KTH Electrical Engineering

Internet Video Transmission

GYÖRGY DÁN

Doctoral Thesis
Stockholm, Sweden, 2006

TRITA-EE 2006:001
ISSN 1653-5146
ISRN KTH/EE–06/001–SE
ISBN 91-7178-256-7

Department of Signals, Sensors and Systems
KTH, Stockholm, Sweden

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen tisdagen den 28 februari 2006 i Sal F3, KTH, Stockholm.

Tryck: Universitetsservice US AB

Abstract

The Internet has rapidly evolved from being a scientific experiment to a commercial network connecting millions of hosts that carries traffic generated by a large amount of applications with diverse requirements. Its architecture was however designed to enable efficient point-to-point delivery of bulk data, and can not provide statistical guarantees on the timely delivery of delay sensitive data such as streaming and real-time multimedia. Thus, applications that require low loss probabilities in today's Internet have to use some end-to-end error recovery mechanism. For delay sensitive applications the introduced latency by the applied schemes has to be low as well. Traffic control functions such as delay limited shaping and forward error correction (FEC), and multiple description coding (MDC) have been proposed for variable bitrate video. Their major drawback is, however, that it is difficult to predict their efficiency, as it depends on many factors like the characteristics of the stream itself, the characteristics of the traffic in the network and the network parameters. Consequently, it is difficult to decide which control mechanisms to employ, how to combine them and to choose the right parameters (e.g. block length, code rate) for optimal performance.

In this thesis we present results on the efficiency of traffic control functions and MDC for video transmission based on mathematical models and simulations. We investigate the efficiency of delay limited traffic shaping and the trade-offs in the joint use of traffic shaping and forward error correction. We identify the packet size distribution of the traffic in the network as an additional factor that may influence the efficiency of FEC, and present a thorough analysis of its possible effects. We present an analytical comparison of MDC versus media-dependent FEC and media-independent FEC, and based on the results we conclude that MDC is a promising error control solution for multimedia communications with very strict delay bounds in an environment with bursty losses. We combine the analytical results with traces from measurements performed on the Internet to evaluate how efficient these error control schemes are under real loss patterns. We compare the efficiency of MDC and media-dependent FEC in the presence of channel estimation errors; we propose a new rate allocation method, which is robust to mis-estimations of the channel state and which improves error resilience on non-stationary channels. Finally we present an analytical model of the performance of an end-point-based multimedia streaming architecture based on multiple distribution trees and forward error correction, and analyze the behavior of the architecture for a large number of nodes.

Acknowledgments

First I would like to thank my advisor Viktória Fodor for her support, for the fruitful conversations and sometimes hard critics, and for introducing me into the world of scientific research. I would also like to thank my second advisor professor Gunnar Karlsson for his helpfulness and his comments on my work. Furthermore I would like to thank all the people at LCN for the nice atmosphere, which made it possible for me to focus my efforts on research.

Additionally I would like to express my gratitude to the organizations and entities that have founded my research, the Affordable Wireless Services and Infrastructure project and the Graduate School of Telecommunications.

I would like to express my special thanks to all my friends in Stockholm who gave me support whenever I needed it, and helped me to spend four wonderful years in Sweden. I would also like to thank to my friends back home and abroad who, though being far away, were always there to help me and made my holidays an infinite source of energy. I am happy to thank Gabriela for her patience and for keeping me in countenance during the ups and downs of the thesis work. Last but not least, I would like to thank to my family for their understanding, and their continuous support.

Contents

Chapter 1

Introduction

Only a couple of decades after the first ever data packet was sent between two computers in a packet switched network, the Internet has evolved to become a network connecting hundreds of millions of nodes that is serving a large amount of diverse applications having different requirements. Originally, the network was meant for machine-to-machine data transfer, but it was soon being used as a medium for personal communications, such as sending e-mails. With the appearance of different multimedia encoding standards the network could also be used for transmission of audiovisual information. Later, as bandwidth and the user base of the Internet grew, the demand for machine-to-human and human-to-human communications appeared. But, while e-mail could offer the reliability of the postal mail service without major architectural implications on the network, the Internet has no mechanisms to support the quality requirements of audiovisual communications that one is used to in the Public Switched Telephone Network (PSTN) and the Integrated Services Digital Network (ISDN).

Background

The Internet is now considered as the universal network for data, voice and video communications. It is recognized that the best effort service provided today is not satisfactory for delay and loss sensitive applications such as live voice and video. It is generally accepted that traffic control functions must be employed to guarantee these applications adequate quality at reasonably high network loads. The introduction of new control functions in the Internet is, however, a critical issue. First, a very high number of networking devices have to be updated or replaced. Second, the complexity of the control functions may limit the scalability and the transmission capacity of the network. Third, it could be in contradiction with one of the principal ideas used in the design of the Internet, the end-to-end argument [1]. Fourth, it might require new business models for the operators who need to charge extra for traffic with quality of service.

Proposed mechanisms like Intserv and Diffserv that aimed at providing some form of quality of service by introducing new functionality in the core of the network have not been deployed globally. To avoid the problem of updating the existing infrastructure, alternative admission control solutions have been proposed based on per–hop or end–to–end measurements with very little or no support in the routers. Measurement-based admission control (MBAC) schemes base the acceptance decision on per-hop real–time measurements of the aggregate traffic intensities [2]. Endpoint measurement-based admission control (EMBAC) schemes decrease the required router support further by involving only the end systems in the admission control process [3, 4, 5, 6]. The idea behind these schemes is to probe the transmission path from the sender to the receiver to estimate the load level in the network and admit new streams only if the load is acceptable. One solution that employed EMBAC is the architecture proposed in [5, 7]. In its controlled load service class, packet scale buffering is used to ensure low network delays and probe based admission control [8] is used to limit the packet loss probability. Packet scale buffering combined with source shaping at the end-nodes was studied in detail in [9, 10, 11].

Though much research and standardization has been done in recent years, there is no support yet from the network side for QoS in form of resource reservation and call admission control. Thus applications that require low packet loss and delay jitter have to employ some end-to-end mechanisms that compensate for the disturbances introduced by the network. There is a wide variety of such mechanisms, traffic control functions, suitable for real-time and streaming multimedia communications, both for point-to-point and multicast.

Delay jitter is often compensated for on the receiver side via adaptive playout algorithms. A good survey of such algorithms can be found in [12]. Packet loss can be compensated for on the sender side and the receiver side. On the receiver side receiver-based error concealment algorithms can be used like insertion and interpolation [13, 14]. On the sender side delay-limited shaping can be used to reduce the burstiness of the traffic before injecting it into the network. At the same time redundant information can be added to the data flow and error resilient source coding can be used. The redundant information can be used to reconstruct lost information. To achieve the best possible perceived quality the ratio of redundancy has to be selected according to the expected network conditions, which leads to the problem of source and redundancy rate allocation.

The set of suitable control functions depends on the type of application, as different applications have different requirements for loss and end-to-end delay. Human-to-human, also called real-time and interactive, communication has the strictest requirements: for toll quality the one way delay should be below 150 ms according to the ITU's G.114 recommendation. Machine-to-human, called streaming, communication is usually more tolerant to delay. Streaming applications can tolerate delays of up to several seconds, unless quick seeking in the streaming content is necessary. The set of available control functions also depends on the number of parties involved. In the case of point-to-point communications feedback might be used to combat errors. In the case of point-to-multipoint communications, called multicast, feedback information from a large number of clients can overload the sender, unless some distributed protocol is used to aggregate feedback information.

Given the set of available mechanisms, one has to decide which mechanisms to use, how to combine them and how to set their parameters for optimal performance based on the estimated channel state. The task is particularly difficult, as the efficiency of these mechanisms depends on many factors, such as the characteristics of the traffic to be inserted in the network and the network state. The work presented in this thesis addresses these issues. We present results that show that traffic shaping can be efficient even with low delays suitable for real-time communication. We investigate the trade-off between traffic shaping and forward error correction (FEC), and show that under certain circumstances combining the two traffic control solutions can improve the quality of the transmission. We identify an additional factor that can influence the efficiency of FEC and present a thorough evaluation of its possible effects. We present a comparison of two solutions for error control, FEC and multiple description coding (MDC), and conclude that MDC is a promising solution for real-time communications on the Internet. We propose a new method for rate allocation that is robust to mis-estimations of the channel state and performs well on non-stationary channels. Finally we investigate the use of FEC to provide robustness to end-point-based multicast streaming.

Chapter 2

Video coding

Digital video consists of a sequence of video frames that are recorded and played back at a given rate, the frame rate. The goal of video coding is to encode the sequence of frames into a stream of data. Frames are usually divided into smaller units, called slices. Slices are in turn groups of macroblocks, which are the basic blocks of the coding process, rectangular areas consisting of typically up to 16 by 16 pixels. There are a large number of different coding standards available today developed mainly by the ITU and the Motion Pictures Expert Group (MPEG) within the ISO. The first codecs developed by the ITU/IEC, such as the H.261 and H.263 were mainly targeted to low bitrate video communications, such as videoconferencing. Early standards developed by the MPEG, such as the MPEG-1 and MPEG-2 (developed together with the ITU) were developed for high quality video, such as digital video storage and broadcasting. A more recent standard developed by the MPEG, the MPEG-4 was designed to be suitable for both low and high bitrates and offers a large set of coding options. The most recent standard in the line of video coders is the H.264/AVC, which is the result of a collaboration between the MPEG and the ITU, and offers superior quality to previous coding standards. While the set of coding parameters and the algorithms used have become more and more sophisticated, the basic idea underlying these hybrid video coders is the same. They achieve high ratios of compression by decreasing the spatial, temporal and statistical redundancy present in video data. The block diagrams of a motion compensated video coder and decoder are shown in Fig. 2.

Spatial redundancy

Spatial redundancy is reduced by spatial prediction, in which case only the difference between neighboring macroblocks of an image is transmitted. Another key component in the reduction of spatial redundancy is the discrete cosine transform (DCT), and the following quantization, which allows to tune the granularity of the transmitted information. The coarser the quantization the less information will be transmitted by the price of increased distortion. The methods used to reduce spatial redundancy are called intra-frame coding.

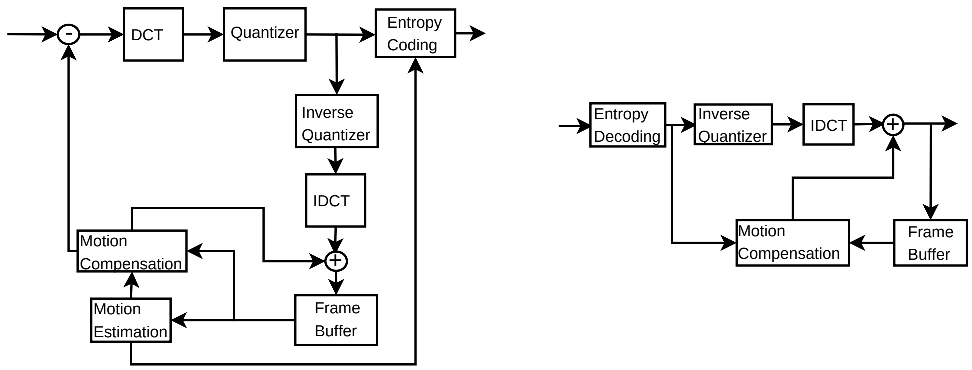


Figure 2.1: Block diagram of a motion compensated video coder and the corresponding decoder.

Temporal redundancy

Temporal redundancy is reduced by temporal prediction. It is usually called motion compensation, and looks for a displacement vector which minimizes the difference between the part of a frame to be coded and its reference frame. Instead of transmitting the frame itself, it is enough to transmit the displacement vector and the difference between the displaced reference frame and the actual frame. If the subsequent images of a video stream are similar, then this technique can reduce the amount of coded information significantly. The methods used to reduce the temporal redundancy are called inter-frame coding. The different standards have diverse options for the choice and the number of reference pictures, with the most recent standard H.264 offering the highest flexibility. Yet, the naming conventions come from the MPEG-1 standard. Frames that are coded using intra-frame coding only, are called I frames, frames that are coded relative to an I frame are called P frames, and frames coded based on two I or P frames are called B frames. Before the H.264 standard whole frames had to be coded using the same inter-frame coding mode, that is, either I, P or B, starting with the H.264 it is possible to select the inter-frame coding mode per macroblock.

Statistical redundancy

Statistical redundancy is reduced by using entropy coding on the intra and inter-coded information. The coding scheme has to provide good compression ratio and at the same time facilitate error detection and concealment. Reversible variable length coding (RVLC) and context adaptive binary arithmetic coding (CABAC) are used in recent video coders for this purpose.

Variable bitrate video

When entire frames are coded using the same inter-frame coding mode, the sizes of the coded frames differ substantially depending on the inter-frame coding mode, unless the quantization parameter is adjusted for each frame individually. I frames are the biggest ones requiring up to an order of magnitude more information for the same distortion as B frames, and P frames have sizes between those of I and B frames. The order of I, P and B frames is usually fixed within a video transmission, and we call frames between two subsequent I frames a group of pictures (GOP). The distinction of different types of frames based on the applied inter-frame coding mode starts to diminish in the H.264.

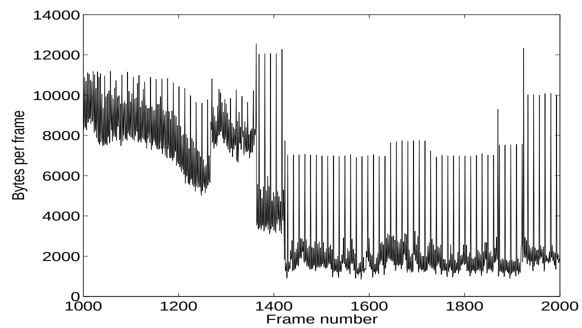


Figure 2.2: Number of bytes per frame for an MPEG-4 coded talk show trace. A scene change can be observed around frame 1400.

The sizes of the coded frames do not only change depending on the coding mode used to encode the individual frames, but also due to the changes in the complexity of the scenes, a sequence of frames having similar video content. Some scenes contain more contextual information, in which case intra-frame coding is less effective. Other scenes can be more motion intensive, in which case inter-frame coding is less effective. Scene changes cause a long term fluctuation of the frame sizes. Figure 2.2 shows the frame sizes of a talk show trace, both short and long term fluctuations of the frame sizes can be observed.

In the case when the quantization parameter is adjusted so that the amount of coded information per GOP is constant, we talk about constant bitrate encoding (CBR), otherwise the encoding is variable bitrate (VBR). CBR video has the advantage that it can easily be transmitted over a channel with fixed capacity, such as the Public Switched Telephone Network (PSTN) or the Integrated Services Digital Network (ISDN), the distortion of the individual frames fluctuates however. Such variation of the distortion usually decreases the perceived quality of the video. VBR video has the advantage of nearly constant quality at the price of the varying size of the coded frames. As the number of streams that can be multiplexed in the network depends on how much the rate of the individual sources fluctuates, it can be beneficial to smooth out the fluctuations of the bitrate by employing traffic shaping. We consider the problem of traffic shaping in Papers A and B in the context of MPEG-4 coded real-time video transmission.

Effects of delay jitter and loss on video

Delay jitter can be compensated for partially on the receiver side by using a playout buffer. Nevertheless, delay jitter higher than the length of the playout buffer results in late delivery of data to the decoder. Data that arrive late are considered to be lost and degrade the quality until they arrive. Once they arrive they might be used for the decoding of subsequent data, e.g. if the delayed data is used as the reference for subsequent data.

In the presence of losses the quality of motion compensated video is degraded due to temporal and spatial error propagation. Temporal error propagation is an effect of the temporal prediction techniques. Macroblocks that are received but were encoded relative to a lost macroblock or damaged macroblock in a previous frame can not be reconstructed correctly. The effects of temporal error propagation can be decreased by sending intra-coded frames more often, which decreases the achievable coding gain. Spatial error propagation is an effect of the spatial prediction techniques. Macroblocks that are received but were encoded relative to a lost or damaged macroblock in the same frame can not be reconstructed correctly. A way to decrease spatial error propagation is to limit spatial prediction to a small subset of macroblocks, this too reduces the achievable coding gain. Figure 2.3 shows how temporal and spatial error propagation aggravate the effects of losses.

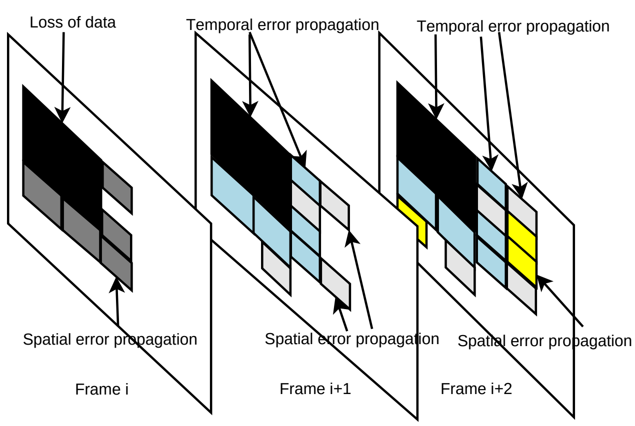


Figure 2.3: Temporal and spatial error propagation aggravate the effects of losses.

Modern coders, like the H.264 support grouping of the coded bitstream according to the importance of the bits for the decoding process. This feature is called data partitioning. The effects of loss are least severe if the least important bits are lost, and they are most severe if the most important bits are lost, in which case the less important bits can not be interpreted. If the sender has the possibility to send the important bits over a channel with lower loss probabilities then the effects of losses can be significantly decreased. A similar approach is layered coding, like MPEG Fine Granularity Scalability (FGS). Layered coding creates several bitstreams, the base layer and the enhancement layers, instead of a single stream of data. For successful decoding the receiver has to receive the base layer plus any number of enhancement layers in increasing order of their importance. The decoding quality is then proportional to the number of layers received. While data partitioning was introduced to leverage prioritized transmission, the goal of layered coding is mainly to support several clients with different bandwidth requirements simultaneously.

Other error resilience features in today's video coders are among others reversible variable length codes, flexible macroblock ordering and adjustable ratio of intra-coded macroblocks [15]. But even with these features it is difficult to achieve good compression and good error resilience at the same time.

2.1 Video traffic modeling

The modeling of video traffic has received much attention in the literature. The proposed models can be divided into two main groups depending on whether they use explicit information (media-aware) about the coder or not (media-unaware).

Media-unaware models are general traffic models applied to variable bitrate traffic. They can be divided into three main groups: regression, self-similar and Markovian models [16, 17, 18].

Regression models attempt to describe the evolution of the traffic intensity based on past values of the traffic intensity and added noise. Simple regression models proposed include discrete time autoregressive [19], discrete autoregressive [20] and autoregressive moving average [21] models. Transform-expand-sample (TES) models are non-linear regression models that aim to fit both the autocorrelation and the marginal distribution of the empirical data. Simple regression models, e.g. autoregressive models, enable a - sometimes approximate - queuing analysis to be performed based on the model as presented in [19, 20]. The complexity of the queuing models resulting from other models does not make it possible however to use them for analytical performance evaluation.

VBR video is known to exhibit long-range dependence (LRD)[22]. As LRD traffic is known to influence the multiplexing behavior a variety of models showing LRD behavior has been proposed. Models belonging to this group are based among others on fractional-ARIMA [22, 23, 24] and shifting level processes [25]. Self-similar models of video traffic give a very good match of the autocorrelation function for both short and long lags, but it is difficult to use them to evaluate the queuing behavior.

Markovian models are widely used despite of the fact that they cannot capture the LRD behavior of video traffic. Although these models fail to match the autocorrelation for long lags, they can be used to derive good approximate results for the queuing performance in the presence of short buffers as they capture the dominant short term correlations [26]. Markovian models, e.g. Markov modulated Poisson processes (MMPP), have been shown to give accurate estimates of the queuing performance of LRD traffic and second-order self similar processes [27, 28, 29]. In this thesis we use a MMPP to model variable bitrate video traffic in Papers C, D and E due to the analytical tractability of the resulting queuing models.

Media-aware models use the information available about the way intra and inter-coded frames follow each other in the coded stream. The model presented in [30] uses the a priori information on the order of I, P and B frames. A similar approach is followed in [31]. The authors in [32] model video traffic by modeling the scene, frame and slice level behavior with separate models. Though such models can be accurate, their use for analytical evaluation is problematic, and thus they are mainly used for simulations.

2.2 Video quality assessment

The goal of video coding is to deliver the best possible quality visual information to the viewer given some constraints on storage capacity or network resources (e.g. capacity). This objective can be reached by maximizing the perceived quality of the decoded video. There are however two problems with this optimization.

The first problem is that even though there has been much research on the human visual system, there is still no generally accepted quantitative model of perceived visual quality. This problem affects both stored (e.g. DVD) and networked (e.g. streaming) video coding. Asking a fairly large number of persons to watch the video sequence is the only way as of today to get an accurate estimate of video quality. The average of the panel of observers' opinions is called the mean opinion score (MOS).

There are however several ways to calculate the approximate fidelity of coded video. The most widely used, and most simple measure of video quality is the mean square error (MSE) and the peak signal to noise ratio (PSNR). The mean square error is the mean of the squared differences between pixels of the original and the reconstructed video. The PSNR is calculated based on the mean square error and the maximum value of the individual pixels ($\hat{p}$, e.g. $\hat{p} = 255$) as $PSNR = 10log_{10}\frac{\hat{p}^2}{MSE}$. The MSE and the PSNR are not a precise measure of perceived video quality, even though they show a good correlation with the MOS [33].

More accurate objective measures of visual quality can be split into two groups. Models belonging to the first group fit a mathematical model to the measured subjective MOS values [34]. Models belonging to the second group are based on the behavior of the human visual system (HVS). The first models were developed for still images, such as the structural similarity index method (SSIM) [35] or the visible difference predictor (VDP) [36]. A detailed review of quality metrics can be found in [37]. Models for video quality assessment have to incorporate the temporal behavior of the HVS. Examples of such models are the perceptual distortion metric (PDM) in [38], the moving pictures quality metric (MPQM) described in [39] or the Video Distortion Meter (VDM) presented in [40].

The second problem is that it is impossible to compare the received, possibly error prone video information to the original video at the encoder. The comparison would allow the encoder to use the right amount of error resilience based on the measured distortion. This problem affects networked video only. No-reference quality metrics that can measure the distortion of video information without the reference are even less developed than quality metrics using reference information. Thus, instead of using no-reference metrics in the decoder, the encoder has to maintain a virtual decoder, which simulates the decoding process of the error prone video stream based on the measured channel state. The coder has to choose the level of error resilience based on the estimated channel parameters and a chosen channel model. Paper F deals with the problem of setting the optimal level of error resilience in presence of channel estimation errors and on a non-stationary channel.

Analytical modeling of video quality

There is a wide range of algorithms that measure the objective video quality based on traces. Models that would allow analytical evaluation of the video quality are however restricted to the mean distortion.

The simplest model is to assume a Gaussian source and use results from distortion-rate theory to calculate the mean distortion bound for a given source rate and with respect to a specific distortion metric. The mean squared error distortion metric is a reasonable choice for video, as it is widely used in video coding. The main advantage of this approach is that the resulting models are analytically tractable. A clear disadvantage is that it is difficult to predict how well results obtained using this approach approximate the behavior of coded video.

A model to calculate the mean distortion of H.263 coded video in the presence of independent packet losses was presented in [41]. The model takes the bitrate, the ratio of intra-coded macroblocks and the packet loss probability as input parameters. The rest of the parameters have to be calculated based on a set of measurements performed with the video trace one would like to model. The assumption of independent losses has been released in a similar model presented in [42]. A measurement based approach was followed in [43], where a simple function was fitted to the measured objective quality of MPEG-2 coded video. Even though these models are rather accurate in predicting the video quality for given network conditions, they have several video trace dependent parameters. Furthermore, not even these models can predict how the distortion behaves when error control is used to recover from errors.

In papers E and F we use results from distortion-rate theory to approximate the distortion of coded video in the presence of losses and various error control solutions. In paper F we use the analytical model developed in [41] to illustrate and support our results.

Chapter 3

Traffic Shaping

Traffic shaping is a way to smooth the bitrate fluctuations of bursty traffic, such as VBR video. Decreasing the fluctuations of the bitrate increases the number of sources that can be multiplexed in the network at a given packet loss probability. In the case of a guaranteed service, it decreases the bandwidth that has to be allocated for the transmission. Shaping can be done either at the end-nodes, before the traffic is injected in the network, or within the network. In this thesis we consider shaping performed in the end-nodes, as we focus on application based traffic control solutions.

For VBR video traffic shaping can be used to decrease the frame to frame bitrate fluctuations. A shaper will typically spread the transmission of large frames over intervals exceeding one frame time, and speed up the transmission of small frames. Hence, traffic shaping introduces some delay. For applications with strict delay constraints the delay introduced by the shaping algorithm can not exceed a given threshold. Smoothing the fluctuations of the traffic with a given delay constraint is called delay limited shaping. The output of a delay limited shaper is not necessarily a constant bitrate stream, as it would be using a leaky bucket.

In the case of stored video, sufficient information is available a priori to achieve optimal shaping. The optimality can be defined in different ways depending on the assumptions on the network cost model and on the assumption on the client buffer size. Possible criteria are to minimize the number of bandwidth changes, to minimize the number of bandwidth increases, to minimize the peak bandwidth requirements and maximize the largest minimum bandwidth requirements or to minimize the variability of the bandwidth requirements [44, 45, 46]. The shaping algorithm has to achieve its goal while avoiding an underflow or overflow of the client buffer. The result of the shaping algorithm is a transmission plan, which consists of a number of fixed rate runs. The size of the client buffer and the playout delay determine how much time in advance large frames can be prefetched and transmitted, and the number of runs. An example showing two possible transmission plans can be seen in Fig. 3.1.

In the case of live video content, like live streaming and real-time communications, there is a trade-off between the delay introduced at the shaper and the

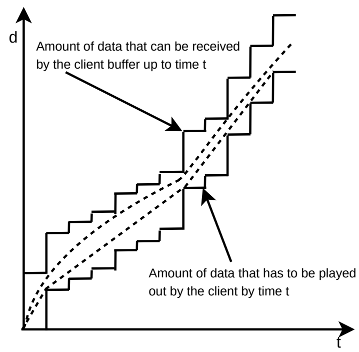


Figure 3.1: Transmission plans for shaping stored video.

effectiveness and complexity of shaping. The higher the delay introduced, the more information is available on upcoming frames, and hence, the closer the shaper can operate to optimality. Consequently, shaping algorithms for live streaming, where delays of several seconds can be tolerated can achieve higher gains than those for interactive applications, where the delay available for shaping is low, in the order of tens of milliseconds.

Shaping algorithms for online traffic usually work by applying existing algorithms for stored video to the set of available frames. The set of available frames is called a window, the higher the number of frames in the window, the smoother the transmission of the video can be. Nevertheless, a longer window increases the delay introduced by shaping, and requires more memory both at the sender and the receiver side. The most straightforward approach to online-shaping is to apply an offline algorithm to non-overlapping windows, such algorithms are called hopping-window based [47]. Sliding-window based algorithms perform the shaping on overlapping windows, and hence can achieve better performance than hopping-window based algorithms [47, 48]. Sliding-window based algorithms are, however, more computationally intensive, as the transmission plan has to be recalculated more often. The performance of online-algorithms can be enhanced by using a priori information about the encoded stream (e.g. frame order, GOP length), which can help to predict the statistics of upcoming frames.

Most of the work done related to traffic shaping focuses on achieving optimal shaping, but less on how much gain can be achieved by shaping. In papers A and B we address this question and show analytical and simulation results using two low-complexity shaping algorithms suitable for the shaping of online traffic.

An application of shaping to call admission control, called bounding interval dependent method, was presented in [49]. In this scheme a piecewise linear approximation of the source's maximum arrival curve, i.e. the maximum amount of data sent by the source over a time interval, is calculated offline for several interval-lengths. Each one of the linear segments corresponds to a leaky bucket. Shaping the source using these leaky buckets the introduced delay is not more than the length of the interval corresponding to the leaky bucket with the lowest rate. The method was adopted to online sources in [50] by renegotiating the traffic parameters whenever needed. If renegotiation is not possible, the source's rate has to be decreased or data has to be discarded.

Chapter 4

Error Control

Error control can be used in error-prone networks to recover from losses. There are two main groups of error control solutions, retransmission protocols and error control coding. These methods are widely used in today's networks in the various layers of the network architecture. The link layer provides error checking to detect errors that occurred during transmission over the physical layer. On wired links the error control implemented in the link layer is typically restricted to error detection and discarding of the erroneous blocks of information. On wireless links error control can include some form of error-correcting codes or retransmission protocols. On the transport layer TCP uses retransmissions to ensure correct delivery of all packets. Application layer error control can include both retransmission protocols and error control coding.

4.1 Feedback based error control

Feedback based error control solutions are mainly used in bulk transfer of data. These solutions use a feedback channel to send information about the reception of data packets (ACK based protocols) or about the loss of data packets (NACK based protocols) to the sender. The sender takes the necessary measures to minimize the effects of the loss. Retransmission based error control mechanisms like Automatic Repeat Request (ARQ) resend the lost packet in the hope that it will not be lost again.

Retransmission based mechanisms can be useful in the case of point-to-point multimedia streaming when used together with a large playout buffer, which compensates for the delay jitter introduced by retransmissions. They are however rarely used in real-time multimedia applications due to the delay introduced by these protocols that is hard to predict. The introduced delay depends on the number of retransmissions required for correct data delivery, and in the case of real-time multimedia, retransmitted packets are likely to miss the playout deadline.

Another form of feedback based error control has however been proposed for low-delay video transmission [51]. The basic idea behind this solution is to use feedback information to stop temporal error propagation. Two proposed ways are to use uncorrupted reference frames as a basis for inter-frame coding (reference picture selection) or to send intra-coded information for the frame segments affected by loss (error tracking). Due to their computational complexity these feedback based solutions are not widely used.

4.2 Forward Error Correction

Forward error correction (FEC) has been proposed to recover from losses in the case of multicast streaming and in real-time applications, where the latency introduced by retransmission schemes is not acceptable. FEC increases the redundancy of the transmitted stream and recovers losses based on the redundant information. There are two main directions of FEC design to recover from packet losses, media-dependent and media-independent FEC.

Media-dependent FEC

In the case of media-dependent FEC (MD-FEC) a redundant copy of the original packet is added to one of the subsequent packets. Figure 4.1 shows the basic concept of MD-FEC. If the packet carrying the primary encoding is lost, data can be reconstructed from the redundant copy, given that it is received. Though the quality obtained from the redundant copy is lower, it is still better than if there is nothing to play out. MD-FEC was primarily proposed for real-time audio and implemented in tools like Rat and FreePhone [52, 53], but has also been included in the H.264 video coding standard [15]. The main advantage of MD-FEC is that it is suitable for applications with low bitrate and low available delay for error control. Several extensions of the original idea have been proposed and evaluated via simulations and measurements, such as sending several redundant copies of the original packet [54], and to increase the spacing between the original and the redundant packets [55]. An analytical evaluation of this scheme has been presented in [56]. It showed that if all connections traversing a bottleneck apply MD-FEC then the quality is worse then if they were not applying it due to the increased load.

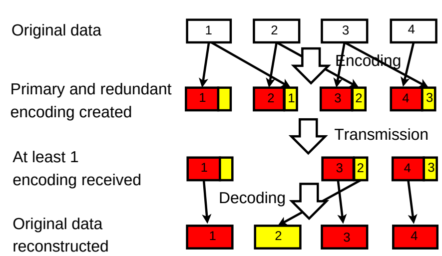


Figure 4.1: Media-dependent FEC scheme. Redundant, low quality descriptions are used to protect data from losses.

Media-independent FEC

Media-independent FEC (MI-FEC) solutions use algebraic coding to create pieces of redundant information that protect the original pieces of information. The pieces of information that algebraic codes operate on are called symbols. A symbol can be one or several bits depending on the code. After transmission on a lossy channel the original symbols can be reconstructed if the ratio or the number of lost symbols is below a certain value. The particular value depends on the parameters and type of the applied coding scheme. The two main types of MI-FEC are based on block codes and convolutional codes.

Block codes

In the case of block based FEC blocks of k symbols are encoded to blocks of n symbols. The code is called systematic, if the encoded block contains the k original symbols; in this case no decoding is needed whenever all symbols are received. Some well known block codes include the binary Hamming codes, the Golay codes, the Reed-Muller codes, the low density parity check (LDPC) codes, the Bose-Chaudhuri-Hocquenghem (BCH) codes and the Reed-Solomon (RS) codes (which are non-binary BCH codes).

Reed-Solomon codes are the most often used among these codes, as they can correct up to $\lfloor (n-k)/2 \rfloor$ errors or $n-k$ erasures. This property is called maximum distance separable. Increasing the block length n of a RS code with a given code rate (k/n) on a channel with random or correlated losses increases its potential to recover from losses on the one hand. On the other hand increasing the block length introduces delay and increases the decoding complexity, a disadvantage for delay-sensitive applications. Originally the value of n for a RS code with q bit symbols is $2^q - 1$. A typical choice in computer communications is to have bytes as symbols $(q = 8)$ in which case the block length is $n = 255$. The length can be increased up to $2^q + 1$ for extended-RS codes by adding parity symbols. To create codes with block lengths $n < 2^q - 1$ a technique called shortening can be used. $2^q - 1 - n$ symbols are padded with zeros before encoding; the zeroes are not transmitted to the receiver. At the receiver the same symbols are padded with zeros. Shortened RS codes combined with interleaving are used for example in compact discs. For an introduction to block codes see for example [57].

While Reed-Solomon codes are typically used to correct bit errors, they can be used to recover lost packets via block interleaving as described in [58]. Given a block of k packets, the packets are prefixed by their lengths in bytes, and packets shorter than the longest one in the block are padded by zeros. Reed-Solomon coding is applied to the i^{th} symbol (typically byte) of each packet (in total k symbols) to form the i^{th} symbols of the c redundant packets. Packets are then transmitted one-by-one, without the padding zeros. The loss of a packet appears as the loss of a symbol in a block of $c + k$ symbols at the receiver and can be corrected as long as the number of lost packets is no more than c. Figure 4.2 illustrates the use of RS codes for error recovery.

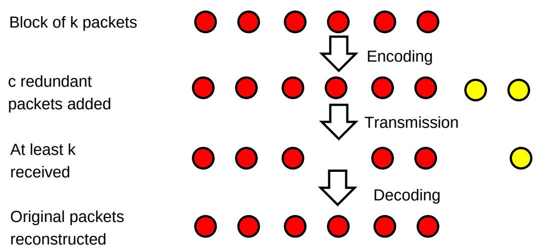


Figure 4.2: Media-independent FEC scheme based on Reed-Solomon codes. If at least k out of $n = k + c$ packets are received, then the original packets can be reconstructed.

The potential of FEC to recover from losses depends on the correlation between losses. Based on earlier results it is clear that the correlation between losses is affected by several factors. The loss process of several sources multiplexed at a single multiplexer was shown not to be independent in [59]. Simulations with video traces [60] and analytical models [61] have both shown that the correlation between losses depends on the level of statistical multiplexing. At the same time the network performance decreases if all sources sharing a bottleneck apply FEC due to the increased load and the resulting increased loss probability [60]. Analytical modeling showed that the burstiness of a source influences the correlations between losses [61].

In Paper B, C and D we add some new aspects to the evaluation of FEC performance. In Paper B we briefly investigate how the burstiness of the background traffic influences the efficiency of FEC. In Papers C and D we show one more factor that can influence the efficiency of FEC, the packet size distribution in the network. We evaluate its effects using mathematical models and via simulations and show that it can influence the performance of FEC both for block based FEC and for MD-FEC.

An interesting area of block codes are the so called rateless codes, which can produce an almost arbitrary number of redundant encodings of the original k symbols of information. Fountain codes [62], which are based on a special type of LDPC codes are an example for such codes. The primary use of such codes is in bulk file transfer, and hence we do not consider them in the thesis.

Convolutional codes

Convolutional codes operate on a continuous flow of symbols. In a convolutional coder each k symbols of information are encoded to n symbols of information, in the encoding the last up to kL symbols are used. The code rate of the coder is k/n and L is called the constraint length of the coder. Codes with higher constraint length provide better protection, but their decoding is more computationally intensive. Codes with arbitrary rates can be achieved by either puncturing rate $1/n$

codes, called mother codes, or by using codes with $k > 1$ and thus at the price of a higher decoding complexity. Convolutional codes are often used in wireless communications with rate 2/3 and below. The most popular algorithm for the decoding of convolutional codes is the Viterbi algorithm, which calculates the maximum likelihood estimate of the original sequence of symbols based on the received sequence of symbols. An advantage of the Viterbi algorithm is that it can incorporate soft decoding easily. Soft decoding assigns probabilities to the output of the decoder, and hence facilitates post-processing of the decoded data. However, the complexity of the Viterbi algorithm grows exponentially with the constraint length of the code. Sequential decoding can be used to decode convolutional codes with high constraint lengths at the price of variable decoding time. The use of convolutional codes in packet communication is not wide-spread, and we do not consider them in this thesis in more detail.

Concatenated codes

Concatenated codes aim to create codes with good performance by combining two or more simple component codes. The concatenated code has the performance of a long code but with significantly smaller decoding complexity. The component codes can be block codes or convolutional codes. A typical solution is to use Reed-Solomon outer code (encoded first, decoded last) and convolutional inner code (decoded first, encoded last). Turbo codes [63] are a type of concatenated codes combined with interleaving and iterative, soft decision decoding. They achieve close to optimal performance while maintaining a moderate decoding complexity. Raptor codes [64], recently adopted by the 3GPP for multicast file delivery, are also a type of concatenated codes, where the innermost code is a Fountain code. Efficient concatenated codes introduce large delays when applied at the packet level and thus we do not consider them in the thesis.

4.3 Joint source-channel coding

The traditional approach to networking, the separation of source and channel coding, was motivated by Shannon's separation theorem [65, 66]. It says that source coding and channel coding can be performed separately while maintaining optimality. However, Shannon's separation theorem assumes that the available delay is unlimited, the channel is stationary and the processing capacity is infinite. None of these criteria are, however, true for real-time applications. While the processing capacity is continuously increasing in accordance with Moore's law, the delay available for multimedia communications is usually limited and the Internet is not stationary [67]. Hence, for delay sensitive multimedia communications over channels with limited bandwidth source and channel coding have to be adjusted jointly.

Rate allocation

Rate allocation deals with the problem of allocating source rate and redundancy rate to achieve minimal distortion given an available total rate and channel conditions. The problem has been addressed for packet level MI-FEC and MPEG-2 coded video based on an empirical mathematical model of video quality in [68]. A similar approach was followed in [41] for bit level FEC and the H.263 video coder. Both studies showed that there exists an optimal allocation of source and channel rate. The problem of finding the optimal rate allocation is, however, problematic in real systems. The first issue is the estimation of the channel state. Errors in the estimate of the packet loss probability and the correlations between losses decrease the gain achievable through optimal rate allocation. On a stationary channel the estimation errors can be made arbitrarily small, and hence in terms of mean distortion the allocation can be optimal, but this approach neglects the short term fluctuations of the channel. Optimization of the rate allocation on a non-stationary channel, whose behavior can not be predicted, is even more problematic.

Another issue is the calculation of the distortion-rate curve in the presence of losses based on feedback from the decoder. Optimal rate allocation assumes that the encoder has a correct model of the transmission channel and hence can simulate the effects of losses on the encoded video stream. We address the issue of optimal rate allocation in Paper G based on analytical models and simulation results with H.264 coded video.

Multiple Description Coding

Multiple description coding (MDC) addresses the problem of joint source and channel coding. Originally it was designed for the transmission of multiple descriptions of a single source over independent channels. If only one of the descriptions is received, it is used for reconstruction with a certain accuracy. If more than one descriptions are received, then the information from the other descriptions can be used to enhance the accuracy (in contrast to MD-FEC, where the redundant copy cannot be used to enhance quality). It has been rediscovered recently for use in packet switched networks [69]. Packets containing the corresponding descriptions can be sent over disjoint paths in the network if it is possible, or instead of using separate channels, one can time-shift the different descriptions, similarly to the case of MD-FEC. Figure 4.3 shows the use of MDC with two descriptions to improve error resilience. In the general case, the amount of information sent over the separate channels (packets) can be different. For single-path packet networks, it can be shown that balanced MDC, i.e. the one sending the same amount of information in all packets, is optimal [70].

There is, however, a trade-off between the distortion when only a subset of the descriptions is received and when all of them are received. The lower the distortion in the case when all descriptions are received the higher the distortion in the case when only a subset of the descriptions is received. This trade-off is illustrated in

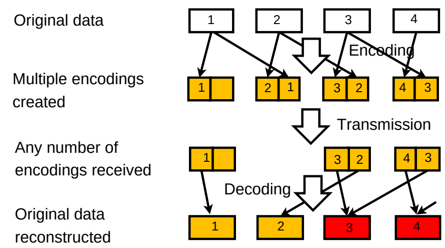
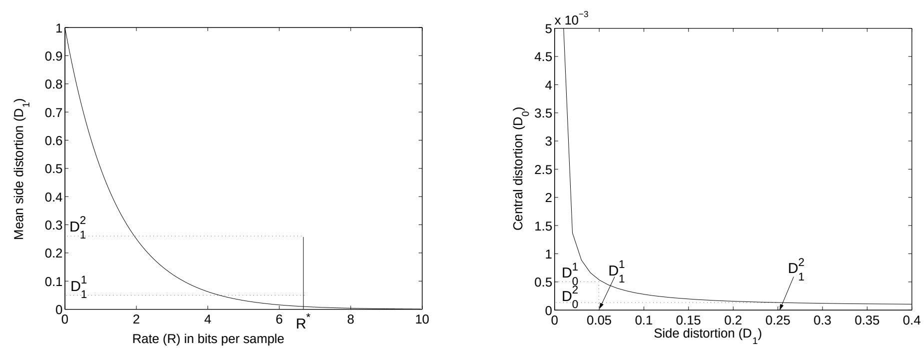


Figure 4.3: Multiple description coding. Several descriptions of the original data are created, the reconstructed quality depends on the number of descriptions received.

Figure 4.4 for the case of two balanced descriptions and a Gaussian source with unit variance [71, 72]. The distortion of the individual descriptions (called side distortion, D_1^1 or D_1^2) can be chosen arbitrarily (as long as it is above the distortion rate curve), but the distortion when both descriptions are received (called central distortion, D_0^1 if the side distortion is D_1^1, or D_0^2 if the side distortion is D_1^2) depends on the choice of the side distortion. If an application is aware of the channel state, i.e. the distribution of the number of lost descriptions, then it can adjust the side and central distortions to minimize the mean distortion.



(a) Distortion-rate curve of the side distortion

(b) Central vs. side distortion for R^*

Figure 4.4: Trade-off in multiple description coding. The choice of the side distortion (D_1^1 and D_1^2 in panel a) influences the central distortion (D_0^1 and D_0^2 in panel b)

There are two main ways of implementing MDC in practice. Algorithms belonging to the first group operate on the source samples. The multiple description scalar quantizer proposed in [69] uses two separate scalar quantizers to create two descriptions. A solution for lattice vector quantizers has been presented in [73]. MD coding based on correlated transforms [74] uses a linear transformation to create two correlated descriptions of two independent random variables after quantization.

The method was extended to an arbitrary number of random variables in [75]. An improvement of the same method was presented in [76], which operates close to the theoretical bound even for high redundancy. The application of these MD coding schemes to motion compensated video raises several issues. The most difficult to handle is that of the prediction errors. Depending on the number of descriptions received, a mismatch condition can exist between the decoder and the encoder until the reception of the next intra-coded frame [77]. The problem of mismatch can be avoided by the price of increased distortion, which makes it difficult to design efficient MD coders with more than two descriptions. The other issue is whether and how to MD code motion vectors and other side information, such as prediction modes, etc. The trivial solution is to repeat such information in each description, which, even though inefficient, is easy to implement [77]. These MD coding methods, despite of their limitations, can be used for interactive multimedia communications as they do not introduce significant delay.

The second approach to implement MDC is based on FEC and layered coding, it is also known as priority encoding transmission (PET) [78]. It uses layered coding to encode a set of frames at once (e.g. all frames of a GOP) and then orders the output bits of the coder in groups in decreasing order of their importance. Block based FEC is used to protect the encoded data, the amount of redundancy added to groups of bits depends on the their importance. The output of the coding scheme is a number of packets, source and redundancy bits belonging to the different groups are distributed among the packets in a way that it is possible to recover the most important bits of the video bitstream upon reception of even a small portion of the packets. To be able to decode the less important bits of the video stream a higher portion of the packets needs to be received. This scheme can achieve good error resilience, it introduces however a considerable amount of delay as its efficiency increases with the number of packets per set of frames. Consequently it is suitable only for streaming video.

4.4 Error control for video transmission

When choosing the best error control solution for video transmission there are basically two constraints that limit the set of available solutions: available bitrate and delay.

For interactive applications, such as videoconferencing, both the available delay and the available bitrate are low. (If it is not limited by the network then it is by the processing power of the user's equipment.) Given a small delay budget for traffic control for such applications, the problem of splitting the delay among the possible functions, such as traffic shaping and FEC, arises. Considering that the efficiency of FEC depends on the burstiness of the traffic and the FEC block length, the optimal allocation is not straightforward. We address this issue in Paper B.

Even interactive applications with strict delay constraints can choose between various different error control solutions to decrease the effects of losses. It is, however, difficult to draw universal conclusions due to the large number of different coding standards. We present a comparison of three error control solutions suitable for interactive communications in Paper E based on analytical models and combine the results of the analytical evaluation with measurement data in Paper F.

For applications that tolerate higher delays than interactive applications, e.g. streaming applications, the available set of error control techniques is bigger. Point-to-point streaming applications can use FEC with large block lengths, which gives strong resilience to losses. They can eventually use retransmission protocols to increase robustness. In the case of multicast streaming retransmission based protocols are not feasible, some form of FEC has to be employed to decrease the effects of losses.

4.5 End-point-based multicast streaming

The simultaneous delivery of information to a large number of destinations puts a big load on the content providers both in terms of processing power and in terms of network bandwidth if the individual destinations are served on a one-by-one basis. Multicast routing in the IP layer was proposed to reduce the load of content providers and use bandwidth resources efficiently. But IP multicast is still not widely available.

Another solution, content distribution networks (CDNs), was deployed commercially instead of IP Multicast. CDNs consist of a large number of servers distributed geographically and share the load between the servers. CDNs are mainly used to distribute static web pages and streaming multimedia content. They offer improved scalability and reliability compared to a single server, and are capable of streaming to a large number of clients. But the costs of streaming increase with the number of clients and sudden increases of the client population, also called flash crowds, often lead to interruptions of the service.

End-point-based multicast is a solution to distribute the costs of large scale simultaneous multicast streaming among the spectators. In end-point-based multicast clients build an overlay for data distribution and forward data to other clients, hence the load of the content provider decreases. Such cooperative overlays can distribute data to a large number of clients in a scalable way, but several issues have to be dealt with.

The first issue is how to minimize the number of free-riders. Free-riders are clients who join the overlay but do not forward data to other clients. Clearly there is a need for some form of incentive-based mechanism or policy enforcement. Possible incentives range from bit-for-bit strategies over distributed reputation management to centralized taxation [79, 80, 81], and are similar in nature to those proposed for file sharing applications. Bit-for-bit schemes allow peers to receive as much data as they send to other peers. Distributed reputation management schemes record the peers' history and allow peers to participate in the overlays based on their history. In the case of centralized taxation resource-rich peers are forced by the content provider to contribute more resources than resource-poor peers.

The second issue is how to handle node departures. Nodes should be allowed to join and leave the overlay at any time, and while joining nodes do not have a disturbing effect on the data distribution, departing nodes can interrupt the flow of data and at the same time can destroy the structure of the overlay. How departures affect the data distribution and the overlay's structure depends very much on the particular solution. Mesh based overlays [82] are usually more robust to node departures, as data can reach nodes on several paths. In tree based overlays the departure of nodes interrupts the flow of data and hence degrades the overlay's performance. However tree based overlays are more scalable and more efficient as data reaches each client only on one path. Some tree-based architectures employ FEC or MDC combined with multiple distribution trees [83, 84] to be more robust to disruptions due to node departures and data loss caused by congestion in the network. How robust these architectures are to node departures and data loss has been the subject of several simulation studies [83, 84, 85]. In Paper G we develop an analytical model for the performance of an overlay based on multiple distribution trees and FEC [84] and use it to evaluate the effects of data loss and node departures in a large overlay.

Chapter 5

Summary of original work

Paper A: On the Efficiency of Shaping Live Video Streams

Gy. Dán, V. Fodor, "On the Efficiency of Shaping Live Video Streams", in Proc. of SPECTS'02, July 2002, San Diego, CA, pp. 49-56.

Summary: An analytical model has been developed and thorough simulations have been performed to evaluate the efficiency of delay limited source shaping for the transmission of live video streams. The goal of the paper is to show that even using computationally simple shaping algorithms significant gain can be reached. We consider MPEG coded video streams multiplexed at a single multiplexer.

Delay limited shaping: We propose two simple shaping algorithms for online shaping of MPEG video with a strict delay bound, one based on the residual transmission delays (RTD), and another, simpler one, which is based on the buffer content only (BC). The efficiency of the two solutions is evaluated by comparing the coefficient of variation (CoV) of the shaped streams. The calculations show that shaping with a delay bound as small as 40 ms results in a significant reduction of the CoV; the marginal gain decreases when increasing the shaping delay further. This gives prospect to the efficient use of traffic shaping in low delay applications. Though the RTD algorithm performs slightly better, we use the BC algorithm throughout our work to show that one can achieve considerable improvements even by using the simplest algorithm.

Analytical model: The analytical model presented in the paper is based on the theory of large deviations and makes it possible to calculate the average loss probability as well as the relative loss probabilities of shaped and unshaped streams. Calculations are made using data derived from actual traces of MPEG videos, and validated using extensive simulations.

Average loss probability: Both the analytical model and simulations show that shaping half of the sources decreases the loss probability of all streams by roughly one order of magnitude. Furthermore streams applying shaping experience a loss probability 20 to 35 percent lower than the average at a load level where the average loss rate is 10^{-5}. The achieved gain might make it desirable for individual users to apply shaping. The calculations show that for the considered traces shaping with a delay limit of 120 ms gives only minor improvement compared to shaping with a delay limit of 40 ms.

Packet loss probabilities in I, P and B frames: It is known that losses in the different types of frames in an MPEG video do not have the same effect on the perceived quality of the reconstructed video. Therefore it is interesting to evaluate the effects of shaping on the losses in the different frame types. The simulations have shown that if shaping is not used, I frames, which have the greatest influence on the perceived quality, have a loss probability which is up to 100 percent higher than the average loss probability experienced by the stream. However by using shaping the packet loss probability among the different frame types can be made equal, which may further improve the performance of the transmission.

Contribution: The original idea came from the second author of the paper. The author of this thesis performed the mathematical analysis based on results in the literature, and carried out the simulations to validate the mathematical model. The article was written in cooperation with the second author of the paper.

Paper B: Quality Differentiation with Source Shaping and Forward Error Correction

Gy. Dán, V. Fodor, "Quality Differentiation with Source Shaping and Forward Error Correction", in Proc. of MIPS 2003, November 2003, Naples, pp. 222-233.

Summary: In this paper the joint use of FEC and delay limited shaping is investigated with respect to delay and loss sensitive video transmission. The results presented in this paper are based on simulations performed with traces of MPEG videos on a single multiplexer.

FEC and shaping for differentiation: The simulation results show that FEC can reduce the loss probability by one to two orders of magnitude depending on the amount of redundancy (overhead) used if only a small portion of the sources applies it. Shaping further decreases the loss probability by 25 to 50 percent. The reason for this is twofold: shaped streams experience a lower loss probability, and the loss process of shaped streams is less correlated which makes FEC more effective.

Optimal allocation of delay between FEC and shaping: Both FEC and delay limited shaping introduce some delay and improve the performance in terms of packet loss as a function of the delay but with a decreasing marginal gain. Finding the right allocation of delay between the two control functions is an optimization problem. The results indicate that even though the lowest average loss probability can be achieved by using FEC only, if we consider the impact of losses in the different frame types on the perceived quality, it is worthwhile to spend a fraction of the available delay on shaping.

FEC redundancy and shaping delay: Given a fixed bandwidth, the use of FEC decreases the effective load due to the overhead imposed. Since shaping and FEC have similar effects it is worthwhile to investigate if shaping can compensate for decreased FEC redundancy. The performed simulations showed that by introducing an additional delay of 60 ms the same loss probability can be achieved at a 33 percent lower redundancy level.

Sensitivity analysis: A weakness of quality differentiation using FEC is its sensitivity to the background traffic characteristics. The considered simulation scenario showed that even though the loss probability of all the streams depends on the burstiness of the background traffic, the difference between the loss experienced by streams with and without FEC is almost constant. By combining FEC with source shaping the robustness of the quality differentiation can be improved further.

Contribution: The original idea came from the second author of the paper. The author of this thesis has carried out the simulations and evaluated the results. The paper was written in cooperation with the second author of the paper.

Paper C: On the Effects of the Packet Size Distribution on the Packet Loss Process

Gy. Dán, V. Fodor, G. Karlsson, "On the Effects of the Packet Size Distribution on the Packet Loss Process", to appear in Telecommunication Systems Journal

Summary: In this paper mathematical models are presented to calculate the probability of j losses in a block of n packets for a bursty stream multiplexed with background traffic with exponential and constant packet sizes. We use the models to investigate whether the packet size distribution in the network affects the correlation structure of losses.

Loss model: The multimedia traffic is modeled by an L-state Markov-modulated Poisson process (MMPP), and the background traffic is represented by a Poisson process. Formulae to calculate the loss probability in a block of packets in the resulting $MMPP + M/M/1/K$ and $MMPP + M/D/1/K$ queues are derived and a numerical method to calculate the probabilities is given.

Model validation: The results given by the models are validated against simulations with an MMPP and with an MPEG-4 coded video trace. The comparison shows a good match between simulation and analytical results.

FEC performance: A comparison of the results obtained with the two models shows that the packet size distribution affects the packet loss process, i.e. the probability of losses in a block and the mean loss run length. We perform the comparison in several scenarios, all of them show that the exponential packet size distribution leads to more correlated losses. The results indicate that from an application's perspective the effects of packet size distribution on the loss process are biggest on access links, where even individual applications can influence the packet size distribution.

Contribution: The idea of developing a mathematical model came from the second author of the paper. The author of this thesis has developed the mathematical models based on results found in the literature, implemented, and carried out the simulations, and analyzed the resulting data. The formulae to calculate the performance measures were derived by the author. The article was written under the supervision of the two co-authors.

Paper D: On the Effects of the Packet Size Distribution on FEC Performance

Gy. Dán, V. Fodor, G. Karlsson, "On the Effects of the Packet Size Distribution on FEC Performance", to appear in Elsevier Computer Networks Journal

Summary: In this paper we present a thorough evaluation of the effects of the packet size distribution on media-dependent and media-independent FEC performance. We present a mathematical model to calculate the probability of j losses in a block of n packets and the consecutive loss probability for a bursty stream multiplexed with bursty background traffic with Erlang-r distributed packet sizes.

Loss model: We model both the multimedia traffic and the background traffic with Markov-modulated Poisson processes (MMPP). We present formulae to calculate the loss probability in a block of packets and the consecutive loss probability in the resulting $MMPP + MMPP/E_r/1/K$ queue.

FEC performance: We compare the performance of FEC for several packet size distributions using the mathematical model and simulations. The results show that the coefficient of variation of the packet size distribution has the biggest effect on FEC performance. The skewness and higher moments have a smaller influence.

Gilbert-model: We evaluate how the Gilbert model's accuracy to predict FEC performance depends on the packet loss probability. We conclude that it mainly depends on it through the average loss probability and the level of statistical multiplexing.

Contribution: The author of this thesis has developed the mathematical model based on results found in the literature, implemented, and carried out the simulations, and analyzed the resulting data. The formulae to calculate the performance measures were derived by the author. The article was written under the supervision of the two co-authors.

Paper E: Are Multiple Descriptions Better than one?

Gy. Dán, V. Fodor, G. Karlsson, "Are Multiple Descriptions Better than one?", in Proc. of 4th IFIP/TC6 Networking 2005, pp. 684-696, May 2005

Summary: We compare the efficiency of three solutions for error control, media-dependent FEC (MD-FEC), media-independent FEC (MI-FEC) and multiple description coding (MDC). We use a simple model from distortion-rate theory and use the mean distortion as performance indicator. To calculate the probability of losses we use the $MMPP + M/E_r/1/K$ model presented in Paper D and a similar model to calculate the conditional loss probability.

MD-FEC vs. MDC: Our results show that MDC always outperforms MD-FEC, i.e. the mean distortion achievable by MDC is always lower than that achievable by MD-FEC. The gain of MDC depends on the packet loss probabilities and the available rate. The results show that below a certain stationary loss probability error control can not decrease the distortion.

MDC vs. MI-FEC: The comparison between MI-FEC and MDC is restricted to MDC with two descriptions. The results show that MDC can outperform MI-FEC if losses are bursty and the available delay for error control is low.

Contribution: The author of this thesis developed the mathematical model based on results found in the literature, and performed the calculations. The article was written under the supervision of the two co-authors.

Paper F: A Rate-distortion Based Comparison of Media-dependent FEC and MDC for Real-time Audio

Gy. Dán, V. Fodor, G. Karlsson, "A Rate-distortion Based Comparison of Media-dependent FEC and MDC for Real-time Audio", to appear in Proc. of IEEE ICC 2006

Summary: We compare the efficiency of MD-FEC and MDC to be used for real-time audio transmission over the Internet. We use a model from distortion-rate theory and the mean distortion as performance indicator. We combine the analytical results for the distortion with statistics derived from a database of packet loss traces measured on the Internet.

Loss model: The evaluation shows that even though MDC always outperforms MD-FEC in terms of mean distortion, the difference is very small if one considers real traces of packet losses. The difference increases as the available rate increases.

Contribution: The author of this thesis processed the database of traces and performed the evaluation. The article was written under the supervision of the two co-authors.

Paper G: Robust Source-channel Coding for Real-time Multimedia

Gy. Dán, V. Fodor, G. Karlsson, "Robust Source-channel Coding for Real-time Multimedia", submitted to ACM Multimedia Systems Journal

Summary: We propose a method for rate allocation that can improve the resilience of multimedia communications in the presence of channel state measurement errors and on non-stationary channels. Instead of minimizing the mean distortion for the measured stationary packet loss probability we propose to minimize the mean distortion based on a model of the channel. We evaluate how estimation errors and the use of simple channel models influence the performance of the rate allocation.

Loss estimation: We use the average loss interval (ALI) method to estimate the channel state. We present formulas for the variance of the estimate on a stationary channel described by the Gilbert-model.

Performance: We compare the mean distortion and the variance of the distortion for the proposed rate allocation method (minmax-α) to that of the minimization of the mean distortion for MD-FEC and MDC. The results show that on a stationary channel minmax-α results in a higher mean distortion but a lower variance of the distortion. On non-stationary channels minmax-α provides smoother degradation of the quality. We apply the minmax-α rate allocation method to H.264 coded video. The simulations show similar results to those obtained with the analytical models based on distortion-rate theory.

Contribution: The author of this thesis has developed the mathematical model, carried out the simulations, and analyzed the resulting data. The article was written under the supervision of the two co-authors.

Paper H: On the Stability of End-point-based Multimedia Streaming

Gy. Dán, V. Fodor, G. Karlsson, "On the Stability of End-point-based Multimedia Streaming", submitted to 5th IFIP/TC6 Networking 2006

Summary: We present a mathematical model of an end-point-based multicast streaming architecture that uses multiple distribution trees and forward error correction (FEC) to improve resilience to node departures and data loss. The model can be used to calculate the portion of the data that a node arbitrarily far away from the root of the multicast receives in the presence of correlated losses due to congestion. We use the theory of discrete dynamic systems to derive conclusion for large overlays.

Static overlay: We validate the mathematical model via simulations, and discuss why there are minor discrepancies under certain conditions between the results. We show how the model can be applied to networks with heterogeneous loss probabilities. The results show that the overlay can deliver data to nodes arbitrarily far away from the root node as long as the loss probability between nodes of the overlay is below a certain threshold. The value of the threshold mainly depends on the parameters of the FEC and the number of trees, the number of nodes per layer has a minor effect. No data is received far from the root node if the loss probability exceeds the threshold. We also show that the overlay is able to recover from severe data losses as long as they do not exceed the threshold of stability.

Dynamic overlay: We show how the model can be used to predict the performance in the presence of node departures. We validate the proposed method via simulations. The results show a good match between the approximative model and the simulations, and suggest that node departures influence the system's performance as a form of correlated losses.

Contribution: The idea of developing a mathematical model came from the second author of the paper. The author of this thesis has developed the mathematical model, implemented, and carried out the simulations, and analyzed the resulting data. The article was written under the supervision of the two co-authors.

Other papers not included in the thesis

Gy. Dán, V. Fodor, *"The Effectiveness of Traffic Shaping in Networks with Small Buffers,"* RVK'02, June 2002

Gy. Dán, V. Fodor, *"Comparison of Shaping and Buffering for Video Transmission,"* in Proc. of NTS 16, pp. 78-87, August 2002.

Gy. Dán, V. Fodor and G. Karlsson *"On the Importance of the Packet Size Distribution in FEC Performance,"* in Proc. of NTS 17, pp. 207-218 August 2004.

Gy. Dán, V. Fodor and G. Karlsson *"Analysis of the Packet Loss Process for Multimedia Traffic,"* in Proc. of ICTSM 12, pp. 83-94, July 2004.

Gy. Dán, V. Fodor and G. Karlsson *"Packet Size Distribution: an Aside?,"* in Proc. of QoSIP 2005, pp. 75-87, Feb. 2005.

Gy. Dán, V. Fodor and G. Karlsson *"On the Asymptotic Behavior of End-point-based Multimedia Streaming,"* to appear in Proc. of International Zürich Seminar on Communications, Feb. 2006.

Chapter 6

Conclusions and future work

This thesis presents an evaluation of how different traffic and error control functions can improve the quality of audiovisual communication. Our results on shaping show that by using traffic shaping the loss probability of the individual streams decreases significantly and losses become more evenly distributed between the different frame types of MPEG coded video. The achieved gain is directly proportional to the introduced delay with a decreasing marginal gain. By taking the influence of the packet losses on the perceived quality into account the use of shaping can bring considerable gains for the individual users. The network operators can benefit from the gradual introduction of shaping in the end nodes as well.

Another benefit of shaping is that it can improve the efficiency of FEC considerably. Even though the optimal combination of shaping and FEC might be difficult to find in practice, it is clear from the results that there exists an optimal combination of shaping and FEC which minimizes the loss of information. Combining the two control functions also increases the robustness of FEC.

The effects of the packet size distribution on the loss process show how difficult it is to predict the performance of the different error control solutions. But knowing the factors that can influence their efficiency helps to design scenarios for the performance evaluation of these schemes. For applications that want to make use of FEC the results give one more aspect of packetization: sending nearly equally sized packets not only decreases the stationary packet loss probability but makes losses less correlated as well.

The results comparing the different error control solutions suggest that research in the field of multiple description coding can give significant benefits, mainly in the field of interactive communications. The achievable gain depends in a large extent on the available rate. Of course, the results presented can not be directly applied to any coding scheme, but they indicate that the achievable gain for low bitrates is minimal. An evaluation with more accurate distortion-rate models could tell how much gain can be expected using particular coding schemes.

Finally the results on end-point-based multicast give a simple means to analyze the behavior of large overlays. The results are encouraging, but at the same time show some of the possible pitfalls of these schemes. Even though in principle the packet reception ratio can be made arbitrary high, the mis-estimation of the network state and the node dynamics can result in all data being lost. Our proposed robust rate allocation method can help to alleviate the effects of mis-estimation by the price of slightly higher mean distortion. We believe that in a cooperative system, where node departures depend partly on the reception quality, it is important to minimize the probability of transient failures as those can lead to increased node departures and hence to the collapse of the overlay.

Future work

The problem of audiovisual communications over the Internet is far from being solved. New coding schemes and error control solutions that provide acceptable quality appear, and how the diverse schemes can be combined to improve the user perceived quality is not clear. Especially in lack of a well defined and understood, easy-to-use model of the human visual system it is not straightforward how quality can be maximized. Complex solutions are often hard to deploy, and hence we have to make the best out of the feasible set of solutions. Questions, like how rate allocation using simple but easy-to-use distortion models can be improved and which combination of the many available features of today's video coders provide the best quality, are still open.

In the field of multicast streaming the peer-to-peer approach can eventually give the opportunity to anyone to become a content provider. It is however not clear, what is the best way to organize nodes into an overlay if timely delivery of the data and robustness are both issues. Another question is how end-point-based multicast would influence the network's performance, as data has to traverse the network several times in an overlay created without any topological considerations. Overlays built using topological information are more sensitive to the failure of single nodes. Another interesting area would be to investigate the use of rateless codes in conjunction with such overlays. A question leading away from the area of networking is the possible set of business models that could make the development and deployment of end-point-based distribution systems economically feasible.

Bibliography

[1] J. H. Saltzer, D. P. Reed, and D. D. Clark, "End-to-end arguments in system design," *ACM Trans. on Computer Systems*, vol. 2, pp. 277–288, November 1984.

[2] L. Breslau, S. Jamin, and S. Shenker, "Comments on the performance of measurement-based admission control algorithms," in *Proc. of IEEE INFOCOM*, pp. 1233–1242, March 2000.

[3] G. Bianchi, A. Capone, and C. Pertioli, "Throughput analysis of end–to–end measurement-based admission control in IP," in *Proc. of IEEE INFOCOM*, pp. 1461–1470, March 2000.

[4] L. Breslau, E. W. Knightly, S. Shenker, I. Stoica, and H. Zhang, "Endpoint admission control: Architectural issues and performance," in *Proc. of ACM SIGCOMM*, pp. 57–69, August 2000.

[5] V. Fodor (née Elek), G. Karlsson, and R. Rönngren, "Admission control based on end-to-end measurements," in *Proc. of IEEE INFOCOM*, pp. 623–630, March 2000.

[6] R. B. Gibbens and F. P. Kelly, "Distributed connection acceptance control for a connectionless network," in *Proc. of the 16th International Teletraffic Congress*, pp. 941–952, June 1999.

[7] G. Karlsson and F. Orava, "The DIY approach to QoS," in *Proc. of IWQoS*, pp. 6–8, May 1999.

[8] I. Más Ivars and G. Karlsson, "PBAC: Probe based admission control," in *Proc. of QofIS, COST 263*, pp. 97–109, September 2001.

[9] M. Reisslein, K. Ross, and S. Rajagopal, "Guaranteeing statistical QoS to regulated traffic: The multiple node case," in *Proc. of IEEE Decision & Control*, pp. 531–538, 1998.

[10] M. Reisslein, K. Ross, and S. Rajagopal, "Guaranteeing statistical QoS to regulated traffic: The single node case," in *Proc. of IEEE INFOCOM*, pp. 1061–1072, 1999.

[11] M. Reisslein, K. Ross, and S. Rajagopal, "A framework for guaranteeing statistical QoS," *IEEE/ACM Trans. Networking*, vol. 10, no. 1, pp. 27–42, 2002.

[12] S. Moon, J. F. Kurose, and D. F. Towsley, "Packet audio playout delay adjustment: performance bounds and algorithms," *Multimedia Systems*, vol. 6, no. 1, pp. 17–28, 1998.

[13] C. Perkins, O. Hodson, and V. Hardman, "A survey of packet loss recovery for streaming audio," *IEEE Network*, 1998.

[14] Y. Wang and Q. Zhu, "Error control and concealment for video communication: a review," *Proc. IEEE*, vol. 86, no. 5, pp. 974–997, 1998.

[15] G. B. G. L. A. Wiegand, T. Sullivan, "Overview of the h.264/avc video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, July 2003.

[16] A. Adas, "Traffic models in broadband networks," *IEEE Commun. Mag.*, vol. 35, pp. 82–89, 1997.

[17] O. Rose, "Simple and efficient models for variable bit rate MPEG video traffic," *Performance Evaluation*, vol. 30, no. 1-2, pp. 69–85, 1997.

[18] M. Izquierdo and D. Reeves, "A survey of statistical source models for variable-bit-rate compressed video," *Multimedia Systems*, vol. 7, pp. 199–213, 1999.

[19] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, and J. Robbins, "Performance models of statistical multiplexing in packet video communications," vol. 36, no. 7, pp. 834–844, 1988.

[20] D. Cohen and D. Heyman, "Performance modeling of video teleconferencing in atm networks," vol. 3, no. 6, pp. 408–420, 1993.

[21] R. Grünenfelder, J. Cosmas, S. Manthorpe, and A. Odinma-Okafor, "Characterization of video codecs as autoregressive movind average processes and related queueing system performance," *IEEE J. Select. Areas Commun.*, vol. 9, no. 3, pp. 284–293, 1991.

[22] J. Beran, R. Sherman, M. Taqqu, and W. Willinger, "Long-range dependence in variable-bit-rate video traffic," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 1566–1579, 1995.

[23] M. Garrett and W. Willinger, "Analysis, modeling and generation of self-similar vbr video traffic," in *Proc. of ACM SIGCOMM*, pp. 269–280, Aug. 1994.

[24] C. Huang, M. Devetsikiotis, I. Lambadaris, and R. Kaye, "Modeling and simulation of self-similar variable bit-rate compressed video: A unified approach," in *ACM SIGCOMM*, pp. 114–125, 1995.

[25] M. Grasse, M. Frater, and J. Arnold, "A new approach to long-range dependence in variable bit rate video traffic," *Telecomm. Systems*, vol. 12, no. 1, pp. 79–100, 1999.

[26] B. Ryu and A. Elwalid, "The importance of long-range dependence of VBR video traffic in ATM traffic engineering: Myths and realities," in *Proc. of ACM SIGCOMM*, pp. 3–14, 1996.

[27] M. Grossglauser and J.-C. Bolot, "On the relevance of long-range dependence in network traffic," *IEEE/ACM Trans. Networking*, vol. 7, no. 5, pp. 629–640, 1999.

[28] S. Robert and J. Le Boudec, "New models for pseudo self-similar traffic," *Performance Evaluation*, vol. 30, no. 1–2, pp. 57–68, 1997.

[29] A. T. Andersen and B. F. Nielsen, "A markovian approach for modeling packet traffic with long-range dependence," *IEEE J. Select. Areas Commun.*, vol. 16, no. 5, pp. 719–732, 1998.

[30] M. Krunz and S. Tripathi, "On the characterization of VBR MPEG streams," in *ACM Sigmetrics*, pp. 192–202.

[31] M. Dai and D. Loguinov, "Analysis and modeling of MPEG-4 and H.264 multi-layer video traffic," in *Proc. of IEEE INFOCOM*, pp. 2257–2267, Mar. 2005.

[32] A. Lazar, G. Pacifici, and D. Pendarakis, "Modeling video sources for real-time scheduling," *Multimedia Systems*, vol. 1, no. 6, pp. 253–266, 1994.

[33] Video Quality Experts Group, "Final report from the video quality experts group on the validation of objective models of video quality assessment, phase II," July 2003.

[34] A. Webster, C. Jones, M. Pinson, S. Voran, and S. Wolf, "An objective video quality assessment system based on human perception," in *Proc. SPIE-Human Vision, Visual Processing and Digital Display*, vol. 1913, pp. 15–26, 1993.

[35] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[36] S. Daly, "The visible difference predictor: An algorithm for the assessment of image fidelity," in *Proc. SPIE*, vol. 1616, pp. 2–15, 1992.

[37] T. Pappas and R. Safranek, *Perceptual criteria for image quality evaluation, Handbook of Image and Video Processing (A. Bovik, ed.).* Academic Press, May 2000.

[38] S. Winkler, "A perceptual distortion metric for digital color video," in *Proc. SPIE*, vol. 3644, pp. 175–184, 1999.

[39] C. van den Branden Lambrecht and O. Verscheure, "Perceptual quality measure using a spatio-temporal model of the human visual system," in *Proc. SPIE*, vol. 2668, pp. 450–461, 1996.

[40] K. Tan, M. Ghanbari, and D. Pearson, "An objective measurement tool for MPEG video quality," *Signal Processing*, vol. 70, pp. 279–294, 1998.

[41] K. Stuhlmüller, N. Färber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Select. Areas Commun.*, vol. 18, June 2000.

[42] Y. Liang, J. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: does burst-length matter?," in *Proc. of IEEE ICASSP*, pp. 684–687, April 2003.

[43] O. Verscheure, X. Garcia, G. Karlsson, and J. Hubaux, "User-oriented QoS in packet video delivery," *IEEE Network*, vol. 12, no. 6, pp. 12–21, 1998.

[44] W. Feng, F. Jahanian, and S. Sechrest, "An optimal bandwidth allocation strategy for the delivery of compressed prerecorded video," *ACM Multimedia Systems*, vol. 5, no. 5, pp. 297–309, 1997.

[45] W. Feng and J. Rexford, "A comparison of bandwidth smoothing techniques for the transmission of prerecorded compressed video," in *IEEE INFOCOM*, pp. 58–66, 1997.

[46] J. McManus and K. Ross, "Video on demand over ATM: Constant-rate transmission and transport," *IEEE J. Select. Areas Commun.*, vol. 14, no. 6, pp. 1087–1096, 1996.

[47] J. Rexford, S. Sen, J. Dey, W. Feng, J. Kurose, J. Stankovic, and D. Towsley, "Online smoothing of live, variable-bit-rate video," in *Proc. of NOSSDAV*, pp. 249–257, May 1997.

[48] S. S. Lam, S. Chow, and D. K. Y. Yau, "An algorithm for lossless smoothing of MPEG video," in *Proc. of ACM SIGCOMM*, pp. 281–293, 1994.

[49] E. Knightly and H. Zhang, "Traffic characterization and switch utilization using a deterministic bounding interval dependent traffic model," in *Proc. of IEEE INFOCOM*, 1995.

[50] H. Zhang and E. Knightly, "RED-VBR: A renegotiation-based approach to support delay-sensitive VBR video," *ACM Multimedia Systems Journal*, vol. 5, pp. 164–176, May 1997.

[51] B. Girod and N. Färber, "Feedback-based error control for mobile video transmission," *Proc. IEEE*, vol. 87, December 1999.

[52] The Multimedia Integrated Conferencing for European Researchers (MICE) Project, "RAT: Robust Audio Tool." http://www-mice.cs.ucl.ac.uk/multimedia/software/rat.

[53] A. V. García and S. Fosse-Parisis, "Freephone audio tool." http://www-sop.inria.fr/rodeo/fphone.

[54] J. C. Bolot, S. Fosse-Parisis, and D. Towsley, "Adaptive FEC-based error control for Internet telephony," in *Proc. of IEEE INFOCOM*, pp. 1453–1460, 1999.

[55] I. Kouvelas, O. Hodson, W. Hardman, and J. Crowcroft, "Redundancy control in real-time Internet audio conferencing," in *Proc. of International Workshop on Audio-Visual Services Over Packet Networks*, September 1997.

[56] P. Dube and E. Altman, "Utility analysis of simple FEC schemes for VoIP," in *Proc. of IFIP Networking 2002*, pp. 226–239, May 2002.

[57] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* Elsevier, New York, 1978.

[58] M. Luby, L. Vicisano, J. Gemmell, L. Rizzo, M. Handley, and J. Crowcroft, "The use of forward error correction (FEC) in reliable multicast," RFC 3453, December 2002.

[59] I. Cidon, A. Khamisy, and M. Sidi, "Analysis of packet loss processes in high speed networks," *IEEE Trans. Inform. Theory*, vol. 39, no. 1, pp. 98–108, 1993.

[60] E. Biersack, "Performance evaluation of forward error correction in ATM networks," in *Proc. of ACM SIGCOMM*, pp. 248–257, August 1992.

[61] K. Kawahara, K. Kumazoe, T. Takine, and Y. Oie, "Forward error correction in ATM networks: An analysis of cell loss distribution in a block," in *Proc. of IEEE INFOCOM*, pp. 1150–1159, June 1994.

[62] J. Byers, M. Luby, and M. Mitzenmacher, "A digital fountain approach to asynchronous reliable multicast," *IEEE J. Select. Areas Commun.*, vol. 20, no. 8, pp. 1528–1540, 2002.

[63] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, 1996.

[64] M. Luby, M. Watson, T. Gasiba, T. Stockhammer, and W. Xu, "Raptor codes for reliable download delivery in wireless broadcast systems," in *Proc. of IEEE Computer Communications and Networking Conference*, Jan. 2006.

[65] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, pp. 379–423, 1948.

[66] S. Vembu, S. Verdu, and Y. Steinberg, "The source-channel separation theorem revisited," *IEEE Trans. Inform. Theory*, vol. 41, pp. 44–54, January 1995.

[67] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker, "On the constancy of internet path properties," in *Proc. of Internet Measurement Conference*, pp. 197–211, 2001.

[68] P. Frossard, "FEC performances in multimedia streaming," *IEEE Commun. Lett.*, vol. 5, no. 3, pp. 122–124, 2001.

[69] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 39, pp. 821–834, May 1993.

[70] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Processing Mag.*, Sept. 2001.

[71] L. Ozarow, "On a source-coding problem with two channels and three receivers," *Bell Syst. Tech. J.*, vol. 59, pp. 1909–1921, December 1980.

[72] R. Zamir, "Gaussian codes and Shannon bounds for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2629–2635, November 1999.

[73] V. A. Vaishampayan, N. J. A. Sloane, and S. D. Servetto, "Multiple description vector quantization with lattice codebooks: Design and analysis," *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 1718–1734, 2001.

[74] Y. Wang, M. T. Orchard, V. A. Vaishampayan, and A. R. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. Image Processing*, vol. 10, no. 3, 2001.

[75] V. K. Goyal and J. Kovačevič, "Generalized multiple description coding with correlating transforms," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2199–2224.

[76] Y. Wang, A. R. Reibman, and H. Orchard, M. T. amd Jafarkhani, "An improvement to multiple description transform coding," *IEEE Trans. Signal Processing*, vol. 50, no. 11, pp. 2843–2853.

[77] A. Reibman, R. Puri, H. Jafarkhani, Y. Wang, and M. Orchard, "Multiple description video coding using motion-compensated temporal prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 193–204, March 2002.

[78] B. Lamparter, A. Albanese, M. Kalfane, and L. M., "PET - priority encoding transmission: A new, robust and efficient video broadcast technology," in *Proc. of ACM Multimedia*, 1995.

[79] B. Cohen, "Incentives build robustness in bittorrent," in *Proc. of P2P Economics Workshop*, 2003.

[80] S. Lee, R. Sherwood, and B. Bhattacharjee, "Cooperative peer groups in NICE," in *Proc. of IEEE INFOCOM*, pp. 1272–1282, Apr. 2003.

[81] Y. Chu, J. Chuang, and H. Zhang, "A case for taxation in peer-to-peer streaming broadcast," in *Proc. of ACM SIGCOMM Workshop on Practice and Theory of Incentives and Game Theory in Networked Systems (PINS)*, 2004.

[82] Y. Chu, S. Rao, S. Seshan, and H. Zhang, "A case for end system multicast," *IEEE J. Select. Areas Commun.*, vol. 20, no. 8, 2002.

[83] M. Castro, P. Druschel, A. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, "Splitstream: High bandwidth content distribution in a cooperative environment," in *Proc. of IPTPS*, 2003.

[84] V. Padmanabhan, H. Wang, and P. Chou, "Resilient peer-to-peer streaming," in *Proc. of IEEE ICNP*, pp. 16–27, 2003.

[85] K. Sripanidkulchai, A. Ganjam, B. Maggs, and H. Zhang, "The feasibility of supporting large-scale live streaming applications with dynamic application end-points," in *Proc. of ACM SIGCOMM*, pp. 107–120, 2004.

On the Efficiency of Shaping Live Video Streams

Gy. Dán and V. Fodor

In Proceedings of the 2002 International Symposium on Performance Evaluation of Computer and Telecommunications Systems (SPECTS '02),
July 2002, San Diego, California

Quality Differentiation with Source Shaping and Forward Error Correction

Gy. Dán and V. Fodor
In Proceedings of the 1st International Workshop on Multimedia Interactive Protocols and Systems (MIPS 2003),
November 2003, Naples, Italy

On the Effects of the Packet Size Distribution on the Packet Loss Process

Gy. Dán, V. Fodor and G. Karlsson
to appear in Telecommunication Systems Journal

On the Effects of the Packet Size Distribution on FEC Performance

Gy. Dán, V. Fodor and G. Karlsson
to appear in Computer Networks

Are Multiple Descriptions Better than one?

Gy. Dán, V. Fodor and G. Karlsson
in Proceedings of 4th IFIP/TC6 Networking
May 2005, Waterloo, Ontario, California

A Rate-distortion Based Comparison of Media-dependent FEC and MDC for Real-time Audio

Gy. Dán, V. Fodor and G. Karlsson
to appear in Proceedings of IEEE ICC 2006
June 2006, Istanbul, Turkey

Robust Source-channel Coding for Real-time Multimedia

Gy. Dán, V. Fodor and G. Karlsson
submitted to ACM Multimedia Systems Journal

On the Stability of End-point-based Multimedia Streaming

Gy. Dán, V. Fodor and G. Karlsson
submitted to 5th IFIP/TC6 Networking
May 2006, Coimbra, Portugal