Quality of Experience in Satellite video streaming transmissions in urban vehicular environment

Alberto Gotta

Wireless Network Lab
CNR-ISTI
Via Moruzzi 1, 56124 Pisa, Italy
alberto.gotta@isti.cnr.it

Erina Ferro

Wireless Network Lab
CNR-ISTI
Via Moruzzi 1, 56124 Pisa, Italy
erina.ferro@isti.cnr.it

Francesco Potortì

Wireless Network Lab
CNR-ISTI
Via Moruzzi 1, 56124 Pisa, Italy
francesco.potorti@isti.cnr.it

Abstract— **In case of video streaming services via satellite towards vehicular clients, very long blockage periods due to road infrastructures, vegetation, and so on, may lead to a complete channel outage, which may result in the loss of all packets transmitted during that period, even in the presence of interleaving and FEC techniques. But if these periods are predictable, as in the presence of known routes traced by means of a GPS navigator, it may be advisable to alert the transmitter in advance, in order to counteract the incoming *outage* interval. We refer to this technique as *Smart Mode*. In the following we will detail how Smart Mode takes advantage of FEC and interleaving techniques, in order to improve the Quality of Experience and to reduce the waste of bandwidth in satellite multimedia streaming.**

Keywords—**Video streaming, Satellite, Mobile, FEC, Interleaver**

I. INTRODUCTION

We consider a hub station, anywhere located, which utilizes a DVB-S2 [1] satellite link to send video streams, together with other information flows (news, advertisements, trains/flights timetables, city and tourist information, bus/train routes and timetables, Internet downloads, music, video, and so on), to a fleet buses in a "personalized" mode: different flows are transmitted to different buses, either in response to explicit requests, or simply to disseminate information.

In this paper, we focus on the video streaming transmission only, which includes real time Video on Demand (VoD), based on the Real Time Streaming Protocol (RTSP) [2], and Live Video Streaming (LVS), such as IPTV [3]. Each bus relays the received streams both as an on-board service to its passengers via dedicated panels or diffusers mounted in the cabinet, or via personal smart phones and multimedia players. The communication bus-hub is done in unicast mode in the VoD case, while there is no restrictions in LVS mode. This work aims at identifying the most suitable techniques for the protection of the information to ensure the most reliable communication for transporting multimedia contents to the vehicle in order to offer both information and entertainment in movement to the passengers. The idea is to spread the information on a longer time span; that is, contiguous blocks of information are transmitted at appropriate time-distant moments in order to make the transmission periods uncorrelated with those of the channel outage. Adding

redundancy to the transmitted data allows to recover the information contents even in the presence of some blocks corrupted when received. This approach is quite commonly used in the context of multimedia streaming. However, it has never been tested and tuned in the mobile satellite communication environment; this makes this study outstanding and innovative.

In [4] we started this study by obtaining some results about the convenience of adopting both an Upper Layer Forward Error Correction (UL-FEC) and an interleaver at the IP level in DVB-RCS/S2 with mobile users. Therefore, the adoption of UL-FEC and interleaving may be desirable, in order to counteract channel's blocked periods. In any case, the presence of interleaving and UL-FEC functionalities introduces a significant delay in the playback of the stream. This delay adds up to the one introduced by the play-out buffer of the video player, whose task is to compensate for the jitter, and which should be filled before starting the video playback. Nevertheless, particularly long blockage periods may lead to a complete channel outage, which may result in the loss of all packets transmitted during the period, overcoming the correcting capacity of the code. If these periods are predictable, as in the presence of known routes traced by means of a GPS navigator, it may be advisable (in the unicast transmission mode, i.e. VoD) to alert the transmitter in advance, and to stop sending packets, which would otherwise be lost, for the specified time interval. We refer to this technique as Smart Mode (SM).

In this paper, we start from the above-described situation, and we propose and analyze parameter settings and transmission techniques for both video on demand and live broadcasting. The main difference in terms of quality of service (QoS) between the two video classes is that the first one may stop playing the stream (and, hence, the transfer), while the second one loses information after an outage period. With live broadcasting, in fact, we cannot pre-load information, which is not available in advance; we can only permit a timeline shift (delay) at start-up, thus introducing a short guard time that can cover just few and short outage periods. In addition to what studied in [5], in this paper we aim at providing video-quality metrics, such as peak signal-to-

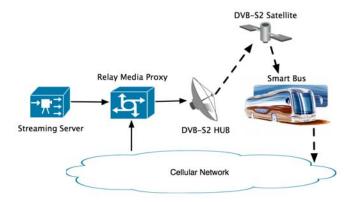noise ratio (PSNR)[(1)], since video quality is tightly coupled with end-user perception.

Figure 1   Video streaming transmission in the "smart" environment

A streaming server (Figure 1) sends the video flows to a Relay Media Proxy, which manages *n* Tx buffers that serve *n* vehicles, each one of these with its own Rx buffer. The Tx and Rx buffers need to store packets, when necessary, during the smart mode operations. We do not deal with the play-out buffer, which is part of the play-out device; we treat the Tx and Rx buffers only. The smart mode behaviour just refers to the set of features in Video on Demand (VoD), driven by feedback commands, that allows the hub to stop the packet transmission during long periods of channel blockage of the bus, and to resume it either at normal rate (r_n) or at a higher rate (r_h), when the receiving buffer of the bus has to be refilled; we assume, in fact, that the hub is able to change the average speed of each streaming connection by changing the transmission frequency of the relevant packets. In good periods, the speed r_h is sufficient to fill the bus Rx buffer as fast as possible. The lower r_h, the higher the probability that the buffer is not filled up before a blocking period; thus, the system could not manage a blocking period longer than the time necessary to fill the receiving buffer at r_h speed, with a resulting increase in the mean PER. It is worth noting that increasing r_h above certain values does not produce any further benefit in terms of PER; therefore, in this paper we will investigate an appropiate threshold for r_h, thus balancing the qualitative satisfaction metric and the waste of allocated resources to the service.

The paper is organized as follows. In section II we provide the reference scenario; section III details on encoding, decoding and interleaving features; section IV provides simulation results; conclusions are addressed in Section V.

II.   REFERENCE SCENARIO

A bus can move in urban or suburban environments, which present very different characteristics both regarding the vehicle speed and the horizon profile, mainly due to buildings.

One of the goals of this paper is to quantify the packet error rate experienced in the non-blocked states of the channel. This is possible by choosing the best operating conditions, i.e., the E_c/N_0 (channel bit energy over two-sided noise power spectral density) ratio, the modulation and coding scheme (MODCOD), and the Reed-Solomon Erasure [6] block length and code rate for the video streaming transmissions. To model the mobile satellite channel at the packet level, we adopted the 3-state Discrete-Time Markov Chain propagation channel model provided by the European Space Agency (ESA) for the Ku-band satellite channel available to mobile users [7], [8]. In each environment, the three states of the Markov chain are named Line of Sight (LoS), Shadowed (Shd), and Blocked (Blk), respectively.
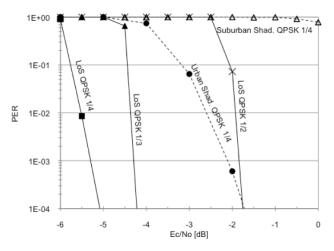

Figure 2   Urban and suburban Packet Error Rate for mobile DVB-S2 in LoS and Shadowed states, at different MODCODs

In [4] we simulated a standard DVB-S2 transmission/reception chain, thus including LDPC and BCH [9], [10] FEC codes, simulating the ESA channel model. In that work we found that by choosing the appropriate MODCOD for the available value of E_c/N_0 (Figure 2), the channel behaviour at the packet level can be seen as a 2-state Discrete-Time process. The two states are a "*Good*" one (LoS) and a "*Bad*" macro-state (Shd + Blk). The urban case has been well studied in [5] and the conclusions are that the study of the urban case can be characterized as follow: when E_c/N_0 is about -2 dB and the MODCOD is QPSK 1/4 (i.e., E_b/N_0 higher than 4 dB), the PER in the Shadowed state is close to 10^{-4}, which is equivalent to that one of a Good state; with QPSK 1/2, which allows operating at double speed with respect to QPSK 1/4, the PER in the Shadowed state is close to 1, which is equivalent to that one of a Bad state.

III.   ENCODING, DECODING, AND INTERLEAVING FEATURES.

Live Video Streaming (LVS) refers to those streaming services that produce live data that cannot be delayed so much, such as radiocast, videocast, news, and so on. In LVS, a live streaming server sends a multimedia flow to the satellite hub, which forwards the contents to vehicles; the VoD case

stands for a server-to-client file transfer with quasi real-time playback. In both cases, we arrange Tx and Rx buffers at the IP level in order to store k information packets of each codeword that are handled within the RSE encoder and decoder. The interleaver, used to scramble the packets before sending them on the channel, is concatenated to the encoder and acts as a large matricial buffer, where each row is made up of k information packets plus n-k redundancy packets that compose an n-ary codeword. Each column of i-length – where i is the interleaver depth – is made up of i packets, which are taken from i different codewords, and it represents the block of data that is sent on the wireless channel. Thus the storing capacity of the interleaver is n x i. In this study we do not focus on the type of interleaving technique (i.e. convolutional, block, helical, etc), because we our interest here is concentrated in its logical function only.

At the vehicle's receiver, both correctly demodulated DVB packets and erasures fill the de-interleaver matrix. Just to give an idea of the decoding process, i codewords will be ready for the decoding process after receiving no less than n transmitted blocks, in order to fulfill the de-interleaver matrix. The RSE (n,k) decoder is able to rebuild k information packets if any k, out of a block of n packets (codeword), are correctly received. In case of *systematic codes*[2], k information packets are followed by n-k redundant ones. Thus, if the first k-ones are correctly received, no decoding actions are needed. If more than n-k erasures out of n occur, the decoder delivers to the player the received information packets only, among the k-transmitted[3]. In LVS, the Minimum Bandwidth Requirements (MBR) on the satellite link is approximately equal to the source rate for the streamed flow multiplied by the channel coding rate, which accounts for both UL-FEC at the IP level and MODCOD applied at the physical level [1]. VoD, instead, is lower bounded by the MBR of live broadcasting, but the file transfer can exceed such a rate, according to the resources available.

Blocked periods represent a waste of bandwidth in both cases, but particularly in VoD, when bandwidth allocation is left free to float over the MBR. When the bandwidth is a precious resource as in satellite and, generally, in wireless communications, such a waste may be disadvantageous, particularly in those cases where pricing policies are applied to the amount of data transferred.

The Smart Mode works well and is suitable in unicast transmissions, while multicast, broadcast and live communications cannot take advantage of the SM features, but need of additive reliable transport protocols [11] that do not conflict with UL-FEC and interleaving techniques. In the following we compare the SM features with a simplified

method (that can be used in multicast and broadcast communication and referred to as Normal Mode, NM) that accounts for interleaver and UL-FEC only.

We note that the latency at the beginning of video playback (*stdelay*) is the sum of three components (in different proportions, according to the various cases): the interleaver filling time, the RSE coding/decoding time, and the Rx buffer filling time (*tstore*). In other words, stdelay is the end-to-end delay in the video reproduction that we pay to reduce the overall packet loss within an acceptable limit. We call *outage* the set of time periods in which certainly there are no good packets for video reproduction; this occurs at all times, when the bad channel duration exceeds the stdelay interval. We also account for a guard time *gaptb* in order to manage the re-establishment of the communication between the bus and the satellite after a blockage period.

In LVS invalid packets fill the whole de-interleaver and no good packet gets out of the RSE decoder, thus resulting in data flow corruption; in VoD, instead, the player is paused up to the reception of new data.

In Normal Mode, stdelay corresponds to the interleaver filling time at speed r_n plus the RSE coding/decoding time. The packet loss is the one resulting by adopting an RSE (n,k) coder plus a n x i interleaver of depth i. In Smart Mode, the time needed to fill the Rx buffer is equal to stdelay, decremented by the interleaver filling time and the RSE coding/decoding time. The interleaver depth can be kept smaller than in NM, because it must only compensate for shorter channel interruptions (up to a certain threshold that we denoted by *maxtb*); emptying the Rx buffer, while the transmission is suspended, compensates for longer interruptions. At the beginning of the process, the Rx buffer filling is done by asking the hub to transmit at speed r_h. When the Rx buffer is completely filled, the reproduction can begin.

When, along the route, the bus predicts a channel blockage longer than *maxtb* seconds, it sends the hub a stop signal, in suitable advance (according to the vehicle speed); the hub suspends the transmission and the Tx buffer begins to be filled. At the bus receiver side, the Rx buffer empties out during the whole duration of the blockage period, and the video reproduction seamlessly continues. When the blockage period ends, the receiver asks the hub to resume the transmission at speed r_h, up to the complete refilling of the Rx buffer. In this case, we stress that the r_h rate is mandatory, because we need to fill up the Rx buffer between two subsequent blockage periods, otherwise we could incur into an outage period.

IV. SIMULATION RESULTS

Performance evaluation has been obtained by developing ad hoc simulative tools within the Mathematica 7 framework. Numerical results are calculated with a 5% maximum confidence interval at 95% level. Table I reports the values of the most important parameters used in the simulations.

In smart mode, the rate r_h should be sufficient to fill the Rx

---

[2] In coding theory, a systematic code is one in which the input data are embedded in the encoded output. Such codes are used to add redundant information to data.

[3] The advantage in using systematic codes is that the redundancy is piggybacked to information packets within a codeword, thus, even if the codeword is not properly decoded information data can still be recovered in part.

buffer as fast as possible within good periods. The *outage* occurs any time the Rx buffer has been emptied and thus there are no more data to play. The buffer filling problem implies a higher outage probability, when the duration of a blockage period is longer than *maxtb*; in the case of blockage periods up to *maxtb* seconds, the buffer filling is treated by the system as a packet loss event and thus UL-FEC and interleaver are used to recover the information. Figure 3 shows the PER and the outage probability as functions of the r_h/r_n ratio.

TABLE I – SIMULATION PARAMETERS

Parameter	Value
Packet rate	160 MPEG pkt/s
stdelay	9 s (sub-urban)
	14 s (urban)
gaptb	0.5 s
maxtb	0.5 s
UL-FEC Codeword	(30, 10) MPEG pkt

In addition to the ACM technique (e.g., MODCOD adaptation), UL-FEC and interleaving may greatly improve the transmission efficiency; these techniques are particularly effective in mobile channels, which are characterized by blockage periods. In any case, the presence of both interleaving and UL-FEC functionalities introduces a significant delay in the reproduction of the stream. In our smart mode case, these functionalities are employed to compensate for channel blockages shorter than *maxtb*, while longer blockages are absorbed by the Rx buffer in the bus, which is filled before starting the reproduction. Therefore, there is a compromise between the interleaver depth (which, multiplied by the size of the corresponding codeword, produces the interleaving storage) and the buffer length; the sum of these two items contributes to the maximum start-up delay (stdelay).
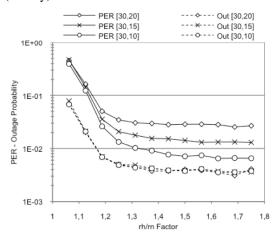

Figure 3   Packet Error Rate and outage probability vs rh filling rate, for a 30-packet codeword in urban QPSK 1/4.

In Figure 4, the residual PER characteristics are reported as a function of the start-up delay, which corresponds to a given interleaver depth (right axis); the figure shows the highway case for different values of the codeword length and coding rate. The comparison is made with sources of independent errors that have the same average PER as the bursty error

channel. Figure 5 refers to the urban case, the most significant one, where very long blockage periods are possible; these impose choosing a long stdelay time (assumed 14s here), in order to yield an acceptable QoS. We compared PER (computed in non-outage periods), *Outage* probability (which accounts for outage periods), and bandwidth occupancy, as functions of the applied RSE code. In case of LVS, *outage* events behave as bursts of lost packets (i.e. not delivered) and they have to be considered in addition to the PER. In case of VoD, *outage* events produce the temporary pause of the timeline, thus without losses, but incurring into a negligible "hiccup" of the play-out.

As we anticipated, the urban case is the most challenging. As previously stated, we can adopt two different MODCODs – QPSK 1/2 and 1/4, respectively - which produce different overheads. Hence, the relative bandwidth occupancy is very different in the two cases. Figure 5 reports also two different curves relevant to the bandwidth occupancy with QPSK ½ and ¼, both with (10,k) RSE code.
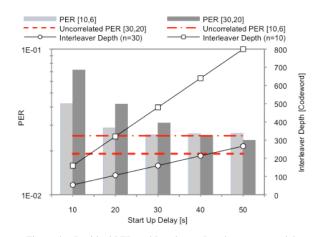

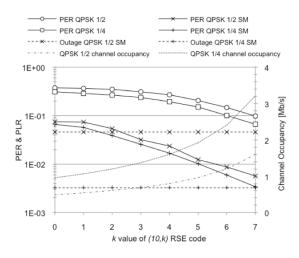Figure 4   Residual PER and Interleaver Depth vs. start-up delay.

Figure 5   Packet Error Rate, Outage probability, and bandwidth occupancy with RSE (10,k) in urban environment.

The most relevant result is that the smart technique outperforms the coupling of UL-FEC and interleaving in all

environments, thus providing a notable gain in terms of PER and *outage* probability.

Without the smart technique, it is impossible to get a PER below the threshold of 2%, which is below the key performance parameter suggested by ITU G.1010 [12] for a video with data rate between 20 and 384 Kbps. There is a trade-off between bandwidth occupancy and *outage* probability in choosing between QPSK ¼ and ½ . In fact, although the 2% threshold is reached with 5 and 6 redundancy packets, in QPSK ¼ and ½, respectively, the bandwidth occupancy is much higher in the former case; this is because of the adoption of ¼ coding rate, which halves the net bandwidth with respect to QPSK ½ . *Outage* is much more likely with QPSK ½ than with QPSK ¼. However, QoS metrics are not sufficient when video and multimedia streaming applications are concerned. Thus, we need an estimation of the video quality, by choosing an appropriate metrics for evaluating the quality of the streamed video in terms of satisfaction on the perceived video quality (Quality of Experience[4]).
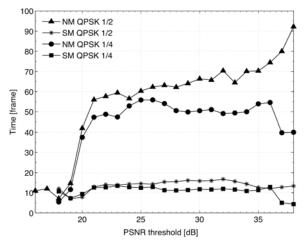

Figure 6   PSNR of "City" video in NM and SM

We chose the RSE (10,5) that keeps the PER below the 2% threshold in SM. Then we produced several video traces in four cases: sequences received in Normal Mode and in Smart Mode both with QPSK ½ and ¼ .

The PSNR in Figure 6 has been computed on the standard "city" test sequence in CIF[5] format which has been coded in MPEG2 TS format at 256 Kb/s constant bitrate using the ffMPEG coder [14] and replicated over 75 times, for a total of over 20000 frames for each of the four depicted cases. A custom script has been used which realistically emulates the TS corruption over the channel, creates frozen frames when appropriate andthen computes the PSNR of the corrupted TS with respect to the uncorrupted one. The script is available by request to the authors.These results show how long (time reported in frame unit) the PSNR has been below the threshold reported on the x-axis. Once again, SM outperforms both with QPSK ½ and ¼, just keeping the PSNR below any threshold no longer than 20 frames.

V.   CONCLUSIONS

This work provides a brief description of a set of features that provide an acceptable level of QoE in video streaming applications for users in movement. These features neither do modify the play-out device, thus not requiring more advanced decoding techniques, nor do impose gap filling techniques to preserve the desired QoE. The Smart Mode technique requires only software upgrade on the gateway units, both onboard and on the server side. Future works will move the attention on broadcast and multicast reliable techniques in order to complete the study of video streaming services toward mobile clients in integrated satellite networks.

REFERENCES

[1]   EN 302 307 - ETSI, DVBS2-74r15, 2004.

[2]   Real Time Streaming Protocol, http://tools.ietf.org/html/rfc2326.

[3]   http://www.openiptvforum.org/.

[4]   Paolo Barsocchi, Alberto Gotta, "Experimental video broadcasting in DVB-RCS/S2 with land mobile satellite channel: a reliability issue", Proc. IEEE International Workshop on Satellite and Space Communications (IWSSC 2008), Toulouse, France, Oct. 2008, http://www.tesa.prd.fr/iwssc08/.

[5]   Nedo Celandroni, Franco Davoli, Erina Ferro, Alberto Gotta, "Smart Buses Used as Access Points to Ubiquitous Broadband Satellite Services in Urban, Suburban and Highway Environmental", in proceeding of KACONF 2008, pp. 653-660, Matera, Italy, Sept., 24-26, 2008

[6]   J. Nonnenmacher, E. W. Biersack, D. Towsley, "Parity-based loss recovery for reliable multicast transmission", IEEE/ACM Trans. Networking, vol. 6, no. 4, pp. 349-361, 1998.

[7]   ESA Project "Feasibility study of a mobile Ku-band terminal", Contract No. 15593/01/NL/DS, Final Report available at http://telecom.esa.int/feasibility_ndsatcom.

[8]   S. Scalise, H. Ernst, and G. Harles, "Measurement and modeling of the land mobile satellite channel", IEEE Trans. Vehic. Technol., vol. 57, no. 2, pp. 693-703, March 2008.

[9]   R. Gallager, "Low density parity check codes", IRE Trans. Inform. Theory ,1962.

[10]   I. S. Reed, X. Chen, Error-Control Coding for Data Networks, Kluwer Academic Publishers, 1999, ISBN 0-7923-8528-4.

[11]   IETF RFC 3940 "Negative-acknowledgment (NACK)-Oriented Reliable Multicast (NORM) Protocol", nov. 2004.

[12]   ITU-T Recommendation G.1010: "End-user multimedia QoS categories", URL: http://www.itu-t.org.

[13]   Xiph.org:http://media.xiph.org/video/derf/

[14]   ffMPEG: http://ffmpeg.org/

---

[4] QoE, in the context of telecommunications is a purely subjective measure from the user's perspective of the overall value of the service provided. QoE cannot be taken a simply the effective quality of the service but must also take into consideration every factor that contributes to overall user value such as suitableness, flexibility, mobility, security, cost, personalization and choice. Apart from its being user dependent, it is invariably be influenced by the user's terminal device, his environment, his expectations, the nature of the content and its importance.

[5] CIF (Common Intermediate Format), is a format used to standardize the horizontal and vertical resolutions in pixels of YCbCr sequences in video signals, commonly used in video teleconferencing systems. It was first proposed in the H.261 standard.