MPEG-4/AVC versus MPEG-2 in IPTV

Stefan Paulsen[1], Tadeus Uhl[1] and Krzysztof Nowicki[2]
[1]*Flensburg University of Applied Sciences, Kanzleistr. 91-93, D-24943 Flensburg, Germany*
[2]*Gdansk University of Technology, Narutowicza 11/12, PL 80952 Gdansk, Poland*

Keywords: Communication Networks, Communication Services, Communication Protocols, Multimedia Applications, IPTV, QoE, PEVQ, MPEG-4/AVC, MPEG-2, ISO/IEC 13818-1 Transport Stream.

Abstract: This paper is essentially a treatment of the theoretical and practical aspects of the new IPTV service. The central part of the paper constitutes a detailed presentation of analysis scenarios and results, and addresses the following issues in particular: What influence does the encoding rate have of on QoE values? What effect does the most obtrusive impairment factor in a network, i.e. packet loss, have on QoE in IPTV? Is the MPEG-2 Transport Stream suitable for encapsulation and transport of MPEG-4/AVC content? Are there alternatives to the ISO/IEC 13818-1 Transport Stream? If so, how do they affect quality of service (QoE)?

1 INTRODUCTION

An ever increasing number of companies are offering television broadcasting and interactive video services such as video on demand (VoD) using digital subscriber line (DSL) technology. Before going any further, it should be pointed out that we are talking about high-speed DSL connections that use the simple metallic pair normally associated with telephones to transmit television programmes. A far better way to carry this service over the last mile is passive optical network (PON) technology. With PON it is possible to achieve last-mile transmission rates in the range of several Gbps. The good old Internet Protocol (with all its drawbacks) is used in the core network and over the last mile as well. The IPTV service itself is supported in the upper layers by the User Datagram Protocol (UDP) or the Real-time Transport Protocol (RTP), or both. Content encoding using MPEG-2 (ISO/IEC 13818-2, 1995) or MPEG-4/AVC (ITU-T H.264, 2007) is done in the highest layer where encoded data is then encapsulated into the MPEG-2 Transport Stream in accordance with ISO/IEC 13818-1 (ISO/IEC 13818-1, 2000). The question arises: Is the MPEG-2 Transport Stream at all suitable for encapsulation and transport of MPEG-4/AVC content? Further questions that need to be answered are: What influence does the encoding rate have of on QoE values? What effect does the most disruptive impairment factor in a network, i.e. packet loss, have on QoE in IPTV? Could the so-called Native RTP technology of MPEG-4/AVC perhaps be more suitable for transporting video content across the networks? This paper describes the search for answers to these questions.

2 MPEG-2 TRANSPORT STREAM

MPEG-2 transport streams according to Rec. ISO/IEC 13818-1 (ISO/IEC 13818-1, 2000) are composed of 188-byte TS packets, each with a 4-byte header. Some TS packets contain an optional Adaption Field, the size of which depends on flags set in the packet header and which may contain timing information, pad bytes, and other data. TS packet payloads may contain program information as well as Packetized Elementary Streams (PES), typically video and audio streams. PES packets are broken into 184-byte chunks to fit into the TS packet payload. So, it is necessary to pad a TS packet that carries the last chunk of a PES packet when there are insufficient PES data to fill it.

A transport stream contains multiplexed data, carrying program stream (PS) packets with payloads from multiple PES packets – again, typically audio and video – and associated program information (PMT: Program Map Table) too. Because PES packet headers contain both Adaption Fields and timing information, no other signalling is necessary to synchronise multiple streams for playback (see

Paulsen S., Uhl T. and Nowicki K..
MPEG-4/AVC versus MPEG-2 in IPTV.
DOI: 10.5220/0004013700270030
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications and Wireless Information Networks and Systems* (SIGMAP-2012), pages 27-30
ISBN: 978-989-8565-25-9

Fig. 1).

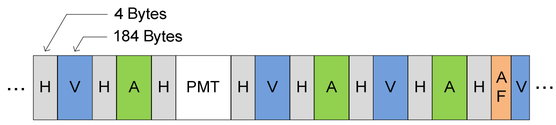

Figure 1: Format of the MPEG-2 Transport Stream (H: TS header; V: video data; A: audio data; AF: adaption field).

There are basically two ways of conveying transport streams through IP networks. The TS packets can either be encapsulated directly into the payload of the UDP datagrams, or they can be transported with the aid of the protocol RTP (IETF RFC 2250, 1998 and IETF RFC 3984, 2005), which supports the synchronisation of real-time services such as IPTV. In either case exactly 7 sequential TS packets are encapsulated in a UDP or RTP packet. The number 7 results from the Maximum Transmission Unit (MTU) in Ethernet-based networks ($1500\,bytes \div 188\,bytes \approx 7$) (see Fig. 2).

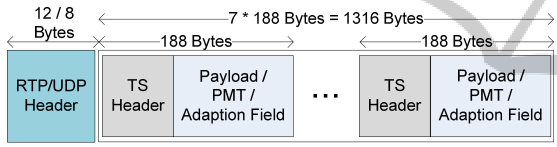

Figure 2: Format of the RTP/UDP packet carrying 7 MPEG-2 TS.

The software tool FFmpeg (FFmpeg, state 2012) was used throughout the study described in this paper to encode and decode and to create the transport streams from the reference file. It is free to use for any non-profit-making purposes and comes with a large number of platform-independent applications and libraries that can be used to record, convert and stream audio and video material.

3 NATIVE RTP IN MPEG-4/AVC

Another way of sending video data over IP-based networks is the native use of RTP packets (IETF RFC 3640, 2003). This technology is very versatile when it comes to mapping independently decodable data blocks, so-called Network Abstraction Layer (NAL) units, into the payloads of individual RTP packets. The video codec MPEG-4/AVC can be divided into a Video Coding Layer (VCL) and the NAL. The VCL fulfils the signal processing tasks such as transformation, quantisation, and motion compensated prediction. The output consists of so-called slices, that contain an integer number of macroblocks. These are then encapsulated by the NAL into corresponding units. Three different packetisation modes can be used to transport these units using RTP: the single NAL unit mode, the non-interleaved mode and the interleaved mode. To maintain conformity with the ITU-T Standard H.241 (ITU-T H.241, 2006), this study has been confined to the single NAL unit mode in which all macroblocks are transported in the decoding sequence and each NAL unit is encapsulated in exactly one RTP packet. Due to the size of the MTU, which in Ethernet-based networks is usually 1500 bytes minus the header overhead, the size of the NAL unit was confined to a maximum of 1400 bytes in this study. This information was passed on to the encoder as a parameter. The transmission of audio data using native RTP is explained in (IETF RFC 3640, 2003) and will not be described here.

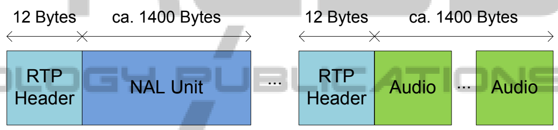

Figure 3: Format of the RTP Packets using Native RTP to carry audio and MPEG-4/AV-encoded video content (single NAL unit mode).

The software tool FFmpeg mentioned in the previous section is capable of encoding the reference video in line with the corresponding maximum NAL unit size and storing it as a byte stream. The individual NAL units within this file can be identified by a unique bit pattern.

4 ANALYSIS SCENARIOS AND RESULTS

Figure 4 shows the numerical investigation environment used in this research.

The AVI file from the company Opticom (Company "Opticom", state 2012), who act as licence holder in Germany for PEVQ, was chosen as the reference file. The file is 8 seconds long with a resolution of 1280x720 (720p HDTV) and a frame rate of 25 fps. As the measurement method for determination of the QoE by IPTV the PEVQ (Perceptual Evaluation of Video Quality) algorithm (Company "Opticom", state 2012) was used.

In the first analysis scenario a lossy IP environment was assumed for the transport of video signals. It was also assumed that packet losses are subject to a binominal distribution and that the burst size is subject to a negative exponential distribution with mean 1. So that the transport via this platform is possible, the video signals are first encoded using MPEG-2 and MPEG-4/AVC (using the default settings of the encoder, i.e. unlimited NAL unit size) and then mapped into the transport stream according to Rec. ISO/IEC 13818-1 (corresponds to the so-called MPEG-2 TS). Figures 5 and 6 show the results obtained here. For all of the following calculations 31 measurements for each determined performance value were used. In this way, it was possible to attain a confidence interval of less than 10 % of the estimated average (with a probability of error of 5 %)

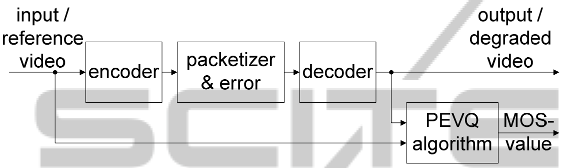

Figure 4: Numerical environment.

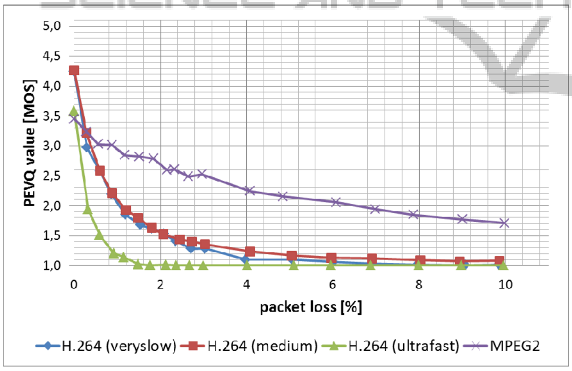

Figure 5: PEVQ values as a function of packet loss for both codecs with MPEG-2 TS and an encoding rate of 5 Mbps.

The curves in Figures 5 and 6 show that the quality of video signals in a loss-free environment improves when the encoding rate is increased. As packet losses increase, the QoE curves will fall more steeply for higher encoding rates than for lower ones. Here too, it could be confirmed that it is perfectly adequate to work with the medium preset in the case of the H.264 codec. The ultrafast preset is extraordinarily sensitive to packet loss: even at a level of packet loss as low as 0.4 % and at an encoding rate of 10 Mbps quality of service drops rapidly to the inacceptable and practically useless value of approx. 1 MOS. It is also obvious that when packet losses are present, the MPEG-2 encoding delivers non-competitive QoE values. Although in a loss-free environment lower QoE values are to be expected in comparison with the MPEG-4/AVC codec, this is very quickly offset when packet losses increase. It is evident that the MPEG-2 TS has been designed and optimised for the transport of video signals encoded according to MPEG-2. The analyses have shown that encapsulation is not to be recommended for the codec H.264. In this case, alternatives must be found. One possible alternative is called Native RTP for MPEG-4/AVC. The following analysis scenarios seek to assess the effectiveness of this alternative.

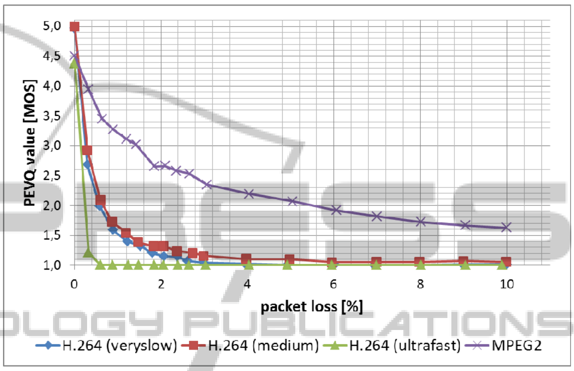

Figure 6: PEVQ values as a function of packet loss for both codecs with MPEG-2 TS and an encoding rate of 10 Mbps.

Figures 7 and 8 show the results obtained in the analysis scenarios described above. For the reasons given in Chapter 3 the single NAL unit mode was used.

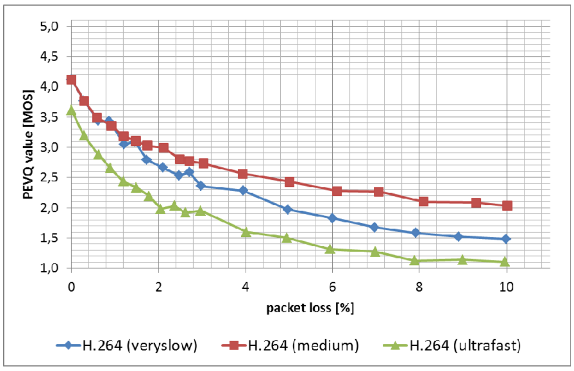

Figure 7: PEVQ values as a function of packet loss for the MPEG-4/AVC codec with Native RTP, a maximum NAL unit size of 1400 bytes and a coding rate of 5 Mbps.

The curves from Figures 7 and 8 show that in the case of native encapsulation of MPEG-4/AVC video content into RTP packets considerably higher QoE values can be achieved than is the case with video content that has been mapped into the MPEG-2 TS. These values are comparable with the qualities gained by using MPEG-2 encoding and mapping into TS ISO 13818-1. These results confirm with hard figures the first, general insights gained from the work described in paper (MacAulay, Felts, Fisher, 2005), in which encapsulation with Native RTP was also investigated. Here too, it is clear that for the codec MPEG-4/AVC it is perfectly adequate to work with the medium preset. The ultrafast preset delivers the worst results by far, and its use should be avoided in practice.

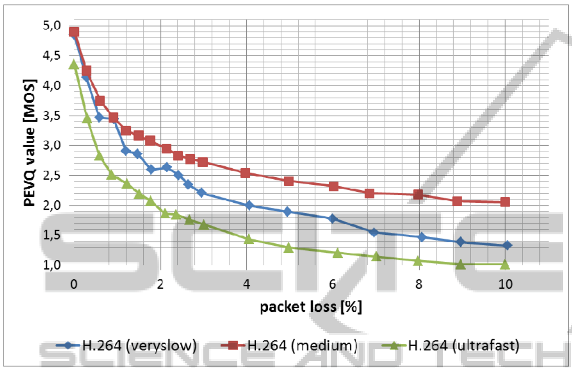

Figure 8: PEVQ values as a function of packet loss for the MPEG-4/AVC codec with Native RTP, a maximum NAL unit size of 1400 bytes and a coding rate of 10 Mbps.

The results gained so far in the course of this study strongly suggest that when the codec MPEG-4/AVC is used, the size of the NAL unit does indeed have a significant influence on QoE. So it makes sense not to use the default mode of the encoder either when using the MPEG-2 TS for contents encoded with MPEG-4/AVC. Instead, the NAL unit size is set to 1400 bytes. All the following analysis scenarios use this setting. For lack of space no further figures are given in this paper. The results obtained here show significantly better QoE values than those gained using the default setting of the codec MPEG-4/AVC (cf. Figs 5 and 6). Here again, the medium preset returns the best QoE values. They are comparable with the levels of quality attained for the MPEG-2 codec. In a loss-free environment the strengths of the MPEG-4/AVC encoder really become evident. It delivers QoE values approx. 0.5 MOS better than the corresponding values for the MPEG-2 codec. Quite clearly it is actually possible to use the MPEG-2 TS to encapsulate MPEG-4/AVC-encoded content as long as the encoder settings have been properly adjusted. This is of immense practical significance.

5 SUMMARY

The focus of this paper has been the subject of quality of service in the service IPTV. A large-scale investigation revealed the strengths and weaknesses of both methods of encapsulating video streams. It became clear that the ISO/IEC 13818-1-formatted transport stream is perfectly suitable for the transport of MPEG-2-encoded video signals. By contrast, MPEG-4/AVC-encoded video signals (using the default settings of the encoder) do have considerable problems with this kind of encapsulation. The study has shown that in this case it makes sense to work either with the encapsulation type Native RTP or, in the case of MPEG-2 TS, to adjust the settings of the encoder (by limiting the size of the NAL unit).

REFERENCES

ISO/IEC 13818-2, 1995. *Information technology -- Generic coding of moving pictures and associated audio information,* http://cutebugs.net/files/mpeg-drafts/is138182.pdf, page last viewed Mai 2012.

ITU-T H.264, 2007. *The Advanced Video Coding Standard*, http://www-ee.uta.edu/Dip/Courses/EE5359/H.264%20Standard2007.pdf, page last viewed Mai 2012.

ISO/IEC 13818-1, 2000. *Generic coding of moving pictures and associated audio information: Systems,* http://mumudvb.braice.net/mumudrupal/sites/default/files/iso13818-1.pdf, page last viewed Mai 2012.

IETF RFC 2250, 1998. *RTP Payload Format for MPEG1/MPEG2 Video*, http://www.ietf.org/rfc/rfc2250.txt, page last viewed Mai 2012.

IETF RFC 3984, 2005. *RTP Payload Format for H.264 Video*, http://www.ietf.org/rfc/rfc3984.txt, page last viewed Mai 2012.

FFmpeg (*current Window builds*), http://ffmpeg.zeranoe.com/builds, page last viewed March 2012.

IETF RFC 3640, 2003. *RTP Payload Format for Transport of MPEG-4 Elementary Streams*, http://www.faqs.org/rfcs/rfc3640.html, page last viewed Mai 2012.

ITU-T H.241, 2006. *Extended video procedures and control signals for H.300-series terminals,* http://www.itu.int/rec/T-REC-H.241-200605-I/en, page last viewed Mai 2012.

Company "Opticom", http://www.opticom.de, page last viewed Mai 2012.

MacAulay, A., Felts, B., Fisher, Y., 2005. *IP Streaming of MPEG-4: Native RTP vs. MPEG-2 Transport Stream*, http://www.envivio.com/files/white-papers/RTPvsTS-v4.pdf, page last viewed Mai 2012.