

Research Article

Protection of Video Packets over a Wireless Rayleigh Fading Link: FEC versus ARQ

Julie Neckebroek, Frederik Vanhaverbeke, Danny De Vleschauer, and Marc Moeneclaeys

Department of Telecommunications and Information Processing (TELIN), Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium

Correspondence should be addressed to Julie Neckebroek, julie.neckebroek@telin.ugent.be

Received 1 October 2007; Revised 25 March 2008; Accepted 8 May 2008

Recommended by David Bull

Video content can be provided to an end user by transmitting video data as a sequence of internet protocol (IP) packets over the network. When the network contains a wireless link, packet erasures occur because of occasional deep fades. In order to maintain a sufficient video quality at the end user, video packets must be protected against erasures by means of a suitable form of error control. In this contribution, we investigate two types of error control: (1) forward error correction (FEC), which involves the transmission of parity packets that enables recovery of a limited number of erased video packets, and (2) the use of an automatic repeat request (ARQ) protocol, where the receiver requests the retransmission of video packets that have been erased. We point out that FEC and ARQ considerably reduce the probability of unrecoverable packet loss, because both error control techniques provide a diversity gain, as compared to the case where no protection against erasures is applied. We derive a simple analytical expression for the diversity gain resulting from FEC or ARQ, in terms of the channel coherence time, the allowable latency, and (for FEC) the allowable overhead or (for ARQ) the time interval between (re)transmissions of copies of a same packet. In the case of HDTV transmission over a 60 GHz indoor wireless link, ARQ happens to outperform FEC.

Copyright © 2008 Julie Neckebroek et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

The internet protocol (IP) allows the provision of a mix of multimedia services (video, audio, voice, data, gaming, etc.) to an end user, by breaking up the bitstreams generated by the various services into IP packets and sending these packets over the network. In this contribution, we consider the delivery of these multimedia services via a wireless channel, and focus on the reliability of the received video data.

The occurrence of fading on wireless channels makes reliable transmission a difficult task, because occasional deep fades give rise to bursts of bit errors at the receiver. IP packets affected by bit errors are erased at the receiver, yielding lost packets at the destination. These lost packets are likely to cause visual distortions when viewing the video content at the destination. Hence, in order to obtain a sufficient quality of experience (QoE) it is imperative to limit the video packet loss rate.

In addition, the frequency selectivity of the wireless channel distorts the transmitted signal. In order to cope with frequency selectivity, we resort to a multicarrier modulation

(orthogonal frequency division multiplexing (OFDM)) [1], which turns the frequency-selective channel into a number of parallel frequency-flat channels.

In order to alleviate the damaging impact of fading, one can reduce the probability of bit errors by means of coding on the physical (PHY) layer. Not only the video, but also the other services that are provided via the same wireless link stand to benefit from this coding. In this contribution, we restrict our attention to orthogonal space-time block codes [2–4], for which the optimum decoding reduces to linear processing and simple symbol-by-symbol detection. When this PHY layer coding is not sufficient to yield a satisfactory QoE related to video, additional protection of the video packets must be envisaged.

In order to provide additional protection of the video packets against erasures, one can resort to forward error correction (FEC) coding [5, 6] or to automatic repeat request (ARQ) protocols [7, 8]; these techniques involve the transmission of redundant packets (in addition to the video information packets) or sending a request for retransmitting erased video packets, respectively. Various proposals have

been formulated for protecting packets against erasures by means of FEC [9–12]; in this contribution we select reed-solomon (RS) codes, because they are able to recover the maximum possible number of erasures for a given transmission overhead [5, 13]. As far as ARQ protocols are concerned, we consider selective repeat (SR) ARQ, which yields the minimum transmission overhead [7, 8]. It is important to keep in mind, however, that these techniques come with a cost. First, both FEC and ARQ introduce transmission overhead (usually higher for FEC than for ARQ) and some latency. Second, there is a complexity increase: ARQ requires a retransmission buffer and a return channel from the receiver to the retransmitting network node, and FEC needs additional encoding/decoding operations.

In this contribution, we investigate to what extent the combination of the RS code or the SR ARQ protocol with the space-time PHY layer code improves the reliability of the video transmission over a wireless channel subject to Rayleigh fading. The paper is organized as follows. In Section 2, we introduce some basic concepts about video compression and transmission over an IP network, and describe the space-time coding on the PHY layer. We detail in Section 3 the RS erasure coding and the SR ARQ protocol that are used as additional protection of the video packets against erasures. We provide in Section 4 the error performance analysis for various scenarios, involving space-time coding or no coding on the PHY layer, with or without protection (RS coding or SR ARQ) of the video packets. In Section 5, we present numerical results, including a case study pertaining to HDTV transmission over a 60 GHz indoor wireless link. Finally, in Section 6 conclusions are drawn regarding system performance and complexity, and some generalizations of the considered assumptions are briefly discussed. A major conclusion is that RS erasure coding and SR ARQ yield the same maximum possible diversity gain, which is determined by the ratio of the allowed latency and the channel coherence time; however, this maximum cannot be achieved because of practical constraints on the allowed overhead (RS erasure coding) or when the time interval between retransmissions exceeds the channel coherence time (SR ARQ).

2. VIDEO SOURCE CODING AND TRANSMISSION

In this section, we describe the video packet transmission from the video server to the end user. First, the video source coding method is considered. Next, the different layers in the protocol stack of the OSI-model, that are relevant to this research, are presented.

2.1. Video source coding

The video stream is encoded (compressed) according to the MPEG-2 standard [14, 15], which is commonly used as the format for digital television. The Video section of MPEG-2 (part 2) is designed to compress the video stream through appropriate coding by exploiting the existing redundancy in space and time. Uncompressed video can be seen as a sequence of picture frames (e.g., 25 frames

per second). Typically, the scenes in successive pictures are very similar. One can take advantage of this similarity to compress the video into three types of frames: intracoded frames (I-frames), predictive-coded frames (P-frames), and bidirectional-predictive-coded frames (B-frames).

An I-frame is a compressed version of a single uncompressed frame. The compression is achieved by exploiting the spatial redundancy in the image and the insensitivity of the human eye to certain changes in the image. P-frames, on the other hand, achieve a higher compression because they take advantage of the resemblance between the picture in the current frame and the picture in the previous I- or P-frame. B-frames are compressed by exploiting both the picture in the preceding I- or P-frame as well as the picture in the following I- or P-frame. These B-frames achieve an even higher compression rate. A commonly used frame pattern is IBBPBBPBBPBB, called a group of pictures (GOPs), which consists of 12 compressed frames and which is repeated. Such a GOP has a duration of 480 milliseconds (25 frames per second).

As the different types of frames achieve different compression rates, their resulting sizes, measured in bits, are not equal. I-frames are larger than P-frames, which in turn are larger than B-frames. Their exact sizes depend on the video content. Typically, the average sizes of I- and P-frames are about 6 and 2 times the average size of a B-frame.

Because of the interdependence of the compressed frames, error propagation occurs: an erroneous I- or P-frame results in errors (after decoding) in the 2 preceding B-frames and in all following frames up to (but not including) the next I-frame. Hence, when an I- or P-frame in a GOP is affected by unrecoverable transmission errors, a visual distortion is likely to occur when viewing the video content. Errors in a B-frame do not propagate to other frames. Hence, when only a B-frame in a GOP is affected by unrecoverable transmission errors, it is possible that no visual distortion occurs through the use of error concealment techniques that exploit the similarity between the erroneous B-frame and surrounding frames.

2.2. Protocol stack

Let us consider the case where video data is sent from the video server to the end user, as shown in Figure 1. A source, the video server, broadcasts the video data. Via an aggregation network, this video data reaches a digital subscriber line access multiplexer (DSLAM). The DSLAM sends the data related to a mix of services (video, audio, voice, data, gaming, etc.), over a digital subscriber line (DSL) [16] to the user home gateway (HG). From the HG, the video data is sent through a wireless LAN to the set top box (STB). Figure 1 also displays the different layers of the protocol stack, that are involved in the operation of each of the network nodes. The network nodes are not able to process information from other layers.

2.3. Application layer

The system section of MPEG-2 (part 1) [15] describes how MPEG-compressed video and audio data streams

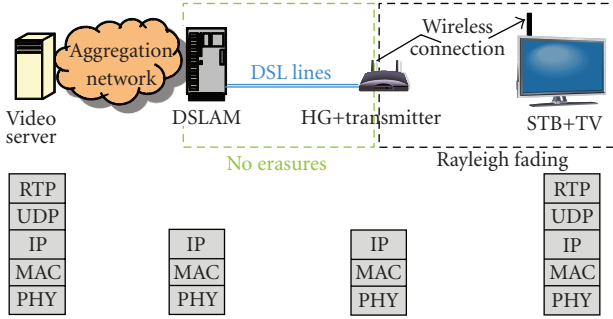


FIGURE 1: Concatenation of DSL connection and wireless connection (DSLAM = digital subscriber line access multiplexer, HG = home gateway, STB = set-top box).

(along with other data, such as teletext, elementary stream identifiers) are multiplexed together to form a single data stream. Basically, the resulting transport stream (TS) consists of a sequence of MPEG-TS packets, that consist of 188 bytes each (including a 4-byte header).

2.4. Session layer

The real-time transport protocol (RTP) [17] is used to deliver audio and video over the Internet. The RTP packets are filled with an integer number of TS packets. In commercial equipment, an RTP packet typically contains 7 TS packets, which is the maximum number of TS packets that fits inside an Ethernet frame (data link layer). The header of an RTP packet contains, among other things, a sequence number and a time stamp. This allows the detection of missing or out-of-order delivery of RTP packets and to perform synchronization, respectively. The header inserted by this protocol is 12 bytes long.

2.5. Transport layer and network layer

The user datagram protocol (UDP) is used on the transport layer to deliver the RTP packets. UDP is well suited for time-sensitive applications that prefer dropped packets to excessively delayed packets.

The UDP packets are passed to the underlying layer, the network layer. This layer uses the IP protocol to deliver the data from source to destination.

2.6. Data link layer

On the medium access control (MAC) sublayer of the data link layer, a header and trailer are added; the latter contains a cyclic redundancy check (CRC). This CRC allows the detection of packets that are corrupted by transmission errors; corrupted packets are not forwarded to the network layer, but are discarded (“erased”). We assume that no ARQ is applied on the MAC layer; the effect of ARQ on the MAC layer is briefly discussed in Section 6.

The structure of a data-link-layer packet is visualized in Figure 2. The packet contains 7 MPEG-TS packets, and the

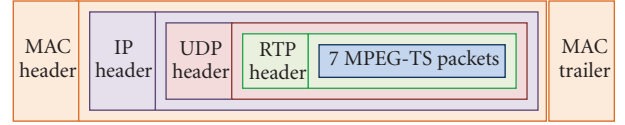


FIGURE 2: The video data is nested in a structure of packets, each packet and corresponding header results from a different layer in the protocol stack.

various headers/trailers that have been added by the different layers in the protocol stack.

2.7. Physical layer

As far as the physical (PHY) layer is concerned, we only consider the wireless link between the HG and the STB. On the PHY layer of the HG transmitter, the L bits to be sent for every data-link-layer packet are mapped onto an M -point signal constellation. The resulting M -ary data symbols are transmitted at a rate R_s (in symbols per second) over the wireless channel; hence the duration of a packet equals $L/(R_s \log_2(M))$. The transmission makes use of orthogonal frequency-division multiplexing (OFDM) [1]. The sequence of data symbols at rate R_s is demultiplexed into N_c parallel symbol streams, each of rate R_s/N_c . These N_c symbol streams are modulated onto N_c distinct subcarriers, that have a frequency separation of (slightly more than) R_s/N_c , and the sum of these modulated subcarriers is transmitted. The transmitted signal can be viewed as a sequence of OFDM blocks. As shown in Figure 3, an OFDM block has a duration of N_c/R_s , and contains N_c data symbols (i.e., one symbol on each of the N_c subcarriers). The bandwidth occupied by the resulting transmitted signal is (slightly more than) R_s . The transmission of an L -bit packet involves $L/(N_c \log_2(M))$ OFDM blocks. Typically, the number N_c of carriers is on the order of 100 to 1000. Because of the large number of subcarriers, OFDM turns the wireless fading channel into a set of N_c flat-fading parallel channels.

For each subcarrier, the fading gain is assumed to be piecewise constant over time; the fading gain does not change over a time interval equal to the channel coherence time T_{coh} , and is statistically independent of the fading gain in other intervals of duration T_{coh} . During an interval T_{coh} , several packets are transmitted, as indicated in Figure 4. Packets from other applications are located in between the packets with video data.

On the PHY layer of the STB receiver, the M -ary data symbols are detected, and demapped to bits. On the MAC sublayer, the recovered bits are grouped into packets of size L , and error detection based on the CRC is performed. When an error is detected, the packet is erased; otherwise, the packet is passed to the higher layers.

Because of fading, the received signal is occasionally strongly attenuated. To alleviate the damaging impact of fading on the detection of the M -ary data symbols, we consider the use of multiple transmit and receive antennas. A multiple-input multiple-output (MIMO) system with N_t transmit and N_r receive antennas allows the introduction

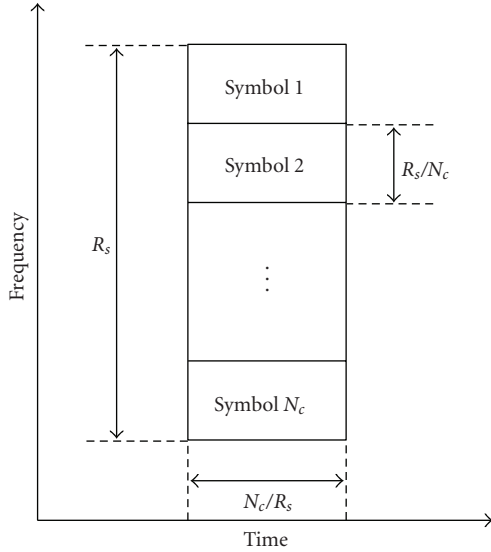


FIGURE 3: Representation of an OFDM block in time and frequency.

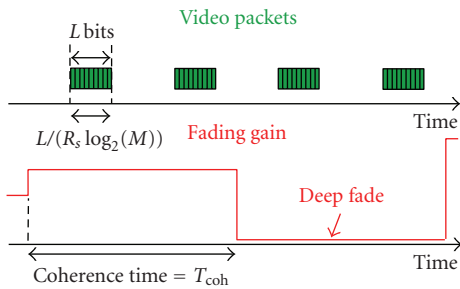


FIGURE 4: Video packet stream and fading gain versus time; in this example, 2 video packets are transmitted during the channel coherence time, in which case a packet group consists of 2 packets.

of space-time coding [2–4]. Whereas an uncoded single-input single-output (SISO) system, that is, $N_t = N_r = 1$, provides only one wireless link between the HG and the STB, the number of wireless links provided by an orthogonal space-time block-coded (OSTBC) MIMO system equals $N_r N_t$. As compared to an SISO system, the larger number of links resulting from OSTBC MIMO gives rise to a considerably higher robustness against fading, and a much better error performance. Using an OSTBC MIMO system does not require additional bandwidth as compared to the SISO system, but comes at a substantial hardware cost that increases with the number of antennas. The space-time coding only marginally increases the latency. Optimum decoding of OSTBC MIMO reduces to linear processing and simple symbol-by-symbol detection at the receiver.

In this paper, we will consider the Alamouti space-time code [2], which requires 2 transmit antennas (and an arbitrary number N_r of receive antennas). Denoting by $s_n(t)$ the signal that corresponds to the n th OFDM block, Alamouti space-time coding involves the transmission of two OFDM blocks during two consecutive intervals (each of

duration N_c/R_s) on two antennas, according to the following scheme:

$$\begin{aligned} \text{interval } 2i: & s_{2i}(t) \text{ (on antenna 1)} \\ & s_{2i+1}(t) \text{ (on antenna 2),} \\ \text{interval } 2i+1: & -(s_{2i+1}(t))^* \text{ (on antenna 1)} \\ & (s_{2i}(t))^* \text{ (on antenna 2),} \end{aligned} \quad (1)$$

where $()^*$ denotes complex conjugate. Hence, each OFDM block $s_n(t)$ reaches the receiver via $2N_r$ wireless links.

3. ADDITIONAL PROTECTION OF THE VIDEO DATA

As mentioned before, packets yielding an erroneous checksum are discarded (erased) on the MAC layer, because they have been affected by transmission errors; the other packets are assumed to be received correctly. Because of video packet erasures, visual distortions may occur when viewing the received video content. In order to guarantee a sufficient QoE to the end user, the rate of video packet erasures should be limited. When the packet erasure rate caused by transmission errors on the wireless link is too large, additional measures are needed to recover erased video packets. In this contribution, we consider the combination of a PHY layer with either no coding or Alamouti space-time coding with 1 or 2 receive antennas, and additional packet protection by means of either RS erasure coding or SR ARQ.

3.1. RS erasure coding

The RS code is defined over the Galois field $GF(2^q)$, which implies that an RS code symbol consists of q bits; typically, $q = 8$. (The RS code symbols are not to be confused with the transmitted data symbols; the former belong to $GF(2^q)$, whereas the latter belong to an M -point signal constellation.) In the sequel, a *video information packet* refers to the MPEG-TS payload (i.e., 7 MPEG-TS packets) of the packet as shown in Figure 2. Per group of K of these video information packets, we transmit $N - K$ parity packets. This results in a packet codeword of N packets. The parity packets are constructed such that taking from each packet the i th block of q bits yields an $RS(N, K)$ codeword, for all $i = 1, 2, \dots, L/q$. This construction is illustrated in Figure 5. Hence, when e packets from the packet codeword are erased, each of the L/q RS codewords is affected by exactly e symbol erasures.

The $RS(N, K)$ code is known to be maximum distance separable (MDS), that is, the code can recover up to $N - K$ erasures, which cannot be outperformed by any other code with the same number $N - K$ of parity symbols (Note that a receiver without an RS decoder can still process the packet stream by simply ignoring the parity packets, at the expense of a performance degradation as compared to a receiver with an RS decoder.) [5, 13]. When the number of erasures is larger than $N - K$, erasure decoding fails and unrecoverable packet loss occurs.

The introduction of erasure coding yields an increase of both overhead and latency.

- (i) Using an (N, K) block code gives rise to a transmission overhead ovh given by $ovh = (N - K)/K$,

because for each K information packets, $N - K$ additional packets must be transmitted. Hence, denoting by R_{pack} (in packets per second) the rate of information packets, the packet transmission rate equals $(N/K)R_{\text{pack}}$. This indicates that because of the coding the fraction of time during which the channel is used for video transmission is increased by a factor N/K , leaving less room for the transmission of packets from other applications.

- (ii) When at most $N - K$ packets are erased, they can be recovered by means of the RS(N, K) code. To perform erasure decoding, at least K packets must be received correctly. Hence, the RS decoder might need to wait until all N packets of the codeword are received, before the erasure decoding can start. Hence, using the (N, K) block code introduces a maximum additional latency T_{lat} which equals the duration K/R_{pack} of a packet codeword. Increasing the latency gives rise to a larger zapping delay, which might unfavorably affect the user's QoE. (The zapping delay is the time that elapses between giving the command to change the TV channel and the appearance of the new TV channel on the screen [18].)

Considering the above, the code parameters N and K should be selected such that the overhead and latency are limited to reasonable values.

It is convenient that the parity packets are generated by the video server, as this is the only network node (besides the STB of the end user) that has access to the video data. In principle, parity packets could instead be generated by the DSLAM or the HG. However, this would require that the DSLAM or the HG has access to the higher protocol layers (beyond IP), which would increase their complexity and cost.

3.2. Selective repeat ARQ

As far as ARQ is concerned, we consider an SR retransmission protocol. The STB receiver sends a retransmission request for each of the erased video packets, and only copies of the erased packets are retransmitted. To limit the round-trip delay, we assume that retransmissions occur from either the DSLAM or the HG. Of course, the functionality of the retransmitting network node needs to be extended beyond the IP layer, in order to be capable of recognizing retransmission requests related to specific video packets; in addition, this node must have a retransmission buffer containing video packets that have not yet been correctly received. Augmenting the functionality of the DSLAM or HG increases their complexity and cost. As the HG is a consumer product, the DSLAM appears to be the economically justified choice for operating as the retransmitting node. However, the HG offers the shorter round-trip delay.

Upon receiving a retransmission request, the retransmitting network node sends a copy of the packet involved. Retransmissions are scheduled such that the time interval T_{retr} between the (re)transmission instants of copies of the same packet is not less than the channel coherence time

T_{coh} . This way, the different copies experience statistically independent fading. When one would select $T_{\text{retr}} < T_{\text{coh}}$, the retransmission of a packet that has been erased because of a deep fade is experiencing the same deep fade, and therefore is likely to be erased as well. Such retransmissions should be avoided, as they are not useful, but rather contribute to the transmission overhead.

The minimum possible time interval $T_{\text{retr}, \min}$ between (re)transmission instants of the same packet is the sum of the packet duration $L/(R_s \log_2(M))$ and the round-trip delay T_{RT} ; the latter is the sum of the two-way propagation delay, the duration of the acknowledgment message, and the processing delays at the receiver and the transmitter [7, 8]. We select $T_{\text{retr}} = \max(T_{\text{retr}, \min}, T_{\text{coh}})$. When $T_{\text{retr}, \min} > T_{\text{coh}}$, this yields $T_{\text{retr}} = T_{\text{retr}, \min}$: the interval between transmission instants is the shortest possible, and (re)transmitted copies of the same packet experience-independent fading. When $T_{\text{retr}, \min} \leq T_{\text{coh}}$, we get $T_{\text{retr}} = T_{\text{coh}}$: the retransmission instant is deliberately delayed by an amount $(T_{\text{coh}} - T_{\text{retr}, \min})$ with respect to the earliest possible retransmission instant, in order that the (re)transmitted copies of the same packet are affected by independent fading gains.

Since each retransmission gives rise to a latency of T_{retr} , the maximum number N_{retr} of allowed retransmissions per packet is given by $N_{\text{retr}} = \lfloor T_{\text{lat}}/T_{\text{retr}} \rfloor$, in order that the total latency caused by the SR ARQ protocol does not exceed T_{lat} .

4. SYSTEM ANALYSIS

In this section, we present the analysis of the system under study. We first investigate the PHY layer, followed by the additional packet protection by means of RS erasure coding or SR ARQ. As a performance measure, we consider the average number of GOPs that are affected by irrecoverable packet loss, over a reference time interval of 12 hours. Finally, analytical results regarding RS erasure coding and SR ARQ are compared.

4.1. PHY layer

We consider the cases of uncoded SISO transmission, and Alamouti orthogonal space-time coding (2 transmit antennas) with 1 or 2 receive antennas. The probability $P_{\text{bit}}(x)$, that a bit is received in error, depends on the instantaneous channel state x . The channel state x is the sum of the squared fading gains that are involved in the transmission of the considered bit (1 fading gain for SISO, and 2 or 4 fading gains for Alamouti with 1 or 2 receive antennas). Limiting our attention to QPSK transmission, $P_{\text{bit}}(x)$ is given by [2, 6]

$$P_{\text{bit}}(x) = \begin{cases} Q\left(\sqrt{\frac{2E_b x}{N_0}}\right) & \text{uncoded SISO,} \\ Q\left(\sqrt{\frac{E_b x}{N_0}}\right) & \text{Alamouti,} \end{cases} \quad (2)$$

where

$$Q(v) = \frac{1}{\sqrt{2\pi}} \int_v^{+\infty} \exp\left(-\frac{u^2}{2}\right) du \quad (3)$$

is the complement of the cumulative distribution function of a zero-mean unit-variance Gaussian random variable. In (2), E_b denotes the transmitted energy per bit of the video packet, and N_0 is the one-sided power spectral density of the noise at the receiver. $P_{\text{bit}}(x)$ equals $1/2$ for $x = 0$, and converges to 0 when $x \rightarrow \infty$; the larger E_b/N_0 is, the faster this convergence occurs. When the fading gains are normalized such that the average energy per bit at each receive antenna also equals E_b , the probability density function $p(x)$ of the channel state is given by [6]

$$p(x) = \frac{x^{D-1} \exp(-x)}{(D-1)!}, \quad (4)$$

with $D = 1$ for uncoded SISO and $D = 2$ or $D = 4$ for Alamouti with $N_r = 1$ or $N_r = 2$. The quantity D is the *diversity* provided by the PHY layer; basically, D equals the number of physical links between the transmitter and the receiver that are exploited by the transmission scheme. As we will shortly demonstrate, the error performance improves with increasing D ; this is intuitively clear, because all D links must fail for a packet erasure to occur.

From (2), the packet erasure probability $P_{\text{pack}}(x)$ conditioned on x equals

$$P_{\text{pack}}(x) = 1 - (1 - P_{\text{bit}}(x))^L. \quad (5)$$

To obtain (5), we have assumed that all N_c subcarriers of the OFDM signal experience the same value of the channel state x , and have taken into account that the packet duration is less than the channel coherence time, so that the channel state is the same for all L bits of a packet. The effect of relaxing this assumption is briefly discussed in Section 6. For $x = 0$, $P_{\text{pack}}(x)$ and $1 - P_{\text{pack}}(x)$ equal $1 - 2^{-L}$ and 2^{-L} , respectively. For $x \rightarrow \infty$, $P_{\text{pack}}(x)$ and $1 - P_{\text{pack}}(x)$ converge to zero and to one, respectively; the speed of convergence increases with increasing E_b/N_0 . Finally, note from (2) that $P_{\text{bit}}(x)$ and $P_{\text{pack}}(x)$ depend on x and E_b/N_0 only through the variable $y = xE_b/N_0$.

Before we consider in the next subsections the cases where RS erasure coding or SR ARQ is used in order to recover erased packets, we now investigate the system performance under the assumption that no such error control measures are taken.

We define a *packet group* as the set of packets that are transmitted consecutively in time during an interval of duration T_{coh} over which the fading is constant. We denote by N_{coh} the number of packets transmitted during the interval T_{coh} . For the example shown in Figure 4, we have $N_{\text{coh}} = 2$. As we consider the case where only information packets and no parity packets are transmitted, we have $N_{\text{coh}} = \lceil T_{\text{coh}} R_{\text{pack}} \rceil$. The probability $P_{\text{group}}(e)$ that e packets are erased within a packet group of size N_{coh} , irrespective of the channel state, is given by

$$P_{\text{group}}(e) = \frac{N_{\text{coh}}!}{e! (N_{\text{coh}} - e)!} \times \int_0^{+\infty} P_{\text{pack}}^e(x) (1 - P_{\text{pack}}(x))^{N_{\text{coh}} - e} p(x) dx, \quad (6)$$

$$e = 0, \dots, N_{\text{coh}}.$$

Considering the behavior of $1 - P_{\text{pack}}(x)$, $P_{\text{group}}(0)$ converges to 1 for large E_b/N_0 . For large E_b/N_0 and $e > 0$, $P_{\text{pack}}^e(x)$ goes to zero much faster than $p(x)$ for increasing x , so that the factor $\exp(-x)$ in (4) can be approximated as $\exp(-x) \approx 1$. Using the approximation in (6) along with the substitution

$$F\left(\frac{E_b x}{N_0}\right) = \frac{N_{\text{coh}}!}{e! (N_{\text{coh}} - e)!} P_{\text{pack}}^e(x) (1 - P_{\text{pack}}(x))^{N_{\text{coh}} - e}, \quad (7)$$

we obtain, for high E_b/N_0 ,

$$P_{\text{group}}(e) \approx \int_0^{+\infty} F\left(\frac{E_b x}{N_0}\right) \frac{x^{D-1}}{(D-1)!} dx$$

$$= \left(\frac{E_b}{N_0}\right)^{-D} \int_0^{+\infty} F(y) \frac{y^{D-1}}{(D-1)!} dy, \quad e = 1, \dots, N_{\text{coh}}. \quad (8)$$

Taking into account that $F(y)$ is not a function of E_b/N_0 , we have $P_{\text{group}}(e) \propto (E_b/N_0)^{-D}$ for $e > 0$.

Let us now compute the probability P_{GOP} that a GOP is affected by unrecoverable packet loss. As no measures are taken to recover erased packets, each erased packet is lost. Denoting by T_{GOP} and N_{GOP} the duration of one GOP and the number of packet groups that fit within the duration of one GOP, respectively, we have $T_{\text{GOP}} = N_{\text{GOP}} N_{\text{coh}} / R_{\text{pack}}$, and

$$P_{\text{GOP}} = 1 - (P_{\text{group}}(0))^{N_{\text{GOP}}}$$

$$= 1 - \left(1 - \sum_{e=1}^{N_{\text{coh}}} P_{\text{group}}(e)\right)^{N_{\text{GOP}}}$$

$$= \sum_{i=1}^{N_{\text{GOP}}} \frac{N_{\text{GOP}}! (-1)^{i-1}}{i! (N_{\text{GOP}} - i)!} \left(\sum_{e=1}^{N_{\text{coh}}} P_{\text{group}}(e)\right)^i \quad (9)$$

$$\approx N_{\text{GOP}} \sum_{e=1}^{N_{\text{coh}}} P_{\text{group}}(e)$$

$$= N_{\text{GOP}} (1 - P_{\text{group}}(0)).$$

The approximation in (9) corresponds to keeping only the term with $i = 1$, which is the dominating term at high E_b/N_0 . Hence, for large E_b/N_0 , we obtain $P_{\text{GOP}} \propto (E_b/N_0)^{-D}$. This illustrates the impact of the PHY layer diversity D : the larger D , the smaller the probability that a GOP is affected by packet erasures.

From (9), we compute the average number $E[\#\text{GOP}_{\text{unrec}}]$ of GOPs that are affected by unrecoverable packet loss in a reference interval T_{ref} of 12 hours. Denoting by N_{ref} the number of GOP intervals in T_{ref} , we have $T_{\text{ref}} = N_{\text{ref}} T_{\text{GOP}} = N_{\text{ref}} N_{\text{GOP}} T_{\text{coh}}$. Hence,

$$E[\#\text{GOP}_{\text{unrec}}] = N_{\text{ref}} P_{\text{GOP}}$$

$$\approx N_{\text{ref}} N_{\text{GOP}} (1 - P_{\text{group}}(0)) \quad (10)$$

$$= \frac{T_{\text{ref}}}{T_{\text{coh}}} (1 - P_{\text{group}}(0)).$$

The approximation in (10) holds for large E_b/N_0 . Note that, at high E_b/N_0 , $E[\#\text{GOP}_{\text{unrec}}]$ is independent of the GOP duration, and proportional to $(E_b/N_0)^{-D}$.

4.2. Packet protection by means of RS erasure coding

Now we consider the case where $(N - K)$ parity packets are added to K information packets, yielding a (N, K) RS packet codeword. The number N_{coh} of packets transmitted during the interval T_{coh} is now given by $N_{\text{coh}} = \lceil (N/K)T_{\text{coh}}R_{\text{pack}} \rceil$, which denotes the size of a packet group. We assume that the N packets of the packet codeword are distributed over N_{group} packet groups, to which we associate the indices $1, 2, \dots$ and N_{group} . We denote by e_n the number of erased packets in the packet group with index n ($n = 1, \dots, N_{\text{group}}$), and introduce the vector $\mathbf{e} = (e_1, \dots, e_{N_{\text{group}}})$. We define by $\Pr(\mathbf{e})$ the probability that the number of erased packets in the groups with indices $1, 2, \dots$ and N_{group} equals e_1, e_2, \dots and $e_{N_{\text{group}}}$, respectively. Assume for simplicity that N is an integer multiple of N_{coh} and that the first packet of the codeword is also the first packet of a packet group; in this case, we have $N_{\text{group}} = N/N_{\text{coh}}$, and each of the packet groups contains exactly N_{coh} packets from the considered codeword. Taking into account that erasures in different packet groups are statistically independent, we obtain

$$\Pr(\mathbf{e}) = \prod_{n=1}^{N_{\text{group}}} P_{\text{group}}(e_n), \quad (11)$$

where $P_{\text{group}}(e)$ is given by (6), but with $N_{\text{coh}} = \lceil (N/K)T_{\text{coh}}R_{\text{pack}} \rceil$. When N is not an integer multiple of N_{coh} and/or the first packet of the codeword is not the first packet of a group, an edge effect occurs: we get $N_{\text{group}} = \lceil N/N_{\text{coh}} \rceil$ or $N_{\text{group}} = \lceil N/N_{\text{coh}} \rceil + 1$, depending on the position of the first packet of the codeword within its packet group; for example, Figure 6 shows a situation with $N = 5$, $N_{\text{coh}} = 3$, and $N_{\text{group}} = 3$. Then (11) must be slightly modified by taking into account that the packet groups with indices 1 and N_{group} might contain fewer than N_{coh} packets from the considered codeword. Recalling that, for high E_b/N_0 , $P_{\text{group}}(e) \propto (E_b/N_0)^{-D}$ for $e > 0$ and $P_{\text{group}}(0) \approx 1$; it follows from (11) that $\Pr(\mathbf{e}) \propto (E_b/N_0)^{-nD}$ with n denoting the number of nonzero entries of \mathbf{e} .

From (11), the probability $P_{\text{RS}}(e_{\text{tot}})$ that e_{tot} erasures occur in the packet codeword is given by

$$P_{\text{RS}}(e_{\text{tot}}) = \sum_{e_1 + e_2 + \dots + e_{N_{\text{group}}} = e_{\text{tot}}} \Pr(\mathbf{e}). \quad (12)$$

Finally, the probability $\Pr(\text{decoding failure})$ that the erasures cannot be recovered by the RS decoder (because e_{tot} is larger than $N - K$) becomes

$$\begin{aligned} \Pr[\text{decoding failure}] &= \sum_{e_{\text{tot}}=N-K+1}^N P_{\text{RS}}(e_{\text{tot}}) \\ &= 1 - \sum_{e_{\text{tot}}=0}^{N-K} P_{\text{RS}}(e_{\text{tot}}). \end{aligned} \quad (13)$$

In order to obtain at least $(N - K + 1)$ erasures in the codeword, at least $\gamma_{\text{RS}} = \lceil (N - K + 1)/N_{\text{coh}} \rceil$ packet groups must contain erased packets; this implies that the vectors \mathbf{e} in (12) must have at least γ_{RS} nonzero entries. Hence, for large E_b/N_0 , $\Pr(\text{decoding failure})$ is proportional to $(E_b/N_0)^{-\gamma_{\text{RS}}D}$. Taking into account that $\text{ovh} = (N - K)/K$, $T_{\text{lat}} = K/R_{\text{pack}}$ and $N_{\text{coh}} = \lceil (N/K)T_{\text{coh}}R_{\text{pack}} \rceil = \lceil NT_{\text{coh}}/T_{\text{lat}} \rceil \approx NT_{\text{coh}}/T_{\text{lat}}$, γ_{RS} can be expressed as

$$\gamma_{\text{RS}} = \left\lceil \frac{N - K + 1}{N_{\text{coh}}} \right\rceil \approx \left\lceil \frac{N - K}{N_{\text{coh}}} \right\rceil \approx \left\lceil \frac{\text{ovh}}{1 + \text{ovh}} \cdot \frac{T_{\text{lat}}}{T_{\text{coh}}} \right\rceil. \quad (14)$$

Note that γ_{RS} is an increasing function of both ovh and T_{lat} .

Now we consider the probability P_{GOP} that a GOP is affected by an unrecoverable packet loss. Denoting by N_{RS} the number of packet codewords in one GOP interval T_{GOP} , we have $T_{\text{GOP}} = N_{\text{RS}}K/R_{\text{pack}}$, and

$$\begin{aligned} P_{\text{GOP}} &= 1 - (1 - \Pr[\text{decoding failure}])^{N_{\text{RS}}} \\ &\approx N_{\text{RS}}\Pr[\text{decoding failure}]. \end{aligned} \quad (15)$$

Similarly, the average number of GOPs that are affected by unrecoverable packet loss during a reference period T_{ref} of 12 hours is given by

$$\begin{aligned} E[\#\text{GOP}_{\text{unrec}}] &= N_{\text{ref}}P_{\text{GOP}} \\ &\approx N_{\text{ref}}N_{\text{RS}}\Pr[\text{decoding failure}] \\ &= \frac{T_{\text{ref}}}{T_{\text{lat}}}\Pr[\text{decoding failure}], \end{aligned} \quad (16)$$

where $T_{\text{ref}} = N_{\text{ref}}T_{\text{GOP}} = N_{\text{ref}}N_{\text{RS}}T_{\text{lat}}$. The approximations in (15) and (16) are valid for large E_b/N_0 . We deduce from (15) and (16) that both P_{GOP} and $E[\#\text{GOP}_{\text{unrec}}]$ are proportional to $(E_b/N_0)^{-\gamma_{\text{RS}}D}$. Hence, as compared to the case where no erasure coding is used, the effect of the RS(N, K) code is to increase the diversity order from D to $\gamma_{\text{RS}}D$: erasure coding introduces a diversity gain of γ_{RS} . According to (14), a tradeoff exists between the achievable diversity gain and the allowable overhead and latency: the smaller the allowable overhead and latency, the smaller the achievable diversity gain.

4.3. Packet protection by means of selective repeat ARQ

With the proposed retransmission strategy, a packet will be lost definitively when it has been erased during the first transmission *and* during N_{retr} successive retransmissions. The probability $P_{\text{ARQ, unrec}}(\mathbf{x})$ of this event is given by

$$P_{\text{ARQ, unrec}}(\mathbf{x}) = \prod_{i=0}^{N_{\text{retr}}} P_{\text{pack}}(x_i), \quad (17)$$

where $P_{\text{pack}}(x)$ is the packet erasure probability corresponding to a channel state x (see (5)), and $\mathbf{x} = (x_0, \dots, x_{N_{\text{retr}}, \text{max}})$ contains the values of the channel state at the first transmission and the subsequent N_{retr} retransmissions of the considered packet. The probability $P_{\text{group, unrec}}(\mathbf{x})$ that at least

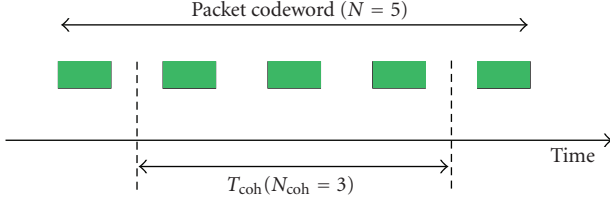


FIGURE 6: Situation where a packet codeword is distributed over 3 packet groups ($N = 5, N_{\text{coh}} = 3, N_{\text{group}} = 3$).

where P_{pack} is the probability that a packet is erased and irrespective of the channel condition

$$P_{\text{pack}} = \int_0^{+\infty} P_{\text{pack}}(x)p(x)dx. \quad (23)$$

For large E_b/N_0 , $P_{\text{pack}} \propto (E_b/N_0)^{-D}$. From (22) we obtain

$$E[\text{ovh}] = P_{\text{pack}} \frac{1 - P_{\text{pack}}^{N_{\text{retr}}}}{1 - P_{\text{pack}}}. \quad (24)$$

For large E_b/N_0 , we have $E[\text{ovh}] \approx P_{\text{pack}} \propto (E_b/N_0)^{-D}$. This indicates that the average overhead resulting from SR ARQ decreases with increasing E_b/N_0 and increasing PHY layer diversity D .

4.4. Comparison of RS erasure coding and selective repeat ARQ

For high E_b/N_0 , given packet transmission rate R_{pack} and a given PHY layer diversity D , the system yielding the largest diversity gain gives rise to the smallest $E[\#\text{GOP}_{\text{unrec}}]$. In the case of RS erasure coding, the highest possible diversity gain $\gamma_{\text{RS,max}}$ equals $\lceil T_{\text{lat}}/T_{\text{coh}} \rceil$, which is achieved for $\text{ovh} \rightarrow \infty$. For SR ARQ, the maximum diversity gain is $\gamma_{\text{ARQ,max}} = 1 + \lceil T_{\text{lat}}/T_{\text{coh}} \rceil$; this gain is obtained when $T_{\text{retr}} = T_{\text{coh}}$, which is the smallest value of T_{retr} that yields statistically independent (re)transmissions of the same packet. Unless T_{lat} is an integer multiple of T_{coh} , we get $\gamma_{\text{RS,max}} = \gamma_{\text{ARQ,max}}$, which indicates that RS erasure coding and SR ARQ yield the same potential diversity gain. However, the achievable diversity gain is limited by practical constraints.

- (i) In the case of RS erasure coding, the allowable overhead ovh is limited by bandwidth constraints. In most practical systems, one imposes the constraint $\text{ovh} < 1$, so that (14) yields $\gamma_{\text{RS}} < \lceil T_{\text{lat}}/(2T_{\text{coh}}) \rceil \approx \gamma_{\text{RS,max}}/2$: under this constraint on the overhead, at most half of the maximum possible diversity gain is achievable.
- (ii) In the case of SR ARQ, $\gamma_{\text{ARQ}} = 1 + \lceil T_{\text{lat}}/\max(T_{\text{coh}}, T_{\text{retr,min}}) \rceil$ so that the maximum diversity gain $\gamma_{\text{ARQ,max}}$ cannot be achieved when $T_{\text{retr,min}} > T_{\text{coh}}$.

Hence, the diversity gain resulting from RS erasure coding is limited by the allowed overhead, whereas in the case of SR ARQ the diversity gain is limited by the ratio $T_{\text{retr,min}}/T_{\text{coh}}$. When $T_{\text{retr,min}} < T_{\text{coh}}$, the system with SR

ARQ yields the largest possible diversity gain $\gamma_{\text{ARQ,max}}$, and outperforms the system with RS erasure decoding. When $T_{\text{retr,min}} > T_{\text{coh}}$, neither RS erasure coding nor SR ARQ achieves the maximum possible diversity gain; when

$$\text{ovh} < \left(\frac{T_{\text{retr,min}}}{T_{\text{coh}}} - 1 \right)^{-1}, \quad (25)$$

the system with SR ARQ outperforms the system with RS erasure coding; otherwise, the system with RS erasure coding yields the better performance. For example, it follows from (25) that RS erasure decoding needs an overhead larger than 50% in order to beat SR ARQ with $T_{\text{retr,min}} = 3T_{\text{coh}}$.

The RS erasure coding introduces a fixed overhead and latency, which are determined by the parameters (N, K) of the RS code. In the case of SR ARQ, the number of retransmissions of a packet is a random number between 0 and N_{tr} . Therefore, the latency and overhead resulting from SR ARQ are also random, with a maximum value determined by N_{tr} , and an average value that decreases with increasing E_b/N_0 and increasing PHY layer diversity D ; typically, these averages are considerably smaller than the fixed overhead and latency resulting from RS erasure coding.

Further, from the complexity point of view, one should take into account that the system with SR ARQ requires the presence of a return channel and an increase of the functionality (beyond the IP layer) of the retransmitting network node (DSLAM or HG). The system with RS erasure coding requires additional complexity for the construction (at the video server) and the decoding (at the STB) of the RS packet codeword.

Finally, we mention that the achieved diversity gain depends neither on the packet size L nor on the packet transmission rate R_{pack} , but solely on the parameters $T_{\text{lat}}/T_{\text{coh}}$ and (for RS erasure coding) ovh or (for SR ARQ) $T_{\text{retr,min}}/T_{\text{coh}}$.

5. NUMERICAL RESULTS

5.1. General numerical results

Assuming that a packet consists of $L = 10^4$ bits and a packet group contains $N_{\text{coh}} = 5$ packets, we have displayed in Figures 7–11 several quantities as a function of E_b/N_0 , for SISO ($D = 1$) and Alamouti with 1 or 2 receive antennas ($D = 2$ or $D = 4$). The presented curves confirm the high E_b/N_0 behavior that we established in Section 4, and illustrate the impact of the PHY layer diversity D on the performance.

- (i) Figure 7 shows the probability P_{pack} from (23) that a packet is erased after transmission over the wireless link. We observe that $P_{\text{pack}} \propto (E_b/N_0)^{-D}$ at high E_b/N_0 .
- (ii) The average number of erased packets in a packet group, conditioned on the event that at least 1 packet from the group has been erased, is shown in Figure 8. Note that even at large E_b/N_0 , packet erasures tend to occur in bursts: as the channel state is constant over the channel coherence time, a small value of the channel state (deep fade) is likely to give rise to multiple erasures within a packet group.

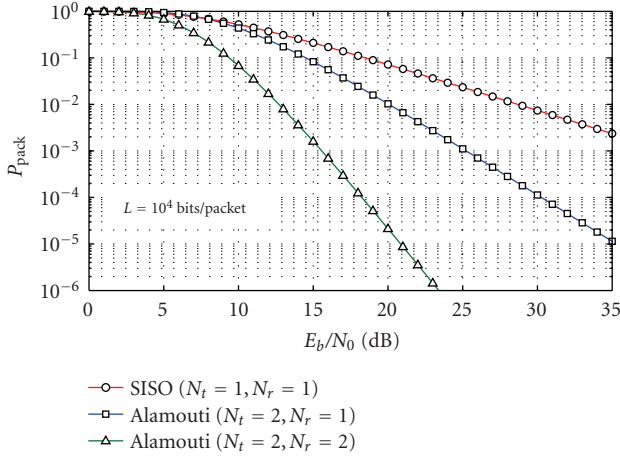


FIGURE 7: Probability P_{pack} that a packet is erased.

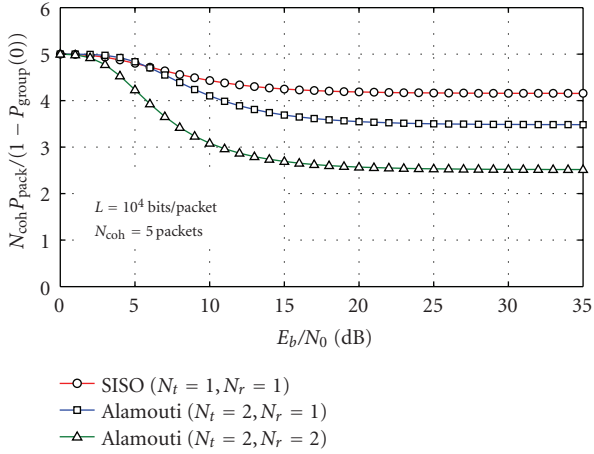


FIGURE 8: Average number of erased packets in a packet group, conditioned on the event that at least one packet in the packet group is erased.

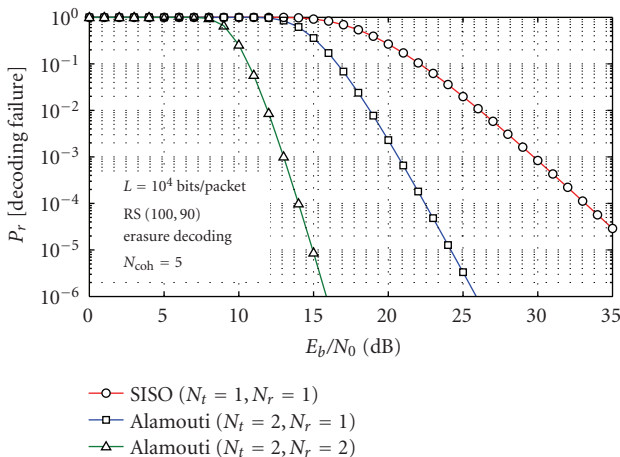


FIGURE 9: Probability of a decoding failure.

- (iii) Figure 9 shows $\Pr(\text{decoding failure})$ (see (13)), for $N = 100$ and $N - K = 10$. As a decoding failure occurs when at least 11 packets in the codeword are erased, a minimum of 3 packet groups is involved in a decoding failure. Hence, according to Section 4, $\Pr[\text{decoding failure}] \propto (E_b/N_0)^{-3D}$ at high E_b/N_0 , which is confirmed by Figure 9.
- (iv) Figure 10 shows the average transmission overhead $E[\text{ovh}]$ from (24), that results from SR ARQ with a maximum of 3 retransmissions. Comparison with Figure 7 reveals that $E[\text{ovh}] \propto P_{\text{pack}}$ at high E_b/N_0 , which confirms our results from Section 4. At small E_b/N_0 , $E[\text{ovh}]$ converges to $N_r = 3$, which corresponds to the case where each packet is retransmitted N_r times.
- (v) Figure 11 shows the probability $P_{\text{group, unrec}}$ (see (19)) that at least one packet from a packet group is definitively lost after 3 retransmissions. Note that $P_{\text{group, unrec}} \propto (E_b/N_0)^{-4D}$ at high E_b/N_0 .

5.2. Results applied to HDTV transmission over a 60 GHz indoor wireless link

Now we consider the transmission of compressed HDTV [19] according to the configuration shown in Figure 1. The compressed video bitrate equals 7.5 Mbps. The link between the HG and the STB is a 60 GHz indoor wireless connection; assuming nonline-of-sight (NLOS) conditions, this connection is modeled as a Rayleigh fading channel, with a coherence time $T_{\text{coh}} = 20$ milliseconds (corresponding to slow motion of about 0.4 m/s) [20]. In order to limit the zapping delay, the latency T_{lat} caused by protecting the video packets against erasures should not exceed 150 milliseconds [21]. The HDTV performance target is a maximum of 1 GOP with unrecoverable packets in 12 hours.

When protecting the video packets by means of an RS packet codeword, we consider transmission overheads of 10%, 20%, and 40%.

When using SR ARQ, we consider two distinct scenarios as far as the location of the retransmission buffer is concerned.

- (i) When the retransmission buffer is located at the HG, $T_{\text{retr, min}}$ is limited to about 5 milliseconds. As 5 milliseconds is less than the 20 milliseconds channel coherence time, the transmitter will defer the retransmission of a packet until 20 milliseconds have elapsed since the previous (re)transmission of the considered packet; hence, this yields $T_{\text{retr}} = 20$ milliseconds.
- (ii) In the case of a low-cost HG, the retransmission buffer is not located at the HG but further upstream, at the DSLAM. The resulting $T_{\text{retr, min}}$ is on the order of 45 milliseconds [22, 23], which exceeds the 20 milliseconds channel coherence time. In this case, we have $T_{\text{retr}} = 45$ milliseconds.

Assuming that the average sizes of an I-frame and a P-frame are 6 times and 2 times the average size of a

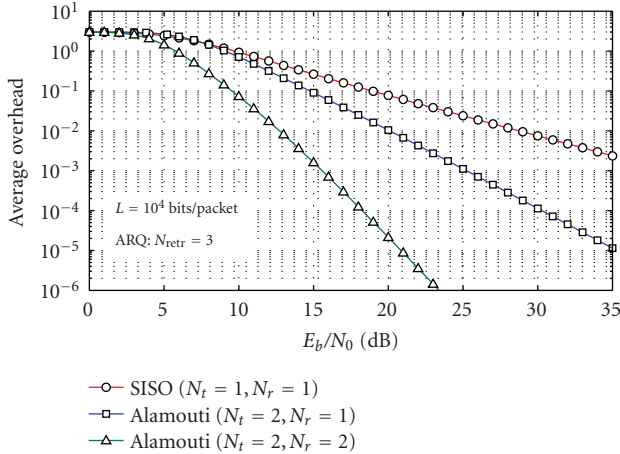


FIGURE 10: Average transmission overhead $E[\text{ovh}]$ from ARQ with maximum 3 retransmissions.

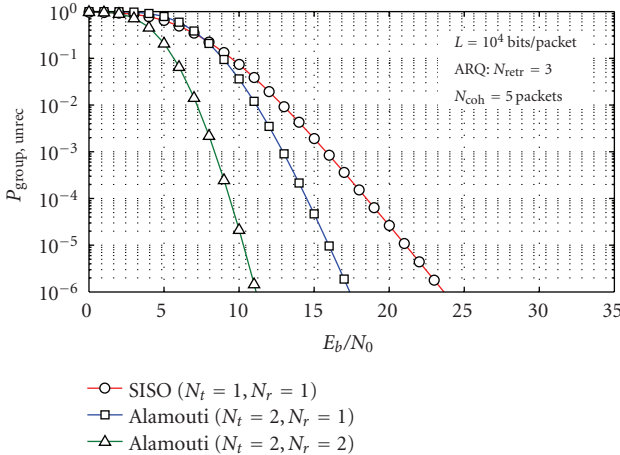


FIGURE 11: Probability $P_{\text{group, unrec}}$ that at least one packet from a packet group is definitively erased (ARQ with maximum 3 retransmissions).

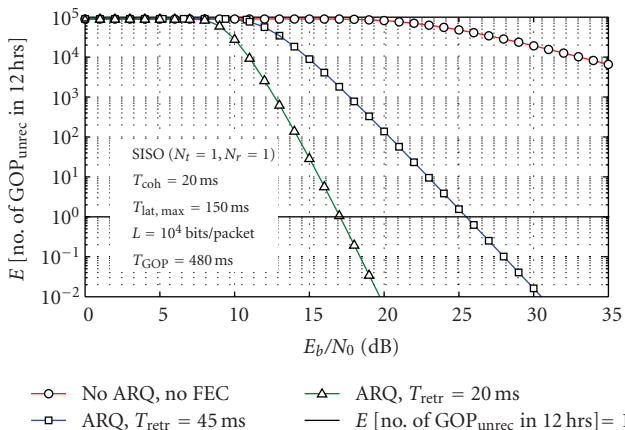


FIGURE 12: Average number of GOPs affected by unrecoverable packet loss in 12 hours (SISO, ARQ).

B-frame, Table 1 shows the average sizes of the different types of frames and of the GOP consisting of the frame sequence IBBPBBPBBPBBP. Note that each type of frame gives rise to multiple IP packets. As the IP packet rate is about 700 packets/s and the channel coherence time is 20 milliseconds, about 14 IP packets fit within the channel coherence time (assuming that IP packets are transmitted at constant regular intervals). Taking into account the propagation of errors from an I- or P-frame to other frames in the GOP, unrecoverable packet loss in an I- or P-frame is very likely to give rise to a visual distortion. Considering that I- and P-frames in a GOP constitute on average 60% of the IP video packets, and packet losses tend to occur in bursts with sizes comparable to the channel coherence time (14 IP packets in our scenario), it follows that when a GOP is affected by an unrecoverable packet loss, the probability that the packet losses occur in I- or P-frame is about 60%. Assuming that packet losses in B-frames are unnoticed but losses in I- or P-frames yield visible distortions, the probability that a GOP affected by unrecoverable packet loss yields a visual distortion is about 60%. (In [20], an experiment is reported which indicates that there is a probability of about 20% that a lost packet yields a visual distortion. However, in [20] the packet losses do not occur in bursts. In the case of bursty packet losses, the probability that a burst of packet losses yields a visual distortion is expected to be larger than 20%.) Moreover, some of the IP packets contain other information (audio, data) related to the HDTV program, that is multiplexed with the video information. The loss of packets containing a multiplex of B-frame information and other HDTV-related information reduces the QoE (because of audible clicks), although the errors in the B-frame do not propagate and could be concealed. Therefore, the average number of GOPs that is affected by unrecoverable packet loss in 12 hours is a meaningful indicator of the QoE.

When conducting the performance analysis, we assumed that the erasure probability on the DSL link is negligibly small as compared to that on the wireless link between the HG and the STB.

Figures 12–18 show the average number of GOPs with unrecoverable packet loss in 12 hours as a function of E_b/N_0 , for the different combinations of PHY layer strategies (SISO and Alamouti with 1 or 2 receive antennas) and packet protection strategies (SR ARQ, RS erasure coding, none). When using SR ARQ, the cases $T_{\text{retr}} = 45$ milliseconds and $T_{\text{retr}} = 20$ milliseconds correspond to diversity gains γ_{ARQ} of 4 (max. 3 retransmission) and 8 (max. 7 retransmissions), respectively. In the case of RS erasure coding, overheads of 10%, 20%, and 40% yield diversity gains γ_{RS} of 1 (i.e., no diversity gain), 2, and 3, respectively. Considering as a performance figure the value of E_b/N_0 that corresponds to $E(\text{no. of GOP}_{\text{unrec}} \text{ in 12 hours}) = 1$, Table 2 collects the performance figure for the different cases. The following observations can be made.

- (i) The highest possible diversity gain is $\lceil T_{\text{lat}}/T_{\text{coh}} \rceil = 8$. This diversity gain is achieved for SR ARQ with $T_{\text{retr}} = T_{\text{coh}}$, that is, when the retransmission buffer is at the HG.

TABLE 1: Average sizes of I-frame, P-frame, B-frame, and GOP.

GOP = {IBBPBBPBBPBB}, 25 frames/s, 7.5 Mbit/s video bitrate			
	size (kbit)	# MPEG-2 TS packets	# IP packets
one I-frame	1080	714	102
one P-frame	360	238	34
one B-frame	180	119	17
one GOP	3600	2380	340

TABLE 2: Value of E_b/N_0 yielding 1 GOP with unrecoverable packet loss per 12 hours.

	$E_b/N_0 @ E[\#GOP_{\text{unrec}} \text{ in 12 hrs}] = 1$					
	no RS, no ARQ	ARQ		RS erasure decoding		
		$T_{\text{del}} = 20 \text{ ms}$	$T_{\text{del}} = 45 \text{ ms}$	ovh = 10%	ovh = 20%	ovh = 40%
SISO	73 dB	17 dB	25.5 dB	71 dB	43 dB	31 dB
Alamouti, $N_r = 1$	43 dB	14 dB	18.5 dB	41 dB	27.5 dB	20.5 dB
Alamouti, $N_r = 2$	25.5 dB	9 dB	12 dB	23.5 dB	16.5 dB	12.5 dB

(ii) Because of their larger diversity gain, the systems with SR ARQ outperform the systems with RS coding. In order to achieve a diversity gain of 4, the transmission overhead of systems with RS coding should be increased to about 70%. A diversity gain of 2 is obtained for the systems with SR ARQ when T_{retr} is between 50 milliseconds and 75 milliseconds.

(iii) Figure 18 compares RS coding and SR ARQ in terms of $E(\text{no. of } GOP_{\text{unrec}} \text{ in 12 hrs})$ for Alamouti with 1 receive antenna, where the system parameters have been selected such that RS coding and SR ARQ yield the same diversity (see Table 3). We observe that the RS code performs worse than SR ARQ. This is because for the RS code the number of dominant erasure patterns yielding irrecoverable packet loss is larger than for SR ARQ.

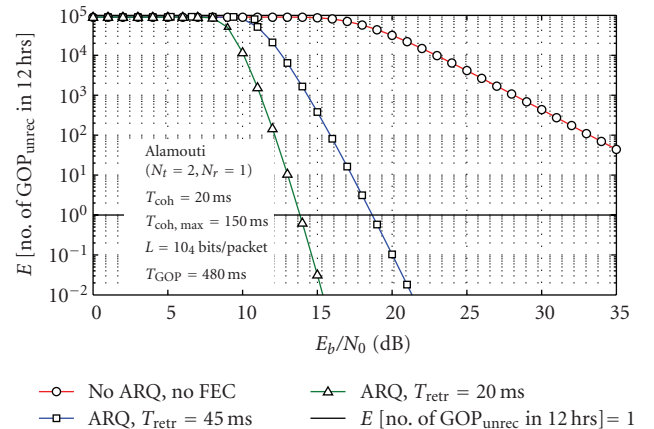
(iv) The performance of the SISO system without any packet protection is very poor. The performance is improved by space-time coding on the PHY layer (which increases the PHY layer diversity D) and/or packet protection by means of RS coding or SR ARQ (which provides additional diversity gain). To some extent, less packet protection can be compensated by using more receive antennas, and vice versa.

6. CONCLUSIONS AND REMARKS

In this paper, we have considered a generic system for video transmission over a wireless link, with space-time coding on the PHY layer and additional video packet protection by means of SR ARQ or RS erasure coding. We have pointed out that SR ARQ and RS erasure coding give rise to a diversity gain yielding improved error performance, and have presented simple analytical expressions for this gain. For both SR ARQ and RS erasure coding, the maximum possible diversity gain equals $\lceil T_{\text{lat}}/T_{\text{coh}} \rceil$. However, when

TABLE 3

	RS	SR ARQ
$\gamma_{\text{RS}} = \gamma_{\text{ARQ}} = 2$	ovh = 20%	$N_{\text{retr}} = 1$
$\gamma_{\text{RS}} = \gamma_{\text{ARQ}} = 3$	ovh = 40%	$N_{\text{retr}} = 2$

FIGURE 13: Average number of GOPs affected by unrecoverable packet loss in 12 hours (Alamouti, $N_r = 1$, ARQ).

using RS erasure coding this maximum diversity gain cannot be achieved because of practical limitations on the allowed transmission overhead. SR ARQ yields the maximum diversity gain provided that $T_{\text{retr}, \text{min}} < T_{\text{coh}}$; otherwise, the actual diversity gain is less. Our theoretical findings have been illustrated in a case study involving HDTV transmission over a 60 GHz indoor wireless link.

The RS erasure coding gives rise to a fixed overhead and latency that are determined by the parameters of the RS code. In the case of SR ARQ, the instantaneous overhead and latency are random; their maximum values are determined

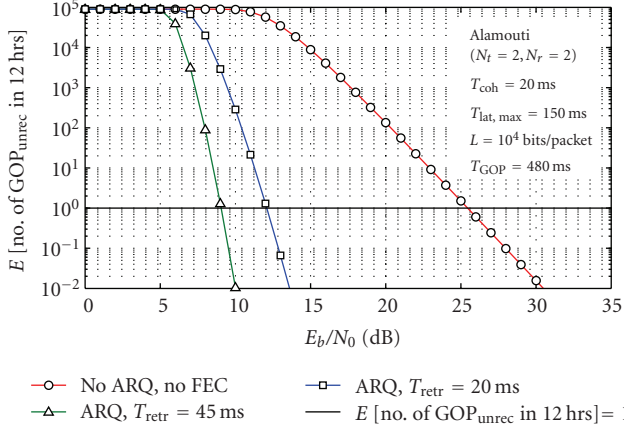


FIGURE 14: Average number of GOPs affected by unrecoverable packet loss in 12 hours (Alamouti, $N_r = 2$, ARQ).

by the maximum number of retransmissions, while their averages decrease with increasing E_b/N_0 are considerably less than the corresponding values for RS erasure coding.

The application of RS erasure coding does not require any modifications of the functionality of the intermediate network nodes, as the construction and the decoding of the RS packet codewords are carried out by the video server and the end user, respectively. Application of SR ARQ involves increasing the functionality (and cost) of the network node where the retransmission buffer is located. From an error performance point of view, the HG should be selected as the retransmitting node, as it provides the smallest round-trip delay and, hence, the largest diversity gain; however, in order to keep the HG a low-cost consumer product, the DSLAM can be selected as the retransmitting node, with the penalty of a larger round-trip delay and a smaller resulting diversity gain. Further, application of ARQ requires the presence of a return channel.

Our performance analysis assumes that the channel state is the same for all OFDM subcarriers. This assumption is valid when the signal bandwidth (R_s) does not exceed the 90% coherence bandwidth of the channel. For the 60 GHz indoor radio channel under NLOS conditions, the 90% coherence bandwidth is about 6 MHz [24], so that our analysis is valid for bitrates up to 12 Mbps (assuming QPSK transmission). When the signal bandwidth is larger than the 90% coherence bandwidth, different subcarriers experience different channel states (which could be exploited to increase the PHY layer diversity by means of frequency-interleaving and coding across the subcarriers of an OFDM block). The detailed analysis of this case is beyond the scope of this paper, but we have been able to verify that the diversity gains γ_{RS} and γ_{ARQ} from Section 4 still apply, so that the main conclusions from this paper remain valid.

WLANs often make use of stop-and-wait (S&W) ARQ on the MAC layer. This form of ARQ has not been included in our performance analysis. We briefly explain how the presence of S&W ARQ on the MAC layer affects the performance. Denoting by $N_{\text{retr, S\&W}}$, and $T_{\text{retr, S\&W}}$ the maximum

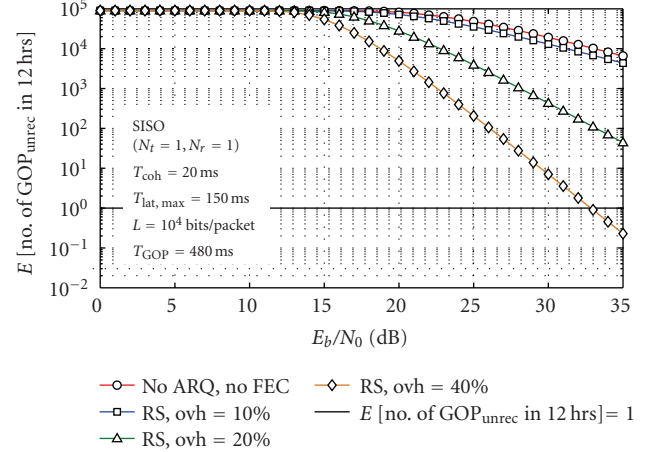


FIGURE 15: Average number of GOPs affected by unrecoverable packet loss in 12 hours (SISO, RS).

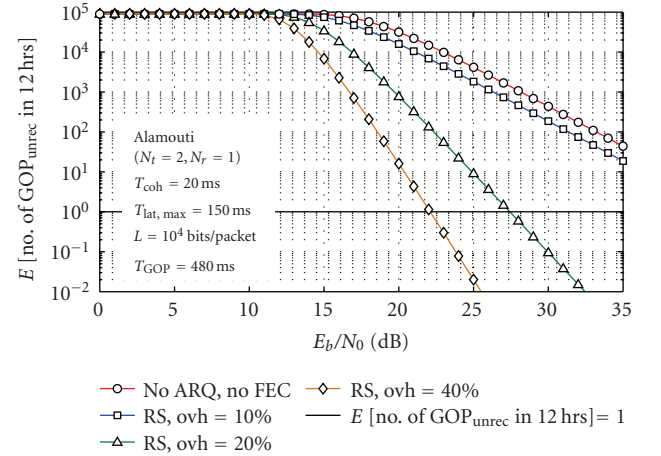


FIGURE 16: Average number of GOPs affected by unrecoverable packet loss in 12 hours (Alamouti, $N_r = 1$, RS).

number of retransmissions and the time interval between (re)transmissions of a same packet, S&W ARQ introduces a maximum latency of $T_{\text{lat, S\&W}} = N_{\text{retr, S\&W}} T_{\text{retr, S\&W}}$. When combined with RS erasure coding, the resulting maximum latency equals $T_{\text{lat}} = T_{\text{lat, S\&W}} + K/R_{\text{pack}}$. When combined with SR ARQ, the resulting maximum latency equals $T_{\text{lat}} = N_{\text{retr, SR}} T_{\text{retr, SR}} + T_{\text{lat, S\&W}}$ with $N_{\text{retr, SR}}$ and $T_{\text{retr, SR}}$ denoting the maximum number of retransmissions and the time between (re)transmissions of the same packet for the SR ARQ protocol; because of the restriction $T_{\text{retr, SR}} > T_{\text{lat, S\&W}}$, we get $T_{\text{lat}} > (N_{\text{retr, SR}} + 1)T_{\text{lat, S\&W}}$. The resulting diversity order is given by $\gamma_{\text{S\&W}} \gamma_{\text{RS}} D$ (RS erasure coding) or $\gamma_{\text{S\&W}} \gamma_{\text{SR}} D$ (SR ARQ), where $\gamma_{\text{RS}} = \lceil (N - K + 1)/N_{\text{coh}} \rceil$, $\gamma_{\text{RS}} = 1 + N_{\text{retr, SR}}$, and $\gamma_{\text{S\&W}}$ is the diversity gain resulting from the S&W ARQ protocol on the MAC layer. As the diversity order does not increase when retransmitted packets experience the same channel state as the packet originally transmitted, the diversity gain from S&W ARQ is evaluated as $\gamma_{\text{S\&W}} = \lceil T_{\text{lat, S\&W}}/T_{\text{coh}} \rceil$.

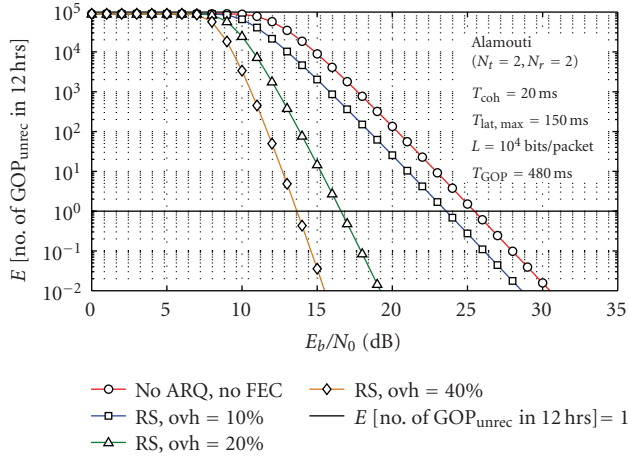


FIGURE 17: Average number of GOPs affected by unrecoverable packet loss in 12 hours (Alamouti, $N_r = 2$, RS).

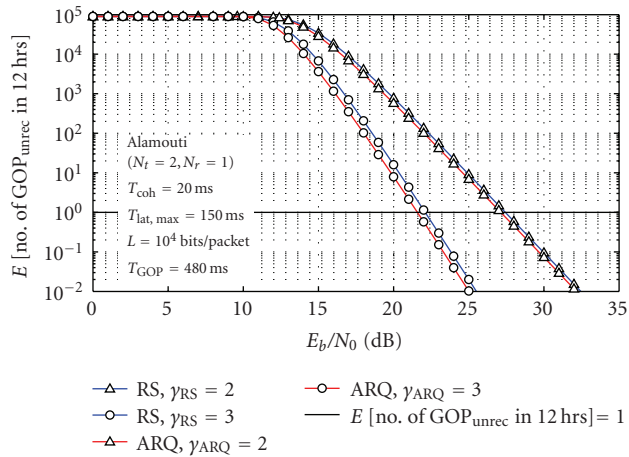


FIGURE 18: Average number of GOPs affected by unrecoverable packet loss in 12 hours, RS versus ARQ (Alamouti, $N_r = 1$).

ACKNOWLEDGMENTS

This work was supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMMUNICATIONS NEWCOM++ (Contract no. 216715). The second author is a Postdoctoral Fellow with the Fund for Scientific Research, Flanders (FWO-Vlaanderen), Belgium.

REFERENCES

- [1] J. A. C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, 1990.
- [2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [3] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*, Cambridge University Press, Cambridge, UK, 2007.
- [4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [5] G. C. Clark Jr. and J. B. Cain, *Error-Correction Coding for Digital Communications*, Springer, New York, NY, USA, 1981.
- [6] J. G. Proakis, *Digital Communications*, McGraw Hill, New York, NY, USA, 2000.
- [7] H. O. Burton and D. D. Sullivan, "Errors and error control," *Proceedings of the IEEE*, vol. 60, no. 11, pp. 1293–1301, 1972.
- [8] R. A. Comroe and D. J. Costello Jr., "ARQ schemes for data transmission in mobile radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 4, pp. 472–481, 1984.
- [9] M. Luby, L. Vicisano, J. Gemell, L. Rizzo, M. Handley, and J. Crowcroft, "The use of forward error correction (FEC) in reliable multicast," IETF RFC 3453, December 2002.
- [10] M. Luby, M. Watson, T. Gasiba, T. Stockhammer, and W. Xu, "Raptor codes for reliable download delivery in wireless broadcast systems," in *Proceedings of the 3rd IEEE Consumer Communications and Networking Conference (CCNC '06)*, vol. 1, pp. 192–197, Las Vegas, Nev, USA, January 2006.
- [11] J. Rosenberg and H. Schulzrinne, "An RTP payload format for generic forward error correction," IETF RFC 2733, December 1999.
- [12] F. Vanhaverbeke, F. Simoens, M. Moeneclaey, and D. De Vleeschauwer, "Binary erasure codes for packet transmission subject to correlated erasures," in *Proceedings of the 7th Pacific Rim Conference on Multimedia (PCM '06)*, pp. 48–55, Hangzhou, China, November 2006.
- [13] S. B. Wicker and V. K. Bhargava, *Reed-Solomon Codes and Their Applications*, IEEE Press, New York, NY, USA, 1994.
- [14] S. R. Ely and C. Eng, "MPEG video coding: a basic tutorial introduction," Research and Development Report BBC RD 1996/3, British Broadcasting Corporation, London, UK, 1996, <http://downloads.bbc.co.uk/rd/pubs/reports/1996-03.pdf>.
- [15] P. A. Sarginson, "MPEG-2: overview of the systems layer," Research and Development Report BBC RD 1996/2, British Broadcasting Corporation, London, UK, 1996, <http://downloads.bbc.co.uk/rd/pubs/reports/1996-02.pdf>.
- [16] Overview of digital subscriber line (DSL) recommendations, ITU-R recommendation G.995.1, February 2001.
- [17] D. Hoffman, G. Fernando, V. Goyal, and M. Civanlar, "RTP payload format for MPEG1/MPEG2 video," IETF RFC 2250, January 1998.
- [18] N. Degrande, K. Laevens, D. De Vleeschauwer, and R. Sharpe, "Increasing the user perceived quality for IPTV services," *IEEE Communications Magazine*, vol. 46, no. 2, pp. 94–100, 2008.
- [19] "Parameter values for the HDTV+ standards for production and international programme ex-change," ITU-R recommendation BT.709-5, 2002.
- [20] S. Kanumuri, P. C. Cosman, A. R. Reibman, and V. A. Vaishampayan, "Modeling packet-loss visibility in MPEG-2 video," *IEEE Transactions on Multimedia*, vol. 8, no. 2, pp. 341–355, 2006.
- [21] M. Watson, "Proposal for evaluation process for forward error correction codes for DVB-IPI," DVB IPI document TM-IPI2084 edition, September 2005.
- [22] A. Gurtov and S. Floyd, "Modelling wireless links for transport protocols," *ACM Computer Communications Review*, vol. 34, no. 2, pp. 85–96, 2004.

-
- [23] C. Hoene, A. Gunther, and A. Wolisz, "Measuring the impact of slow user motion on packet loss and delay over IEEE 802.11b wireless links," in *Proceedings of the 28th Annual IEEE International Conference on Local Computer Networks (LCN '03)*, pp. 652–662, Bonn, Germany, October 2003.
- [24] H. Yang, P. F. M. Smulders, and M. H. A. J. Herben, "Channel characteristics and transmission performance for various channel configurations at 60 GHz," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, Article ID 19613, 15 pages, 2007.