Performance Evaluation of Video Transmission over IEEE 802.11b WLANs from the Video Quality Perspective

Peter Počta, Robert Hudec, Martin Bojmír
University of Žilina, FEE, Dept. of Telecommunications and Multimedia, Univerzitná 1, SK-010 26, Žilina, Slovakia
E-mail: pocta@fel.uniza.sk

Abstract. This paper deals with measurements of the impact of background traffic on video quality in WLAN (IEEE 802.11b) environment. The simulated background traffic consists of two types of the current traffics in telecommunication networks such as the data transfer service and Web service. The background traffic was generated by means of the accomplished Distributed Internet Traffic Generator (D-ITG). The investigation of the impact of these types of traffic and traffic load on video quality using two different types of video sequences is the aim of this paper. The assessment of video quality is carried out by means of the accomplished objective full-reference assessment criteria, such as Peak Signal-to-Noise Ratio (PSNR), Video Quality Metric (VQM), and Structural SIMilarity (SSIM).

Keywords: IEEE 802.11, Video over WLAN, video quality, PSNR, SSIM, VQM

1 Introduction

Wireless multimedia transmission across Wireless Local Area Networks (WLANs) has been gaining attention in the recent years because of the proliferation of technologies like Bluetooth, IEEE 802.11, 3G, and WiMAX. In particular, IEEE 802.11 WLAN [1] has emerged as a prevailing technology for (indoor) broadband wireless access because it supports real-time conversational multimedia applications like the Voice over Internet Protocol and video conferencing.

To provide person-to-person (instead of place-to-place) connections anywhere and anytime, the Internet is expected to penetrate the wireless domain. A very promising wireless network is the Wireless Local Area Network (WLAN), which has shown the potential to provide high-rate data services at low cost over local area coverage. Working in the license-exempted 2.4GHz industrial, scientific, and medical (ISM) frequency band, the IEEE 802.11b WLAN and IEEE 802.11g WLAN offer a data rate up to 11Mbps and 54 Mbps, respectively, while IEEE 802.11a WLAN and European Telecommunications Standard Institute (ETSI) Hiperlan/2 can support data rates up to 54 Mbps in the 5GHz frequency band. Nowadays, the work on upcoming IEEE 802.11n WLAN standard is in progress. Its publication is currently expected in September, 2008 [2]. The IEEE 802.11n will be able to support a data rate up to 248 Mbps (2 streams) in the 5GHz or 2.4GHz frequency band. As a wireless extension to the wired Ethernet, WLANs typically cover a small geographic area, in hotspot local areas where the traffic intensity is usually much higher than in other areas.

One major challenge for Video over WLAN is Quality of Service (QoS) provisioning. Originally designed for high-rate data traffic, WLANs may experience bandwidth inefficiency when supporting delay-sensitive video traffic. Hence, it is essential to enhance the QoS support capability of current WLAN standards, such as the most popular IEEE 802.11 standard. The assessment of video quality in the current WLAN standards is very desirable opportunity for design, evaluation, verification, and testing of enhanced QoS methods for supporting video services in this type of very fast developing wireless telecommunication networks.

Video quality is jointly affected by various network-dependent and application-specific factors. For instance, packet losses and delay jitter (which also translates into losses in the playout buffer) are the major network-dependent factors, while video codec and loss recovery technique, coding bit rate, packetization scheme and
content characteristics are the major application specific factors that affect video quality and the sensitivity to networks errors. The above mentioned factors can usually cause some distortions and artifacts in the video signal, such as image and sound defects, temporal and spatial artifacts etc.

Some works have studied the performance evaluation of video transmission over IEEE 802.11b WLANs. Particularly, [3–5] examined performance of IEEE 802.11 in Peak Signal-to-Noise Ratio (PSNR) domain. In [3], MPEG-4 video transmission over IEEE 802.11b wireless channel was evaluated with simulation scenarios in ns-2. Different data flows patterns were used in order to investigate stream QoS parameters in terms of PSNR, end-to-end delay and jitter. Simulations results show that the quality of received stream can be affected by network behavior even if high quality encoding characteristics are chosen. Therefore, the encoding technique must be chosen based not only on the network state but also on the content characteristics of the sending stream. In [4], the important problem of delay-sensitive transmission of video over IEEE 802.11a/e WLANs is addressed. The simple cross-layer strategy using subflows at application layer, frame length, retry limit at the MAC layer, and modulation schemes at physical layer was illustrated to improve the quality of multimedia. Finally, the concept of air/time fairness was proposed, that enables a fair division of resources among competing wireless stations when different cross-layer strategies are deployed. In [5], authors present an application-level perceptual ARQ algorithm for video streaming over 802.11e wireless networks. A simple and effective formula is proposed to combine the perceptual and temporal importance of each packet into a single priority value, which is then used to drive the packet-selection process at each retransmission opportunity. Video streaming of H.264 test sequences has been simulated with ns-2 in a realistic 802.11e home scenario, in which the various kinds of traffic flows have been assigned to different 802.11e access categories according to the Wi-Fi alliance WMM specification. Extensive simulations show that the proposed method consistently outperforms the standard link-layer 802.11 retransmission scheme, delivering PSNR gains up to 12 dB while achieving low transmission delay and limited impact on the concurrent traffic. Some works focused on real-time monitoring of video quality in IP networks, such as [6]. In [6], authors introduced an approach for on-line estimation of quality of video transmitted over network paths. The proposed solution would allow the large-scale monitoring of video quality using only simple measurements of network performance. The feasibility and accuracy of the proposed solution were assessed through extensive simulations and experiments.

Here, we focus on the impact of background traffic on video quality of transmission sequences in the environment of IEEE 802.11b networks. The background traffic was generated by means of Distributed Internet Traffic Generator (D-ITG) [7]. The simulated background traffic consists of two types of the current traffics. The current traffics are: the data transfer service and Web service. Increasing traffic load causes the increasing jitter and packet loss. In general, video quality drops with increasing packet loss and jitter. The impact of these types of traffic and traffic load on video quality using two different types of video sequences is studied in this paper. The video quality is assessed by means of the accomplished objective full-reference objective criteria, such as PSNR, Video Quality Metric (VQM) and Structural SIMilarity (SSIM).

The rest of the paper is organized as follows. Section 2 briefly reviews the currently used methods for the video quality assessment. Section 3 describes the experimental scenario. Section 4 presents the experimental results. Section 5 concludes the paper and suggests some future studies.

2 Overview of the currently used video quality assessment methods

From the point of view of measurement of video quality, there are two major approaches, namely, subjective and objective ones. In general, the correct evaluation of visual information can be offered by subjective criteria only. Being directly based on human eyes perception, they are more expensive. On the other hand, the objective criteria like SSIM or VQM try to substitute the subjective approaches. Moreover, they are widely used worldwide as they offer results closer to the response of Human Visual System (HVS).

2.1 Subjective approaches

There are a lot of subjective methodologies standardized by ITU-R BT.500 [13] and ITU-T P.910. Among these criteria have one for example: Double Stimulus Impairment Scale (DSIS), Double Stimulus Continuous Quality Scale (DSCQS), Single Stimulus Continuous Evaluation (SSCQE), and Simultaneous Double Stimulus for Continuous Evaluation (SDSC) or Subjective Assessment Methodology for Video Quality (SAMVIQ) methodology. Following the reasons given above, the subjective criteria will be replaced by the highly correlated objective criteria.

2.2 Objective approaches

The main problem of the objective approaches regarding correct evaluation of the visual information is modeling the human eye, which is a difficult task. The objective criteria are expected to unify and repeatability of the evaluation process. These criteria can by divided into two classes.
In the first class there are some simple criteria like Mean Absolute Error (MAE), Mean Absolute Error Reduction (MAER), Mean Square Error (MSE), Noise Reduction (NR), Signal-to-Noise Reduction (SNR), Color Difference (CD), Motion Sum of Absolute Differences (MSAD), PSNR, Blurring measure, Blocking measure, etc., which are focused on a special image parameter. For example, MAE measures preservation of image details and edges, MSE measures suppression of image artifacts, CD measures changes in image chromaticity, etc. As to color image quality evaluation with the objective criteria closer to HVS, than one image parameter like the image contrast, luminance, color, structure, texture etc. must be included.

In the second class have one enhanced criteria like New Quality Index (NQI), VQM or SSIM. They combine more than one image parameter enabling them to be closer to response of HVS than the first criteria category. In our future experiments we will use the PSNR, VQM and SSIM criteria only.

2.2.1 PSNR

The PSNR criterion is derived from the MSE criterion. It expresses the ratio between the maximum possible power of a signal and the power of error/noise corrupting the original image signal. As images represent a signal of a wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale.

$$PSNR(x,y) = 20 \log_{10} \left( \frac{L}{\frac{1}{NM} \sum_{i=1}^{M} \sum_{j=1}^{N} (x_{i,j} - y_{i,j})^2} \right),$$ (1)

where L represents the number of gray scale image levels (for 8bit image representation it is equal to 255) and N and M define the image dimension. If the measured dynamic image sequence contains small differences compared to the original sequence one, the PSNR is high.

2.2.2 VQM

VQM is a metric derived in 1991 from Digital Video Quality (DVQ) and based on Discrete Cosine Transform (DCT) coefficients that indicate a 95% correlation to the subjective criteria [11, 12]. The main idea consists in applying Spatial Contrast Sensitivity Function (SCSF) matrix for static frames and dynamic frames in one step. The original and received/measured images are transformed to the DCT coefficients. Further, they are converted to the Local Contrast (LC) as follows

$$LC(i,j) = DCT(i,j) \left( \frac{DC}{1024} \right)^{0.65},$$ (2)

where DC is a DC coefficient and the 0.65 value is the best parameter for fitting psychological data. Moreover, all these steps are identical to the Watson’s DVQ criterion. For the static SCSF matrix, the DCT coefficients are multiplying with corresponding element in default MPEG quantization matrix [12]. Likewise, for the dynamic matrix, each entry in the static SCSF matrix is raised to a power to account for the temporal property of SCSF and the power is determined by the frame-rate of video sequences [12].

$$VQM(x,y) = 1000 \frac{\text{diff}}{\text{max} \{\text{diff} \}} + 0.005 \times 1000 \times \left[ \frac{\text{diff}}{\text{max} \{\text{diff} \}} \right],$$ (3)

where the value 0.005 is the maximum distortion weight and value 1000 is the standardization ratio.

2.2.3 SSIM

The SSIM criterion represents another approach to the video quality measurement. It is based on the fact that HVS is highly focused on the structural information inside the image and is not on the errors. This is why all the criteria based on the structural distortion (including the SSIM), offer a higher correlation with the subjective criteria.

Let x and y be digitized images defined by follow statistical averages $\mu_x$ and $\mu_y$, variances $\sigma_x$ and $\sigma_y$, and covariance $\sigma_{xy}$

$$\mu_x = \frac{1}{N} \sum_{i=1}^{N} x_i, \quad \mu_y = \frac{1}{N} \sum_{i=1}^{N} y_i,$$ (4a)

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)^2},$$ (4b)

$$\sigma_y = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (y_i - \mu_y)^2},$$ (4c)

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y).$$ (4d)

The measurement made with SSIM is divided into three parts, e. g. the l-luminance, c-contrast and s-structure measurements. The overall similarity measure is defined as follows

$$S(x,y) = f(l(x,y), c(x,y), s(x,y)).$$ (5)

The partial elements are relatively independent and can be defined as follows

$$l(x,y) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1},$$ (6a)

$$c(x,y) = \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2},$$ (6b)
\[ s(x,y) = \frac{2\sigma_{xy} + C_1}{\sigma_x \sigma_y + C_3}, \quad (8c) \]

where \( C_1 = (K_L)^2 \), \( C_2 = (K_L)^2 \) and \( C_3 = C_2/2 \) are constant so as to avoid instability and \( L \) define the dynamic range of image analogously to the PSNR criterion. After combining all parts, the SSIM index is defined by the final formula

\[ SSIM(x,y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \quad (9) \]

The averaged value of the SSIM index is for the overall image quality defined as follows

\[ Q = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} SSIM_{i,j}}{NM}. \quad (10) \]

3 Experimental scenario

3.1 Experimental setup

A one-way video session was established between two hosts (VLC Server and wireless VLC Client), via the Access Point (AP), in IEEE 802.11b WLAN (see Figure 1). Two wireless stations (ITG Sender and wireless ITG Receiver) equipped with the accomplished D-ITG traffic generator [7] were used to generate and receive background traffic. ITG Sender generated the Transmission Control Protocol (TCP) packets of length of 1024 bytes. The background traffic is described in Chapter 3.3. The video traffic was generated using VLC media player clients [8]. Two VLC clients were deployed for the purpose of these experiments. One of them was configured as a server and the second one as a client. In the case of this video streaming session, the MPEG-TS format was used.

The measurements have been performed for five different testing conditions. The video sequences described in Chapter 3.2 are utilized for transmission through the given Video over WLAN connection. Finally, the video quality has been assessed by the currently accomplished objective full-reference criteria, such as PSNR, VQM and SSIM.

The adjusting of wireless station positions and antennas was being performed until the signal strength and link quality achieved by all the stations were roughly similar. All the wireless stations had a fixed position (no mobility) during the measurements. The signal strength and link quality were kept in the range from 85% to 100% (excellent) for all the performed measurements.

3.2 Reference video sequences

The reference video sequences, which are of 300-seconds long, were stored on the server in video codec v.3 MPEG-4 format. In [9], it was pointed out that the use of longer test sequences (e.g. 5 minutes) would enable the examination of network loss and burst size under more realistic test conditions. It was our reason for the using 300-seconds long test sequences, in the case of these experiments.

Table 1. Description of audio-video content for the sequences used in experiments

<table>
<thead>
<tr>
<th>Video Sequence</th>
<th>Video Content</th>
<th>Audio Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.1</td>
<td>Grand prix Monaco, a few dramatic situations during monoposto racing</td>
<td>Race reports (anchor), some dramatic situations on the race</td>
</tr>
<tr>
<td>No.2</td>
<td>Sitcom, people sitting on a sofa having conversation</td>
<td>Free talk in living-room</td>
</tr>
</tbody>
</table>

The parameters of both test sequences were as follows: video resolution 512x384 pxl, frame rate 25fps and bit rate 1 Mbps for video stream and 64 kbps for audio stream.

In these experiments, we used two different types of video sequences. The content of each sequence is provided in Table 1. Selection of the material was based on the typical entertainment-oriented video streaming content.

3.3 Background traffic

The background traffic has been generated by the D-ITG traffic generator. The primary goal of the background traffic is two-fold. Firstly, it simulates the standard traffic that appears in current wireless networks, which includes data transfer services via Hypertext Transfer Protocol (HTTP) and File Transfer Protocol (FTP). Secondly, it affects video transmission by changing of Video over WLAN connection network performance parameters such as delay, jitter and packet loss. The simulated background traffic includes two types of communication:
• “Data transfer service”, which includes FTP and other non-specified services, is represented as information stream with a constant bit rate based on TCP.

• “Web service” that is simulated as a sequence of separated data bursts with Poisson distribution of the packet rate. The active period of burst is 400 ms and the bursts appear periodically every 10 seconds. TCP was used for the purpose of this service.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.08</td>
<td>10.5</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0.16</td>
<td>19.5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.24</td>
<td>28.5</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.32</td>
<td>37.7</td>
</tr>
</tbody>
</table>

As mentioned in Section 3.1, the measurements have been performed for five different testing conditions. The selected bit rates of the two above mentioned types of communication and average traffic load of the background traffic are described in Table 2. The calculation of the average traffic load was based on the 100% channel rate, which is 11Mbps for IEEE 802.11b technology. The traffic load has been measured by means of the Wireshark network analyzer [10].

4 Experimental results

The measurements were independently performed ten times under the same testing condition for both types of the investigated video sequences (Grand prix, Sitcom). The video quality results (PSNR, VQM and SSIM values) were averaged. In Figures 2, 3 and 4, it can be seen that the background traffic has an impact on the overall video quality in the case of all investigated video sequences.

On the basis of our assumption, we expect that the slope of the curves would be significantly changed in the case of the different traffic loads, expressed by different testing conditions in this paper. Especially, in the transmission of the sequence with lower motion vector like the test sequence No. 2, no significant change of the curves slope would be achieved. On the other hand, in a dynamic sequence, the significant change of curves slope would be observed. From this reason, the objective criteria will be compared from above mentioned assumption perspective. The experimental results for the examined video quality assessment criteria will be described in more detail below.

Figure 2 shows the measurement results for both types of the investigated video sequences. The graphs represent the dependence of the PSNR change on the testing conditions. The testing conditions represent a few types of network conditions. Each network condition is described by the traffic load. The increasing traffic load causes jitter and also packet loss to increase. In general, video quality drops with the increasing packet loss and jitter. The transmission rates for the given testing conditions are described in Table 2.

As seen from Figure 2, the PSNR criterion is less sensitive to transmitted content changing (type of the video sequence) from our assumption point of view. This lower sensitivity is described by similar slopes of the investigated sequences curves (Seq. No.1 and Seq. No.2). In our experiments, the constant shifting approximately 4 dB in the PSNR domain has been achieved from the transmitted content changing point of view. Based on the experiments, the PSNR criterion is not suitable for future video quality evaluation because of the low correlation with the subjective tests [15] thus has been pointed out in some previous works. For instance in [15], the correlation was only around 40%. Probably, the above mentioned lower sensitivity could be one of the reasons for this weak correlation. Naturally, that is a point for a future investigation in this area because it requires a more precise elaboration.

Figure 2: PSNR as a function of the background traffic for both types of the investigated video sequences.

Figure 3: VQM as a function of the background traffic for both types of the investigated video sequences.
In principle, the VQM and SSIM criteria are based on other approaches to quality assessment than PSNR. Thanks to high correlation with the HVS [11, 12], VQM or SSIM curves well take into account the transmitted content changing. Moreover, the both facts are reason for higher sensitivity to traffic load changing too. This fact can be seen in Figures 3 and 4.

Figure 4: SSIM as a function of the background traffic for both types of the investigated video sequences.

There are small differences achieved among the VQM and SSIM curves, but we can not exactly say which criterion is better. This decision can be resolved by the subjective tests and exhaustive investigation of the packet lost impact on the video quality.

5 Conclusion

This paper investigated the impact of the background traffic on video quality in Video over WLAN applications. Different traffic testing conditions were used for the purpose of the measurements. Each testing condition consists of the two types of the current traffics, which exist in the telecommunication networks. The current traffics are the data transfer service and Web service. The video quality has been assessed by the currently used objective full-reference criteria, such as PSNR, VQM and SSIM.

The results show that the background traffic has an impact on overall video quality in the case of all investigated video sequences. In the case of the PSNR criterion, the lower sensitivity to the transmitted content changing has been achieved. The constant shifting approximately 4 dB in the PSNR domain has been observed from the transmitted content perspective (type of the video sequence). VQM or SSIM curves well take into account the transmitted content changing by traffic load. That is caused by two facts. Firstly, the VQM and SSIM criteria are based on other approaches to video quality evaluation than PSNR and secondly, these criteria have higher correlation with subjective tests [11, 12]. A future work will focus towards the following issues. Firstly, we would like to verify our results by subjective tests and afterwards propose the final findings that could suggest the most exact full-reference assessment criterion for the purpose of the objective video quality evaluation in real Video over WLAN scenarios. Secondly, we plan to realize an exhaustive investigation of an impact of packet loss on video quality for the different reference video sequences and packet loss patterns and codecs parameters from the PSNR, SSIM and VQM perspective. Thirdly, we want to investigate the impact of lower sensitivity of PSNR criterion to transmitted content changing (type of the video sequence) on its final estimations from subjective perspective.

6 Acknowledgements

The work has been partially supported by the VEGA grant agency projects no. 1/0313/08, “The investigation of the methods of detection of the critical conditions in telecommunication networks from the speech quality point of view” and no. 1/4066/07, “New systems and principles of semantic description and retrieval of multimedia content”.

7 References


Peter POČTA was born in 1981, in Nové Zámky, Slovakia. He received his M.S. and Ph.D. degrees from University of Žilina, Faculty of Electrical Engineering, Slovakia in 2004 and 2007, respectively. During his Ph.D. study, he realized a few fellowships. Firstly, he spent 3 months as an Erasmus student in the Department of Electrical Engineering and Information Technology, Chair of Telecommunications at Dresdner University of Technology, Germany. He collaborated on testing principles over ADSL access lines. Secondly, he was with Alcatel-Lucent, R&D center, Network integration department, Stuttgart, Germany. He was entrusted with investigation of some impacts on speech quality in WiMAX system. He is currently as an Assistant Professor in the Department of Telecommunications and Multimedia at University of Žilina and European Telecommunications Standards Institute/Speech Transmission Quality (ETSI/STQ) working group member. Areas of his interest include speech quality assessment, access networks, convergent networks, VoIP, VoWLAN, VoWiMAX and cross-layer optimization.

Róbert HUDEC was born in Revúca in 1974, Slovakia. He received his M.S. and Ph.D. degrees from the Technical University in Košice, Faculty of Electrical Engineering and Informatics, Slovakia, in 1998 and 2003, respectively. Nowadays, he is an Associated Professor at the Department of Telecommunications and Multimedia, University of Žilina. Areas of his research includes digital image processing with special focus laid on filtration of mixed noise by using adaptive order-statistics filters and combined low-level image description for MPEG-7 standard. From 2005 his research interests covers also systems, services and terminals of e-/m-health applications. He received Werner von Siemens Excellent Award and Jozef Murgaš Award for his research on image filtration in 2003 and 2007, respectively, and Award of „Vice premier and minister of school for science and research” in the category: celebrity of research and development till 35 years old for his research on e/m-health systems and terminals in 2006.

Martin BOJMÍR was born in Žilina in 1985, Slovakia. He received his B.S. degree in Telecommunications from University of Žilina, Faculty of Electrical Engineering, Slovakia, in 2007. Currently, he is master student at the same faculty. His areas of interest include video quality assessment and networking.