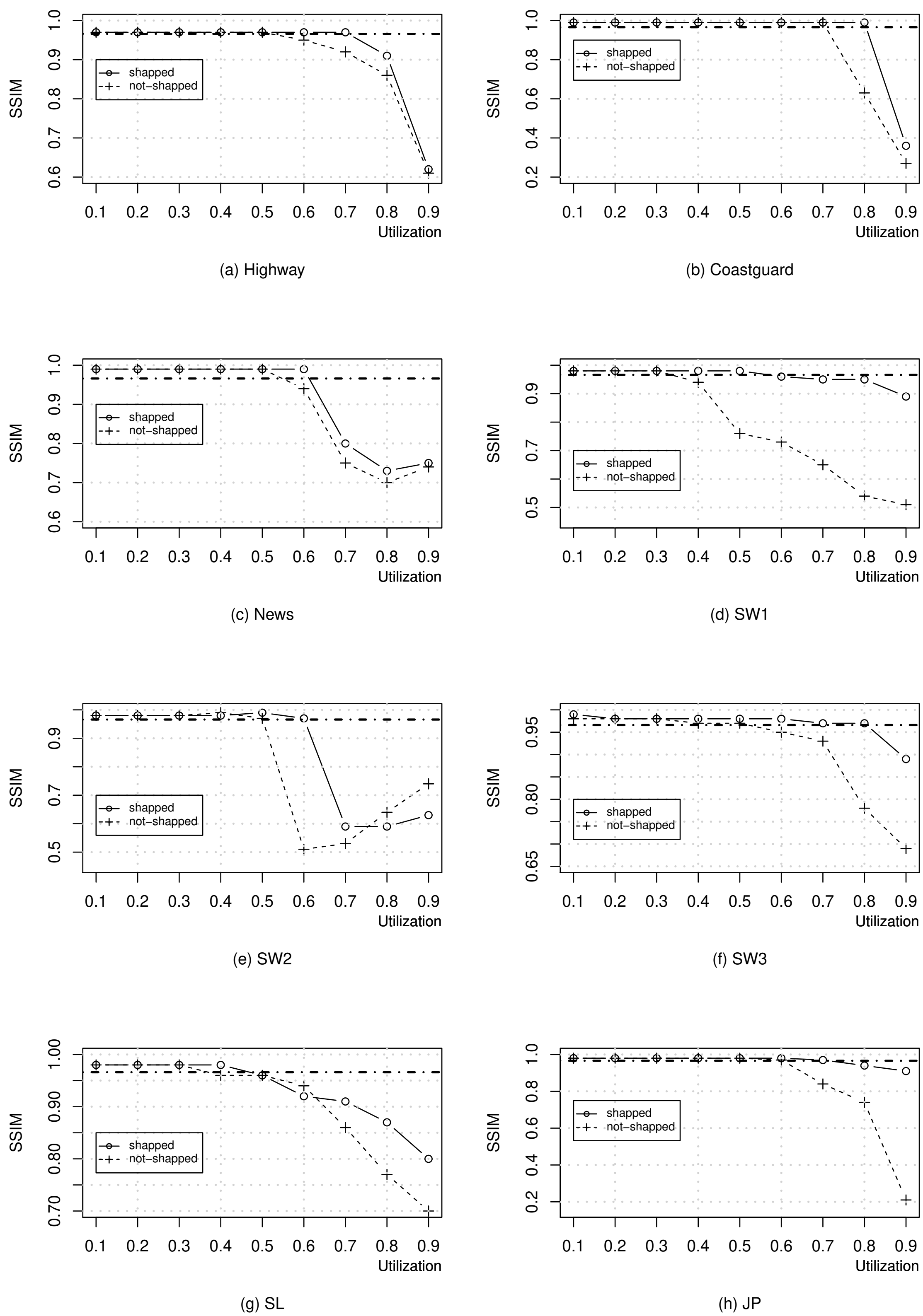

Figure 6: Simulation of video streaming encoded with MPEG-4 part 10 using the video News as background traffic for (a) Highway, (b) Coastguard, (c) News, (d) SW1, (e) SW2, (f) SW3, (g) SL, and (h) JP.

Table 2: Mapping between PSNR, SSIM and MOS compiled from [17] and [25]

PSNR	SSIM	MOS	Meaning
$PSNR > 37dB$	$SSIM > 0.995$	5	Excellent
$31dB < PSNR \leq 37dB$	$0.99375 < SSIM \leq 0.995$	4	Good
$25dB < PSNR \leq 31dB$	$0.965 < SSIM \leq 0.99375$	3	Fair
$20dB < PSNR \leq 25dB$	$0.875 < SSIM \leq 0.965$	2	Poor
$PSNR \leq 20dB$	$SSIM \leq 0.875$	1	Bad

and SSIM, respectively. The graphs also show a horizontal dashed line indicating that the threshold of the SSIM metric for equivalent quality of MOS was equal to 3. All videos and movie snippets were encoded with H.264/AVC, with GOP length of 12 frames. In this test, we used a background traffic consisting of a video streaming. The video News was used to produce the background traffic. To avoid synchronization, the starting time of transmission was chosen randomly between 0 and 1/25 seconds. It was observed that as the level of utilization increases the viewer-perceived quality reduces. The use of the proposed method result in a good quality even at higher utilization levels. This behavior was observed in all videos. However, the level of utilization where degradation starts was different in each case. This happened because the videos had different patterns of movement.

Table 5 shows the quality evaluation using SSIM for video streaming encoded with MPEG-4 part 2 using as background traffic a file transfer over TCP (Transmission Control Protocol), at several utilization levels, for videos Coastguard, Highway, and News. The videos under test were encoded with GOP lengths of 8, 12, 18, and 30 frames. Table 6 shows the corresponding quality evaluation using PSNR. In this test, the quality was evaluated by comparing the raw video in YUV format with the received video. The TCP is a friendly protocol, with its congestion control mechanism reducing the data transmission rate according to network conditions. Congestion control uses TCP Reno [8], which is one of the most used TCP implementations. Although friendly, it is expected that the background TCP traffic impairs the video quality due to variation in delay caused by TCP. The results show that the use of the proposed method benefits the quality of video streaming in all cases. It is also important to note that the algorithm benefited the quality of all GOP lengths in the study, that is, the variation of GOP does not affect significantly the performance of the proposed method.

Table 7 shows the quality evaluation using SSIM for video streaming encoded with H.264/AVC using as background traffic a file transfer over TCP (Transmission Control Protocol), at several utilization levels, for videos Coastguard, Highway, and News. Table 8 shows the corresponding quality evaluation using PSNR. In this test, the quality was evaluated by comparing the raw video in YUV format with the received video. The videos were encoded with GOP lengths of 8, 12, 18, and 30 frames. The video News was used to produce the background traffic. The results show a modest increase in quality using the proposed method. Comparing these results with videos encoded with MPEG-4 part 2, it is observed that quality degradation starts at lower utilization levels. In this case, the GOP length does not affect the performance of the proposed method.

Table 9 shows the quality evaluation using SSIM for video streaming encoded with H.264/AVC using as background traffic a video streaming, at several utilization levels, for videos Coastguard, Highway, and News. The videos were encoded with GOP lengths of 8, 12, 18, and 30 frames. Table 10 shows the corresponding quality evaluation using PSNR. In this test, the quality was evaluated by comparing the raw video in YUV format with the received video. The video News was used to produce the background traffic. To avoid synchronization, the starting time of transmission was chosen randomly between 0 and 1/25 seconds. The utilization was evaluated using the ratio of the average video data rate and link data rate. The results show that the use of the proposed method greatly benefits the quality of video streaming in all cases. In this case, the bottleneck bandwidth is almost twice the bandwidth of corresponding simulation with FTP/TCP background traffic. Because of this, the quality degradation starts at higher utilization levels. However, the proposed method greatly improves the quality of video streaming in this case. As in the previous tests, the GOP length does not seem to have an important role in the performance of the proposed method.

Table 3: Evaluation of PSNR (dB) for video background traffic using MPEG-4 part10.

		without shaping								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
JP	12	47.2	46.4	47.6	48.0	37.6	35.9	21.0	18.4	10.6
SL	12	41.7	40.9	41.0	37.1	35.4	32.5	26.4	21.6	21.5
SW1	12	45.4	45.4	46.0	26.0	11.9	13.9	11.5	11.4	13.1
SW2	12	46.9	46.7	47.3	47.8	32.4	14.5	14.9	15.7	18.4
SW3	12	46.5	46.6	46.2	38.5	34.0	30.0	28.0	20.7	20.0
		proposed method								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
JP	12	47.2	46.4	47.6	47.6	47.6	36.6	36.5	28.3	25.9
SL	12	41.7	40.9	41.0	41.0	36.4	32.6	32.2	29.7	23.9
SW1	12	45.4	45.4	46.0	46.0	37.4	29.2	25.9	26.4	20.9
SW2	12	46.9	46.7	47.3	47.3	38.4	32.2	13.8	15.5	26.4
SW3	12	46.5	46.6	46.2	46.2	39.4	36.2	34.4	32.1	24.0

Table 4: Evaluation of SSIM for video background traffic using MPEG-4 part10.

		without shaping								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
JP	12	0.98	0.98	0.98	0.98	0.98	0.97	0.84	0.74	0.21
SL	12	0.98	0.98	0.98	0.96	0.96	0.94	0.86	0.77	0.70
SW1	12	0.98	0.98	0.98	0.94	0.76	0.73	0.65	0.54	0.51
SW2	12	0.98	0.98	0.98	0.99	0.97	0.51	0.53	0.64	0.74
SW3	12	0.98	0.98	0.98	0.97	0.97	0.95	0.93	0.78	0.69
		proposed method								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
JP	12	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.94	0.91
SL	12	0.98	0.98	0.98	0.98	0.96	0.92	0.91	0.87	0.80
SW1	12	0.98	0.98	0.98	0.98	0.98	0.96	0.95	0.95	0.89
SW2	12	0.98	0.98	0.98	0.98	0.99	0.97	0.59	0.59	0.63
SW3	12	0.99	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.89

Table 5: Evaluation of SSIM for TCP/FTP background traffic using MPEG-4 part2.

		without shaping								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
Coastguard	8	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.76	0.28
	12	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.63	0.27
	18	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.55	0.26
	30	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.57	0.28
Highway	8	0.97	0.97	0.97	0.97	0.97	0.95	0.94	0.87	0.61
	12	0.97	0.97	0.97	0.97	0.97	0.95	0.92	0.86	0.61
	18	0.97	0.97	0.97	0.97	0.97	0.95	0.91	0.85	0.61
	30	0.97	0.97	0.97	0.97	0.97	0.93	0.90	0.84	0.61
News	8	0.99	0.99	0.99	0.99	0.99	0.99	0.90	0.70	0.73
	12	0.99	0.99	0.99	0.99	0.99	0.94	0.75	0.70	0.74
	18	0.99	0.99	0.99	0.99	0.99	0.91	0.69	0.71	0.74
	30	0.99	0.99	0.99	0.99	0.98	0.91	0.64	0.66	0.73

		without shaping								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
Coastguard	8	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.33
	12	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.36
	18	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.32
	30	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.36
Highway	8	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.91	0.62
	12	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.91	0.62
	18	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.90	0.58
	30	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.90	0.59
News	8	0.99	0.99	0.99	0.99	0.99	0.99	0.92	0.74	0.74
	12	0.99	0.99	0.99	0.99	0.99	0.99	0.80	0.73	0.75
	18	0.99	0.99	0.99	0.99	0.99	0.98	0.80	0.74	0.75
	30	0.99	0.99	0.99	0.99	0.99	0.93	0.76	0.67	0.75

Table 6: Evaluation of PSNR (dB) for TCP/FTP background traffic using MPEG-4 part2.

		without shaping								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
Coastguard	8	41.2	41.2	41.2	41.2	41.2	41.2	41.2	24.4	15.6
	12	41.1	41.1	41.1	41.1	41.1	41.1	41.1	20.1	15.5
	18	41.0	41.0	41.0	41.0	41.0	41.0	41.0	18.0	15.5
	30	41.0	41.0	41.0	41.0	41.0	41.0	41.0	19.0	15.5
Highway	8	42.6	42.6	42.6	42.6	42.6	35.1	33.0	26.2	18.0
	12	42.5	42.5	42.5	42.5	41.7	35.0	30.3	25.2	18.0
	18	42.5	42.5	42.5	42.5	42.4	34.1	27.5	23.7	17.8
	30	42.5	42.5	42.5	42.5	41.7	32.3	25.4	23.4	18.0
News	8	43.9	43.9	43.9	43.9	43.9	43.9	23.8	18.3	18.9
	12	43.7	43.7	43.7	43.7	43.7	27.6	19.6	18.1	19.3
	18	43.7	43.7	43.7	43.7	43.7	24.0	19.0	18.6	19.3
	30	43.6	43.6	43.6	43.6	35.7	24.4	18.0	18.4	19.0

		proposed method								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
Coastguard	8	41.2	41.2	41.2	41.2	41.2	41.2	41.2	41.2	15.7
	12	41.1	41.1	41.1	41.1	41.1	41.1	41.1	41.1	16.0
	18	41.0	41.0	41.0	41.0	41.0	41.0	41.0	41.0	15.7
	30	41.0	41.0	41.0	41.0	41.0	41.0	41.0	41.0	16.0
Highway	8	42.6	42.6	42.6	42.6	42.6	42.6	42.6	28.8	18.0
	12	42.5	42.5	42.5	42.5	42.5	42.5	42.5	28.3	17.9
	18	42.5	42.5	42.5	42.5	42.5	42.5	42.5	27.6	17.3
	30	42.5	42.5	42.5	42.5	42.5	42.5	42.5	27.5	17.6
News	8	43.9	43.9	43.9	43.9	43.9	43.9	26.5	19.2	19.4
	12	43.7	43.7	43.7	43.7	43.7	43.7	20.3	18.8	19.5
	18	43.7	43.7	43.7	43.7	43.7	34.8	19.5	19.2	19.5
	30	43.6	43.6	43.6	43.6	43.6	26.8	18.8	18.2	19.5

Table 7: Evaluation of SSIM for TCP/FTP background traffic using MPEG-4 part10.

		without shaping								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
Coastguard	8	0.98	0.98	0.98	0.60	0.42	0.42	0.42	0.42	0.42
	12	0.98	0.98	0.65	0.43	0.42	0.42	0.42	0.42	0.41
	18	0.98	0.98	0.48	0.43	0.43	0.43	0.43	0.42	0.41
	30	0.98	0.71	0.46	0.45	0.44	0.43	0.42	0.41	0.41
Highway	8	0.95	0.95	0.93	0.85	0.68	0.61	0.59	0.59	0.59
	12	0.95	0.95	0.90	0.80	0.64	0.59	0.59	0.59	0.59
	18	0.95	0.95	0.91	0.74	0.62	0.59	0.58	0.58	0.58
	30	0.95	0.95	0.86	0.69	0.57	0.58	0.57	0.57	0.57
News	8	0.98	0.98	0.98	0.60	0.42	0.42	0.42	0.42	0.42
	12	0.98	0.98	0.65	0.43	0.42	0.42	0.42	0.42	0.41
	18	0.98	0.98	0.48	0.43	0.43	0.43	0.43	0.42	0.41
	30	0.98	0.71	0.46	0.45	0.44	0.43	0.42	0.41	0.41

		proposed method								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
Coastguard	8	0.98	0.98	0.98	0.59	0.42	0.41	0.41	0.41	0.42
	12	0.98	0.98	0.74	0.45	0.42	0.42	0.42	0.43	0.41
	18	0.98	0.98	0.58	0.43	0.43	0.43	0.43	0.43	0.41
	30	0.98	0.84	0.47	0.45	0.44	0.43	0.42	0.41	0.41
Highway	8	0.95	0.95	0.95	0.95	0.72	0.59	0.59	0.59	0.59
	12	0.95	0.95	0.95	0.92	0.67	0.59	0.59	0.59	0.59
	18	0.95	0.95	0.95	0.87	0.60	0.58	0.58	0.58	0.58
	30	0.95	0.95	0.95	0.72	0.56	0.57	0.57	0.57	0.57
News	8	0.98	0.98	0.98	0.59	0.42	0.41	0.41	0.41	0.42
	12	0.98	0.98	0.74	0.45	0.42	0.42	0.42	0.43	0.41
	18	0.98	0.98	0.58	0.43	0.43	0.43	0.43	0.43	0.41
	30	0.98	0.98	0.58	0.43	0.43	0.43	0.43	0.43	0.41

Table 8: Evaluation of PSNR (dB) for TCP/FTP background traffic using MPEG-4 part10.

		without shaping								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
Coastguard	8	35.2	35.2	35.2	35.2	17.8	13.7	13.8	13.9	13.9
	12	35.1	35.1	35.1	35.1	15.5	13.4	13.2	13.2	13.2
	18	35.1	35.1	35.1	35.1	15.5	13.4	13.2	13.2	13.2
	30	35.0	35.1	35.1	35.1	11.5	12.2	12.2	12.2	12.2
Highway	8	35.2	35.2	35.2	35.2	17.8	13.7	13.8	13.9	13.9
	12	35.1	35.1	35.1	35.1	15.5	13.4	13.2	13.2	13.2
	18	35.1	35.1	35.1	35.1	15.5	13.4	13.2	13.2	13.2
	30	35.1	35.1	35.1	35.1	11.5	12.2	12.2	12.2	12.2
News	8	39.6	39.6	39.6	13.3	11.3	11.3	11.3	11.3	11.4
	12	39.8	39.8	13.9	11.5	11.3	11.3	11.3	11.4	11.3
	18	40.0	40.0	11.9	11.5	11.5	11.5	11.5	11.6	11.4
	30	40.2	15.5	11.6	11.6	11.6	11.5	11.4	11.4	11.3

		proposed method								
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Video	GOP									
Coastguard	8	35.2	35.2	35.2	35.2	35.2	13.7	13.8	13.9	13.9
	12	35.1	35.1	35.1	35.1	24.1	11.9	13.2	13.2	13.2
	18	35.1	35.1	35.1	35.1	12.7	12.1	12.6	12.6	12.6
	30	35.1	35.1	35.1	35.1	14.8	12.2	12.2	12.2	12.2
Highway	8	35.2	35.2	35.2	35.2	35.2	13.7	13.8	13.9	13.9
	12	35.1	35.1	35.1	35.1	24.1	11.9	13.2	13.2	13.2
	18	35.1	35.1	35.1	35.1	12.7	12.1	12.6	12.6	12.6
	30	35.1	35.1	35.1	35.1	14.8	12.2	12.2	12.2	12.2
News	8	39.6	39.6	39.6	13.4	11.3	11.3	11.3	11.3	11.3
	12	39.8	39.8	16.8	11.9	11.3	11.3	11.3	11.6	11.4
	18	40.0	40.0	13.4	11.5	11.5	11.5	11.7	11.7	11.4
	30	40.2	20.6	12.0	11.6	11.6	11.5	11.5	11.5	11.3

Table 9: Evaluation of SSIM for video background traffic using MPEG-4 part10.

		without shaping									
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
Video	GOP										
Coastguard	8	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.63
	12	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.64
	18	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.83	-
	30	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.91	-
Highway	8	0.95	0.95	0.95	0.95	0.94	0.93	0.92	0.89	0.87	-
	12	0.95	0.95	0.95	0.95	0.94	0.93	0.90	0.94	0.86	-
	18	0.95	0.95	0.95	0.95	0.94	0.93	0.91	0.88	0.84	-
	30	0.95	0.95	0.95	0.95	0.93	0.91	0.89	0.86	0.85	-
News	8	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.95	-
	12	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.96	0.91	-
	18	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.96	0.81	-
	30	0.98	0.98	0.98	0.98	0.98	0.98	0.95	0.77	0.70	-

		proposed method									
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
Video	GOP										
Coastguard	8	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
	12	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
	18	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	-
	30	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	-
Highway	8	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	-
	12	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	-
	18	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.94	-
	30	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.94	-
News	8	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	-
	12	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.96	-
	18	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.97	0.95	-
	30	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.95	0.84	-

Table 10: Evaluation of PSNR (dB) for video background traffic using MPEG-4 part10.

		without shaping									
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
Video	GOP										
Coastguard	8	35.2	35.2	35.2	35.2	35.2	35.2	35.2	35.2	35.2	18.4
	12	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	17.6
	18	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	27.0	-
	30	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	22.0	-
Highway	8	38.6	38.6	38.6	38.6	29.0	25.9	22.2	19.8	17.8	-
	12	38.7	38.7	38.7	38.7	27.7	26.2	21.8	20.5	17.7	-
	18	38.7	38.7	38.7	38.7	28.0	22.9	23.0	20.3	17.5	-
	30	38.7	38.7	38.7	38.7	26.1	22.0	20.2	17.0	16.7	-
News	8	39.6	39.6	39.6	39.6	39.6	39.6	39.6	39.6	25.5	-
	12	39.8	39.8	39.8	39.8	39.8	39.8	39.8	25.0	20.9	-
	18	40.0	40.0	40.0	40.0	40.0	40.0	40.0	23.9	16.9	-
	30	40.2	40.2	40.2	40.2	40.2	40.2	24.0	16.8	14.9	-

		proposed method									
Utilization		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
Video	GOP										
Coastguard	8	35.2	35.2	35.2	35.2	35.2	35.2	35.2	35.2	35.2	35.2
	12	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1
	18	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	-
	30	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	-
Highway	8	38.6	38.6	38.6	38.6	38.6	38.6	38.6	38.6	38.62	-
	12	38.7	38.7	38.7	38.7	38.7	38.7	38.7	38.7	38.66	-
	18	38.7	38.7	38.7	38.7	38.7	38.7	38.7	38.7	32.18	-
	30	38.7	38.7	38.7	38.7	38.7	38.7	38.7	38.7	32.03	-
News	8	39.8	39.8	39.8	39.8	39.8	39.5	39.5	39.5	39.5	-
	12	39.8	39.8	39.8	39.8	39.8	39.8	35.6	33.7	31.2	-
	18	40.0	40.0	40.0	40.0	40.0	33.0	33.1	32.0	27.2	-
	30	40.2	40.2	40.2	40.2	40.2	40.2	40.2	28.1	18.9	-

4 Conclusions

In this paper, we propose the use of traffic shaping to mitigate the effects of network impairments in viewer's perceived quality of video streaming over IP networks. The shaping method used was the token bucket algorithm. The setting of bucket size was performed according to a proposed heuristic, which is based on the characteristics of the encoded video.

The improvement in viewer's perceived video quality is demonstrated through computer simulations by comparing SSIM and PSNR of the original image with the received one. The results show that the proposed approach effectively improves the perceived quality, mainly in situations of higher levels of link utilization. The results also show that videos encoded with H.264/AVC are more affected by background traffic as compared to videos encoded with MPEG-4 part 2.

The video background traffic was more harmful to the user's perceived quality than the TCP background traffic. This was expected because the transmission of streaming video is performed using the UDP protocol, which is not user friendly in terms of the congestion control. When the proposed method is applied in the concurrent transmission of two videos, a significant improvement in the perceived quality was observed.

The proposed method is simpler to implement as compared to existing methods. Moreover, due to the availability of token bucket algorithm in routers and servers, it is not necessary to perform additional changes in server/routers, and the proposed method can be immediately made available for implementation.

Acknowledgements: The research was supported by the Brazilian National Council for Scientific and Technological Development (CNPq), research grant 455584/2012-5. The authors would like to thank Evandro Luis Copercini and Guilhermo Tonin Agranionih for their work in developing programs and testing.

References:

[1] ISO/IEC 14496-10:2003 Information technology - coding of audio-visual objects - part 10: Advanced video coding, September 2004.

[2] Triple-play services quality of experience (QoE) requirements, December 2006. TR-126. On-line: http://www.broadband-forum.org/technical/download/TR-126.pdf.

[3] Recommendation ITU-R BT.1788, January 2007. Methodology for the subjective assessment of video quality in multimedia applications.

[4] Recommendation ITU-T P.910, April 2007. Subjective video quality assessment methods for multimedia applications.

[5] Requirements for the support of IPTV services, January 2009. Recommendation Y.1901, Int. Telecommunication Union (ITU).

[6] Measuring broadband America, February 2013. Federal Communications Commission (FCC). On-line: http://www.fcc.gov/measuring-broadband-america/2013/February.

[7] Video trace library, January 2013. Arizona State University. On Line: http://trace.eas.asu.edu/.

[8] A. Afanasyev, N. Tilley, P. Reiher, and L. Kleinrock. Host-to-host congestion control for TCP. *Communications Surveys Tutorials, IEEE*, 12(3):304–342, Third 2010.

[9] M. F. Alam, M. Atiquzzaman, and M. A. Karim. Traffic shaping for MPEG video transmission over the next generation Internet. *Computer Communications*, 23(14-15):1336–1348, August 2000.

[10] R. Braden, D. Clark, and S. Shenker. Integrated services in the Internet architecture: an overview, June 1994.

[11] L. Breslau, D. Estrin, K. Fall, Sally Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Ya Xu, and H. Yu. Advances in network simulation. *IEEE Computer*, 33(5):59–67, 2000.

[12] Yung-Mu Chen, Tein-Yaw Chung, Ming-Yen Lai, and Chih-Hung Hsu. MACD-Based motion detection approach in heterogeneous networks. *EURASIP J. Wireless Comm. and Networking*, 2008, july 2008.

[13] Faheem, F. Humayun, M.IK. Babar, M.H. Zafar, and M.F. Zuhairi. Performance analysis of a token bucket shaper for MPEG4 video and real audio signal. In *Smart Instrumentation, Measurement and Applications (ICSIMA), 2013 IEEE International Conference on*, pages 1–4, Nov 2013.

[14] Jason Greengrass, John Evans, and Ali C. Begen. Not all packets are equal. *IEEE Internet Computing*, 13:74–82, March 2009.

[15] Sungwoo Hong and Youjip Won. Incorporating packet semantics in scheduling of real-time multimedia streaming. *Multimedia Tools Appl.*, 46:463–492, January 2010.

[16] Yaqing Huang, Roch Guérin, and Pranav Gupta. Supporting excess real-time traffic with active drop queue. *IEEE/ACM Transactions Networking*, 14:965–977, October 2006.

[17] Jirka Klaue, Berthold Rathke, and Adam Wolisz. Evalvid - a framework for video transmission and quality evaluation. In *13th Int. Conference on Modeling Technology for Computer Performance Evaluation*, pages 255–272, 2003.

[18] J. Maisonneuve, M. Deschanel, J. Heiles, Wei Li, Hong Liu, R. Sharpe, and Yiyan Wu. An overview of IPTV standards development. *IEEE Transactions on Broadcasting*, 55(2):315–328, June 2009.

[19] Carlos Eduardo Maffini Santos, Eduardo Parente Ribeiro, and Carlos Marcelo Pedroso. The application of neural networks to improve the quality of experience of video transmission over IP networks. *Engineering Applications of Artificial Intelligence*, 27:137–147, 2014.

[20] Michael Schier and Michael Welzl. Selective packet discard in mobile video delivery based on macroblock-based distortion estimation. In *28th IEEE Int. Conf. on Computer Communications Workshops*, INFOCOM'09, pages 212–217, Piscataway, NJ, USA, 2009. IEEE Press.

[21] R. Serral-Gracià, E. Cerqueira, M. Curado, M. Yannuzzi, E. Monteiro, and X. Masip-Bruin. An overview of quality of experience measurement challenges for video applications in IP networks. In *8th International Conference on Wired/Wireless Internet Communications*, WWIC'10, pages 252–263, Berlin, Heidelberg, 2010. Springer-Verlag.

[22] K. Silpa and Aruna Mastani. Comparison of image quality metrics. *International Journal of Engineering Research and Technology (IJERT)*, 1(4):1–5, 2012.

[23] Andrew Tanenbaum. *Computer Networks*. Prentice Hall Professional Technical Reference, 4th edition, 2002.

[24] I Wahidah, I Iwut, S. Naning, N. Adhi, and A Widya. Comparison analysis of H.264 and MPEG-4 digital video transport over wireless IPv4 networks. In *Wireless and Optical Communications Networks, 2009. WOCN '09. IFIP International Conference on*, pages 1–6, April 2009.

[25] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[26] J. Watkinson. *The MPEG Handbook: MPEG-1, MPEG-2, MPEG-4*. Broadcasting and communications. Focal Press, 2001.

[27] Yang Xiaogang, Liu Jiqiang, and Li Ning. Congestion control based on priority drop for h.264/svc. In *Int. Conf. on Multimedia and Ubiquitous Engineering - MUE'07*, pages 585–589, April 2007.