2010-08-04

Scalable Video Transport over IP Networks

Dian Fan
University of Miami, dian_fan2002@yahoo.com

UNIVERSITY OF MIAMI

SCALABLE VIDEO TRANSPORT OVER IP NETWORKS

By

Dian Fan

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

August 2010

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

SCALABLE VIDEO TRANSPORT OVER IP NETWORKS

Dian Fan

Approved:

James W. Modestino, Ph.D. Professor of Electrical and Computer Engineering	Terri A. Scandura, Ph.D. Dean of the Graduate School
Akmal Younis, Ph.D. Associate Professor of Electrical and Computer Engineering	Manohar N. Murthi, Ph.D. Associate Professor of Electrical and Computer Engineering
Xiaodong Cai, Ph.D. Assistant Professor of Computer Science	Dilip Sarkar, Ph.D. Associate Professor of Computer Science

FAN, DIAN (Ph.D., Electrical and Computer Engineering)

Scalable Video Transport Over IP Networks (August 2010)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor James W. Modestino.
No. of pages in text. (113)

With the advances in video compression and networking techniques, the last ten years have witnessed an explosive growth of video applications over the Internet. However, the service model of the current best-effort network was never engineered to handle video traffic and, as a result, video applications still suffer from varying and unpredictable network conditions, in terms of bandwidth, packet loss and delay. To address these problems, a lot of innovative techniques have been proposed and researched. Among them, scalable video coding is a promising one to cope with the dynamics of the available bandwidth and heterogeneous terminals. This work aims at improving the efficacy of scalable video transport over IP networks.

In this work, we first propose an optimal interleaving scheme combined with motion-compensated fine granularity scalability video source coding and unequal loss protection schemes, under an imposed delay constraint. The network is modeled as a packet-loss channel with random delays. The motion compensation prediction, ULP allocation and the depth of the interleaver are jointly optimized based on the network status and the delay constraint.

We then proceed to investigate the multiple path transport technique. A unified approach which incorporates adaptive motion compensation prediction, multiple description coding and unequal multiple path allocation, is proposed to improve both

the robustness and error resilience property of the video coding and transmission system, while the delivered video quality is improved simultaneously.

To analytically investigate the efficacy of error resilient transport schemes for progressively encoded sources, including unequal loss protection, best-effort and FEC transport schemes, we develop evaluation and optimization approaches for these transport schemes. In this part of the work, the network is modeled as an $M/D/1/K$ queue, and then a comprehensive queueing analysis is provided. Armed with these results, the efficacy of these transport schemes for progressively encoded sources are investigated and compared.

To my parents and my wife

Acknowledgements

First, I would like to express my sincere gratitude to my advisor Professor James W. Modestino, who introduced me to the research field of video transport and provided me his encouragement and advice during my doctoral program. This dissertation would have never been possible without Dr.Modestino's help.

I would like to thank my Ph.D. Dissertation Committee members: Professor Akmal Younis, Professor Manohar N. Murthi, Professor Xiaodong Cai, and Professor Dilip Sarkar for their valuable input for my work, and for the time they have spent with me on this dissertation.

I also would like to thank Dr. Yee Sin Chan and Dr. Xunqi Yu for their helpful suggestions and discussions during my research.

Last but not least, I would like to thank my parents and my wife. Their support and encouragement help me go through all stages of my life.

Dian Fan

University of Miami

August 2010

Table of Contents

List of Figures

CHAPTER 1

Introduction

1.1 Background and Related Works

With the advances in video compression and networking techniques, the last ten years have witnessed an explosive growth of video applications over the Internet. Especially video streaming applications, such as YouTube, are used every day by people around the world. However, the service model of the current best-effort Internet was never engineered to handle video traffic and, as a result, video applications still suffer from varying and unpredictable network conditions, in terms of bandwidth, packet loss and delay.

In most of the video coding standards, from the early H.261 standard to the latest paradigms, such as H.264 and MPEG-4, motion-compensated prediction techniques have been widely used [3–10]. In motion-compensated prediction, a video frame can be predicted by reference frames, which can be previous in time or from the future. The motion-compensated prediction can efficiently remove the temporal redundancy across successive frames and greatly improve the coding efficiency [11]. However, using motion-compensated prediction leads to the fact that proper decoding of such inter-coded frames depends on the error-free reception and reconstruction of their reference frames. As a result, typical encoded video streams are highly structured

and prioritized, and the packets associated with a reference frame are more important than those containing information of frames predicted by the reference frame. If a packet associated with a reference frame is lost, error will be propagated between frames, which is known as the drifting phenomenon.

In contrast to applications such as web text/image and email, video applications are usually delay sensitive. The packets that do not arrive at the receiver in time are considered useless and discarded. Especially, in real-time interactive applications, such as video conference and video phone, the delay constraint is kept on the order of several hundred milliseconds. Excessive delay could severely impair communication interactivity. These characteristics of video applications - hierarchical structure of encoded streams, drifting phenomenon due to packet losses and timing considerations - play an important role in design paradigms and research emphases.

Characterized by the "best effort" service model, the deployment of current networks cannot provide reliable transport of packets and the quality-of-service (QoS) cannot be guaranteed. Furthermore, each packet is treated equally regardless of its priority and content. Packets may be lost due to network congestion in bottleneck nodes or be discarded due to excessive bit errors and failure to pass the cyclic redundancy check (CRC) at the link layer. For wired networks with highly reliable transmission media, network congestion in bottleneck nodes is the major reason for packet losses. Therefore, a packet loss is usually followed by several successive losses, which is characterized as a bursty packet loss process [12–14].

The received packets also experience variable network delay, which includes transmission, propagation and queueing delay. The transmission delay is the amount of time required for the routers to sent out the packet. It is a function of the packet's length and the transmission rate of the link. The propagation delay is the time it

takes a bit to propagate from one router to the next. It is a function of the distance between the two routers. In modern high-speed networks, both transmission delay and propagation delay are usually very small. The key component of network delay is the queueing delay, which is the time that a packet waits to be sent out in buffers of the routers. Queueing delay can vary significantly from packet-to-packet. Because queueing delay is caused by network congestion in bottleneck nodes, a packet with large delay is usually followed by several packets with relatively large delay as well. This phenomenon is referred as the "delay spike" [12, 15, 16].

In a nutshell, the importance of video packets are different while best-effort networks treat every packet equally; video applications require bounded network delay while the best-effort networks provide no delay guarantee. As a result, video applications suffer from the QoS limitations of best-effort networks. Moreover, as mentioned previously, a packet loss in the network is usually followed by several successive losses, which is referred as a bursty packet loss process. This bursty packet loss is another major challenge of video transport because the bursty loss could produce a much larger distortion than an equal number of isolated losses [17].

To address the challenges of video transport over best-effort networks, research efforts have been directed toward communication efficacy, error-robustness, low latency, and scalability [18–24]. And various techniques have been proposed and researched in both the academic and industrial communities.

Because packet loss is unavoidable in best-effort networks, error-control mechanisms need to be employed to mitigate the effects of packet losses. Typically, error-control mechanisms can be classified into four categories: retransmission, error resilience, error concealment and forward error correction (FEC). Retransmission-based schemes, such as automatic repeat request (ARQ), retransmit the packets that are

lost by the network. These retransmission-based schemes are typically dismissed as a way to protect real-time video applications because they introduce excessive delay for each lost packet. Error resilience schemes enhance the capability of the encoder to produce a compressed bitstream better able to resist channel errors [25–30]. Error concealment is a post-processing technique employed in the decoder. It can hide the glitches from the viewer so that a more visually pleasing rendition of the decoded video can be obtained [31–34]. FEC is a very useful technique to protect real-time video streams over packet loss networks [23, 35–41]. With FEC schemes, redundant packets are transmitted along with the information packets so that the lost information packets can be recovered, at least in part, from the redundant packets. However, the bursty nature of the packet loss process in the Internet can limit the efficacy of FEC schemes [12].

To cope with the dynamics of the available bandwidth and heterogeneous terminals, scalable or layered video coding schemes have been proposed in [42–50]. These coding schemes partition the compressed bitstream into a base layer (BL) and one or more enhancement layers (ELs). The BL contains the most important information, such as coding mode, low frequency coefficients and motion vectors (MV). The ELs contain the information that can improve the video quality. The BL can be decodable independently, but the ELs must be decoded cumulatively, i.e., layer i can only be decoded along with layers 1 to $i-1$, which is also referred to as progressive refinement. A typical scalable video encoder can provide three types of scalabilities: temporal scalability, spatial scalability and signal-to-noise ratio (SNR) scalability. With temporal scalability, a video sequence is encoded into several layers at the same spatial resolution, but different frame rates. The BL is at a lowest frame rate. The ELs provide higher frame rates. With spatial scalability, a video sequence is encoded

into several layers at the same frame rate, but different spatial resolutions. The BL is coded at a lowest spatial resolution. Typically, the reconstructed base-layer picture is up-sampled to form the prediction for the high-resolution picture in the ELs. SNR scalability, also referred to as quality scalability, is a technique to code a video sequence into several layers with the same spatio-temporal resolution, but different quality. The BL is coded at a lowest quality and used to predict the higher quality ELs. In traditional scalable video there is no partition within an EL, which means an EL is either completely decoded or discarded. To get more flexibility, the MPEG-4 committee developed the Fine Granularity Scalability (FGS) profile [50]. In this scheme, a single EL is coded progressively such that it can be truncated at any arbitrary location, which provides fine granularity of reconstructed video quality proportional to the number of bits actually decoded. But the coding efficiency is decreased because of the lack of exploiting temporal dependency at the FGS EL. To address this shortcoming, motion compensation prediction loops are employed in the coding procedure of ELs [51, 52], which is known as motion-compensated FGS (MC-FGS) scheme. However, one of the major disadvantages of scalable coding is that the dependency across the compressed bitstream is increased even more by introducing the predication across layers.

Recently, multiple description (MD) coding, which is sometimes referred to as source-based diversity [53], has emerged as a promising technique to transport video in packetized networks [30]. Like so much of communication technology, MD coding was invented at Bell Laboratories in connection with communicating speech over the telephone network [54]. Earlier work on MDC from the perspective of information theory can be found in [55–65]. Recently, MD coding techniques have been widely applied in video communication [30, 66–69]. In MD coding, the source is encoded into

multiple bitstreams (descriptions) of approximately equal importance such that each description can be decoded independently. This property makes it better suited to the current deployment of the Internet, which treats each packet equally. Moreover, the loss of a particular description does not jeopardize the decoding process while the end-quality improves as the number of description received increases. In other words, it can provide adequate quality without requiring retransmission of any lost packets. Hence, MD coding is particularly appealing for real-time video applications, for which retransmission is unacceptable because of the excessive delay it introduces. Various state-of-the-art source coding techniques, such as subsampling in the spatial, temporal, or frequency domain, MD quantization, and MD transform coding, have been developed to generate the multiple independently decodable codes [59, 69–72]. A disadvantage of these MD-based source coding algorithms is that the design of the encoder becomes very complex when more than two descriptions are needed. Instead of designing the source encoder to yield MDs directly, one can generate MDs by applying unequal loss protection (ULP) to different parts of a scalable bitstream, which is commonly known as MD-FEC [2,73]. MD-FEC can be applied to any scalable video stream to generate an arbitrary number of descriptions with approximately equal importance [1, 2, 73–78].

Multiple path transport (MPT) is another powerful approach to combat the bursty packet losses, which is particularly efficient when combined with MD coding [79–83]. Specifically, instead of using only a single transmission path, multiple independent routing paths are employed to transport packets in MPT. This is sometimes referred to as path diversity. By taking advantage of the uncorrelated characteristics of the loss process on different network paths, the virtual average path exhibits a smaller variability in communication quality [70]. MPT technology has attracted intensive

interest from the video transport community. In [84], it is shown that MPT transforms a single path with bursty loss behavior into multiple paths with uniform loss for which FEC is quite effective. Similar research is reported in [85], where a single description of a video sequence is transported by multiple paths combined with packet-level FEC and the efficacy of path diversity is evaluated. To maximize the benefit, most works on MPT are combined with some form of MD coding. In the video communication system proposed in [70], odd and even frames of a video sequence are independently encoded and transmitted along network paths. This encoding scheme falls in the category of MD subsampling in the temporal domain. A model-based performance analysis for this MD video coding scheme combined with MPT is presented in [86], which demonstrates that sending multiple descriptions along multiple independent paths can effectively reduce the end-to-end video distortion compared to conventional single description coding transmitted over a single path. Recognizing that the transport capability of network paths are variant, video traffic is allocated into different paths based on their transport capability in [87–89]. Another advantage of MPT is it can alleviate the problem that the default path determined by the routing algorithm is not optimum, which might often be the case according to [90]. In [91], an optimal multiple path selection scheme for MD video streaming is proposed, which shows that avoiding joint links does not guarantee the best quality of video and high quality joint links should not be sacrificed.

Packet interleaving is a traditional way to cope with bursty losses in communication systems, which is furthermore used widely in audio and video transport [92–97]. In a typical block interleaving scheme, packets are buffered into the interleaver column-by-column and then sent out to the network row-by-row. Consecutive packets from the source encoder are then spaced in time when they are sent to the

network, which reduces the bursty losses and produces lower total distortion. Obviously, the effectiveness of an interleaving scheme depends on the interleaving depth, which is defined as the number of transmission intervals between successive application packets, i.e., the number of columns in the block interleaver. The larger the interleaving depth, the better the expected performance. When the interleaving depth is infinite, the bursty channel is converted into an independent loss channel. However, the larger the interleaving depth, the greater the delay that will be introduced into the transmission.

In this work we aim at improving the efficacy of scalable video transport over best-effort networks and formulating methodologies to evaluate and optimize these video transport schemes. This work is organized as follows. In Chapter 2, we study optimal interleaving combined with MC-FGS and ULP schemes under an imposed delay constraint. The interleaver, the video source coder and the ULP allocation are jointly optimized based on the network state and subject to an imposed delay constraint.

In Chapter 3, we propose a unified approach incorporating adaptive motion compensation prediction, multiple description coding and unequal multiple path allocation so as to improve both the robustness and error resilience property of the video coding and transmission system, while simultaneously improving the delivered video quality.

In Chapter 4 we propose and investigate a model-based analytical approach for evaluating the efficacy of error resilient transport schemes for a scalably encoded source operating over a congested network, including unequal loss protection (ULP), best-effort and optimal FEC schemes. Firstly, we formulate appropriate evaluation and optimization approaches for these three transport schemes. Then, a queueing

analysis is developed based on the $M/D/1/K$ queue model for describing and characterizing each of the transport networks. Finally, the efficacy of these three schemes is evaluated using a common modeling framework.

Chapter 5 presents the conclusions and points out future directions for extending the current work.

1.2 Summary of Major Contributions

The major contributions of this work are summarized as follows:

- A video transport system which integrates MC-FGS video source coding, ULP and interleaving schemes is proposed. Under an imposed delay constraint, the source coder, ULP allocation and interleaver are jointly optimized. The simulation results demonstrate that the proposed system can substantially improve the video transport efficacy.

- A packet loss model is developed, which can be used to derive the block error density function at any packet sending rate from the channel average packet loss rate and average burst time.

- A unified approach incorporating adaptive motion compensation prediction, multiple description coding and unequal multiple path allocation is developed to improve both the robustness and error resilience property of the video coding and transmission system, while simultaneously improving the delivered video quality.

- Explicit queueing analysis approaches are formulated based on an $M/D/1/K$ queueing model to evaluate the corresponding end-to-end performance of three

transport schemes, including ULP, best-effort and optimal FEC transport schemes.

- The efficacy of ULP, best-effort and optimal FEC transport schemes are quantitatively analyzed and compared within the common modeling framework under both idealized source modeling assumptions as well as for real-world video sources.

CHAPTER 2

Delay-Constrained Motion-Compensated FGS Video Transport With Optimal Interleaving

2.1 Motivation

To cope with the dynamics of the available bandwidth and support of heterogeneous terminals, the MPEG-4 committee developed the fine granularity scalability (FGS) profile [50]. In this FGS video coding scheme, a single Enhancement Layer (EL) is coded progressively without motion-compensated prediction (MCP) such that it can be truncated at any arbitrary location, which provides fine granularity of reconstructed video quality adaptive to fluctuations in the network state. However, because MCP is not used at the FGS EL, the coding efficiency is compromised. To address this shortcoming, it has been proposed that a portion of the EL should be used to predict the subsequent frames through MCP. This is known as motion-compensated FGS (MC-FGS) [51].

The MC-FGS video bitstreams are highly structured and prioritized while the Internet treats all packets equally regardless of their priority and content. To provide a better match to the current transport capabilities of the Internet, MC-FGS video coding can be combined with ULP using maximum-distance separable (MDS) erasure

codes, such as Reed-Solomon (RS) codes, to generate an arbitrary number of packets that can be decoded independently in a non-hierachical manner [1, 2, 78]. Through this approach, the information content of each video frame is distributed into multiple independent descriptions, and the bitstream is transformed into a non-prioritized packet stream. As a result, the quality of a reconstructed video frame becomes a nondecreasing function of the number of received packets, regardless of the location of lost packets. More details on the MC-FGS and ULP schemes will be discussed in what follows.

In the Internet, a packet loss is usually followed by several successive losses, which is characterized as a bursty packet loss process. Bursty packet losses can be very harmful to video applications since a burst loss generally produces a larger video distortion than an equal number of isolated losses [17]. Packet interleaving is an effective way to combat bursty losses. For example, in a block interleaver, packets are buffered into the interleaver column-by-column and then sent out to the network row-by-row. Consecutive packets from the source encoder are then spaced in time when they are sent to the network, which reduces the bursty losses and produces lower total distortion. Obviously, the effectiveness of an interleaving scheme depends on the interleaving depth, which is defined as the number of transmission intervals between successive application packets, i.e., the number of columns in the block interleaver. The larger the interleaving depth, the better the expected performance. When the interleaving depth is infinite, the bursty channel is converted into an independent loss channel. However, the larger the interleaving depth, the greater the delay that will be introduced into the transmission. For delay-constrained video applications, the packets are useless and will be discarded if they arrive at the receiver after the playout deadline of the corresponding video frames. Hence, the interleaver depth

should be optimized to mitigate the effect of burst errors subject to an imposed delay constraint.

In the previous literature, interleaving schemes have been used extensively to transport video over bursty loss channels. In [94], given a constant delay budget for interleaving, the interleaver is optimized to transmit video over a wireless network. In [95], a Group Of Picture (GOP) level interleaving scheme is combined with ULP and retransmission to transport 3D-SPIHT video over the Internet. In this Chapter, we study optimal interleaving combined with MC-FGS and ULP schemes. The interleaver, the MC-FGS video source coder and the ULP allocation are jointly optimized based on the network state and subject to an imposed delay constraint.

The remainder of this Chapter is organized as follows. In Section 2.2, we give a brief description of the MC-FGS video coding scheme and the ULP scheme. Section 2.3 presents the joint optimization scheme of source coder, ULP and interleaver. Section 2.4 provides simulation results illustrating the resulting improvement. Finally, Section 2.5 concludes this Chapter.

2.2 Preliminaries

2.2.1 MC-FGS Video Coding

The video coder employed in this work is based on the MPEG-4 fine granularity scalable (FGS) video coder. The FGS video coding scheme was developed to cope with dynamic channel conditions and increasingly heterogeneous network environments [50]. It partitions the compressed bitstream into a base layer (BL) and a progressively (embedded) encoded FGS enhancement layer (EL) such that the EL can be truncated at any arbitrary location. The BL carries the most important infor-

mation, such as coding mode, low-frequency coefficients and motion vectors (MVs), while subsequent ELs contain information that can improve the final video quality if reliably received. Generally, as the loss of the BL would cause a substantial degradation the BL should be transmitted in a well-protected channel. This can be achieved by, for example, using excessive channel coding.

However, in the MPEG-4 FGS coder, while the BL is generated by MCP, which effectively exploits the temporal correlation between video frames, there is no MCP for the FGS EL. This structure provides an inherent robustness against channel impairments and completely avoids error propagation (drifting effects) due to corruption of the EL. However, due to the lack of MCP in the EL, conventional FGS coding suffers from reduced compression efficiency. To fix this shortcoming, several algorithms have been proposed in the literature [51, 98, 99]. In particular, motion-compensated FGS (MC-FGS) coding was proposed in [51] which re-introduces an MCP loop in the EL (FGS layer) by using the motion vectors (MVs) and prediction modes from the BL. Fig. 2.1 illustrates the general framework for a motion compensated FGS (MC-FGS) coder.

As illustrated in the figure, an MC-FGS coder subdivides the single EL into two portions: EL-MCP and EL-Extra. The information below the MCP point (EL-MCP) is used for the prediction of the next EL frame while the upper portion of the EL (EL-Extra) is used solely for enhancing the video quality without being used for MCP. In the rest of the Chapter, we follow the terminology used in [1] and refer to the information used for MCP as EL-MCP while the EL portion used for enhancing the video quality only as EL-Extra. In general, the higher the MCP point, the greater is the compression efficiency and the lesser the error resilience to channel errors. Hence, there is a tradeoff associated with the selection of the MCP point.

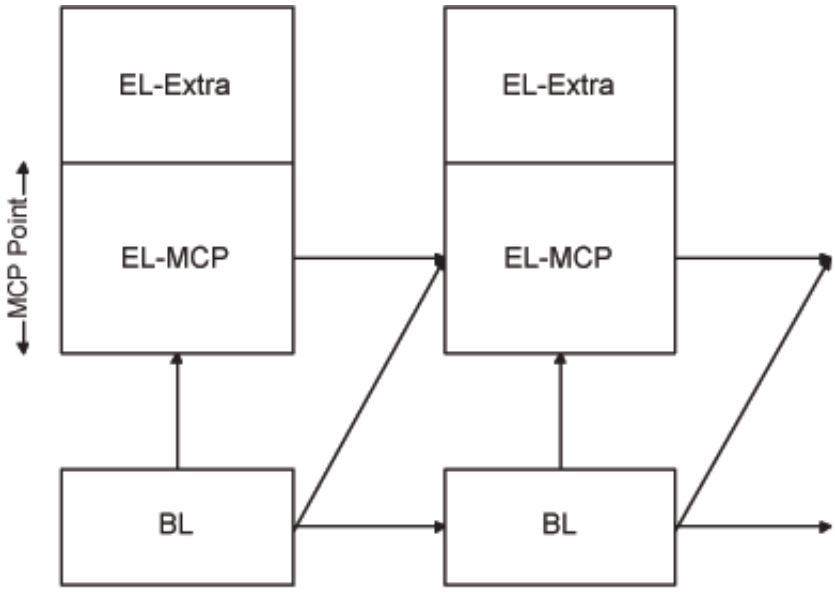

Figure 2.1: A Motion-Compensated FGS hybrid coder [1].

2.2.2 The Unequal Loss Protection (ULP) Scheme

As discussed above, scalable video bitstreams, such as MC-FGS, are highly structured and prioritized. However, the current Internet provides only a simple first-in-first-out (FIFO) queuing/scheduling policy regardless of the priority and content of packets. So it is important to convert a scalable, prioritized bitstream into an arbitrary number of equally important packets that can be decoded independently so that it could better "match" the Internet. An unequal loss protection (ULP) scheme is first described in [2], and later extended in [73], which assigns different amounts of protection to the different layers or parts of a bitstream based on their importance. In this way, a scalable, prioritized bitstream can be converted into an arbitrary number of packets with approximately equal importance [2, 73], which is considered to be better matched to the best-effort transport properties of current networks. Using this approach, the reconstructed quality of the source only depends on the actual number of received packets and not the relative positions of the losses.

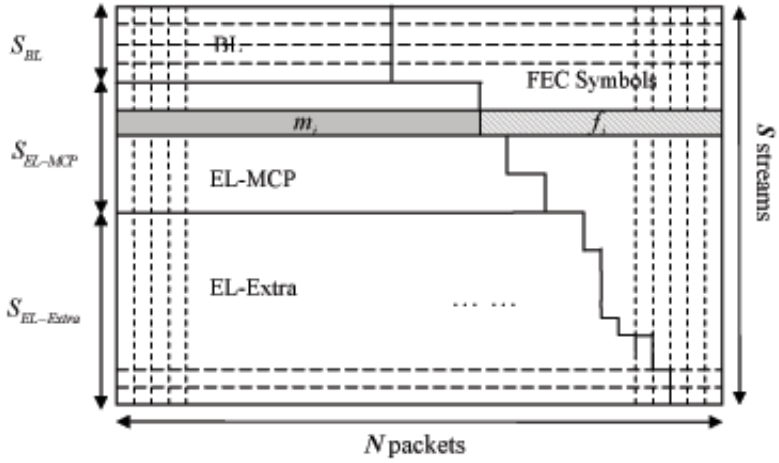

Figure 2.2: The Unequal Loss Protection (ULP) approach [2].

The ULP scheme for the MC-FGS video bitstream is illustrated in Fig. 2.2. Here, the encoded information associated with a video frame is arranged in an array with S rows and N columns. Each row is defined as a stream and each column represents the payload of a packet. The MC-FGS bitstream is then represented by S streams, with each stream consisting of one symbol from each of N packets and each stream constituting a channel coding block. As shown in Fig. 2.2, the symbols to the left of the boundary are information symbols, those to the right of the boundary are parity symbols. In each stream i, $1 \leq i \leq S$, there are m_i information symbols and $f_i = N - m_i$ parity FEC symbols. Therefore, with an $RS(N, m_i)$ FEC code, up to f_i packet losses can be recovered in the $i'th$ stream. These streams can be further classified into S_{BL} streams carrying BL information, S_{EL-MCP} streams carrying EL-MCP information and $S_{EL-Extra}$ streams carrying EL-Extra information. Because the data in the ELs are progressively encoded, we assume the number of parity symbols f_i is nonincreasing with stream index i. More specifically, we assume that $f_{i+1} \leq f_i$

for $i = 1, 2, ..., S-1$ [1]. Under this assumption, it follows that if stream i can be decoded then all streams j, for $j \leq i$ can be successfully decoded. The progressive encoding scheme is then reduced to a non-hierachical coding scheme where each of the N packets provides an independent description of the video frame.

Let the vector $\mathbf{f} = (f_1, f_2, ..., f_S)$ represent the FEC redundancy assignment across streams. If n, $1 \leq n \leq N$, packets are received, it follows that all the streams with $m_i \leq n$ can be successfully decoded. For a given $\mathbf{f}$ and i, we define $M(i, \mathbf{f}) = \sum_{j=1}^{i} m_j$ as the number of information symbols that can be recovered if the i'th stream can be successfully decoded. In this ULP scheme, the quality of a reconstructed frame depends only on the number of packets received.

The FEC parity vector $\mathbf{f} = (f_1, f_2, ..., f_S)$ can be optimized based on the channel state to minimize the expected video distortion. The average length of the BL is used in the optimization. Since the length of the BL can vary frame-by-frame, S_{BL} is adjusted while f_{BL}, S_{EL-MCP} and $S_{EL-Extra}$ remain fixed. Appropriate optimization methods are presented in [2, 73, 77, 100]. In this Chapter, we follow the optimal allocation framework proposed in [2].

2.3 Delay-constrained Optimal Packet Interleaving

2.3.1 Packet Interleaver

In this Chapter, frame l, $l = 1, 2, ..., L$, where L denotes the length of this video sequence, is encoded by the MC-FGS video coder and then transcoded into N packets with the ULP scheme. Let $P_{l,k}$ denote the k'th packet of frame l, where $k = 1, 2, ..., N$, and T_F denote the frame period. Because the encoding and decoding time is hardware-

[1] Actually, $f_i = f_{BL}$ for $i = 1, 2, ..., S_{BL}$ in the approach considered here, although we will describe the problem more generally.

dependent, and typically very small compared to the delay caused by interleaving and network transport, it is neglected in this Chapter. Hence, at time $(l-1)T_F$, packet $P_{l,k}$ is ready to be buffered into the interleaver. With an end-to-end delay constraint D, frame l has to be rendered at the playout deadline $T^d_{l,k} = (l-1)T_F + D$, and the packets arriving at the receiver after this playout deadline are considered useless.

Before being transmitted to the network, these packets are buffered into a block interleaver column-by-column. As soon as the interleaver is filled up, these packets are sent to the network row-by-row with a constant packet sending rate. The size of this interleaver is $N \times M$ with M defined as the interleaving depth, *i.e.*, the number of packet transmissions separating successive packets in a video frame.

Consider packet $P_{l,k}$ where $l = nM+i$, with $n = \lfloor (l-1)/M \rfloor$ and $i = l - nM$. The contents of the interleaver when it contains $P_{l,k}$ is then as illustrated in Fig. 2.3, where each column represents an entire video frame of N packets. Since packets are sent to the network at a constant rate, the time between any two successive transmitted packets is $\Delta = T_F/N$ so that the interleaver drain rate is the same as the fill rate[2]. The sending time of packet $P_{l,k}$ is then $T^s_{l,k} = nMT_F + (M-1)T_F + [(k-1)M + (i-1)]\Delta$, with n and i as defined above.

2.3.2 Network Model And Analysis

In this Chapter, the network is modeled as a packet-loss channel with random delays. The Gilbert model [101] is employed to capture the packet-loss process. As Fig. 2.4 indicates, a Gilbert model is a first-order, discrete-time, stationary, Markov chain with two states, where state 0 indicates a packet is successively received at the

[2]In the proposed scheme we assume there are dual buffers so that one can be filled while the other is being drained. This avoids the latency in waiting until the buffer is completely empty before it can be refilled.

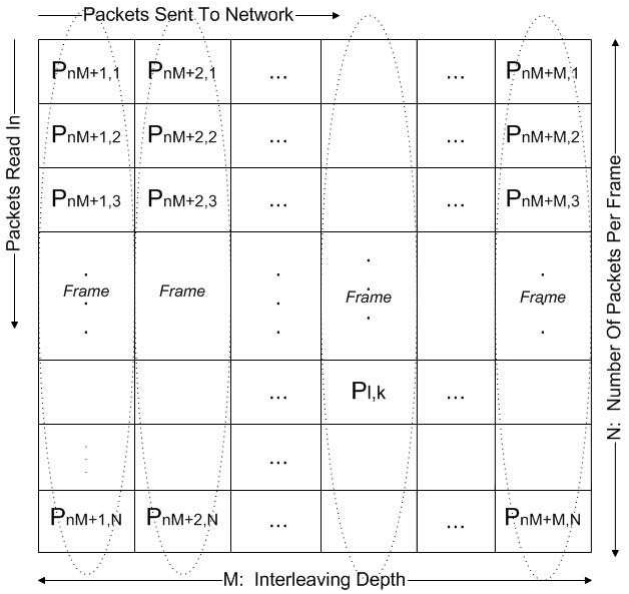

Figure 2.3: A block interleaver containing packet $P_{l,k}$.

receiver while state 1 indicates a packet loss. The parameters p and q denote, respectively, the transition probabilities from reception to loss and from loss to reception states. The Average Packet Loss rate (PL) and average Burst Length (ABL) can be estimated using control packets sent from receiver to sender, so these two transition probabilities can be evaluated according to:

$$p = \frac{PL}{1 - PL}\frac{1}{ABL}\ ;\ q = \frac{1}{ABL}\ . \tag{2.1}$$

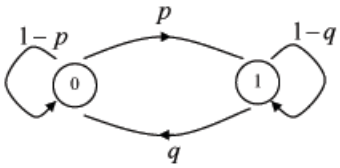

Figure 2.4: Gilbert loss model.

Because the packets in the interleaver are sent to the network row-by-row, the probability of losing or receiving the packet $P_{l,k+1}$ given reception or loss of the packet

$P_{l,k}$, for $1 \leq k \leq N-1$, can be determined from the M-step transition probability matrix associated with the underlying Gilbert model and can be computed as [102]:

$$p_M = p \cdot \frac{1-(1-p-q)^M}{p+q}, \tag{2.2}$$

and

$$q_M = q \cdot \frac{1-(1-p-q)^M}{p+q}, \tag{2.3}$$

where p_M denotes the probability of losing the packet $P_{l,k+1}$ given the packet $P_{l,k}$ is received, and likewise q_M denotes the probability of receiving the packet $P_{l,k+1}$ given the packet $P_{l,k}$ is lost.

The successfully received packets also experience random transport delay. We assume this delay process is *i.i.d.* and independent of the packet loss process. The cumulative distribution function (*cdf*) of the transport delay is denoted as $F(\tau|received)$, which is the probability that the transport delay is no larger than τ given that the packet can be delivered successfully. Hence, the probability that the packet $P_{l,k}$ arrives at the receiver before it's playout deadline given it is not lost by the network is $F(T^d_{l,k} - T^s_{l,k}|received)$.

Given the interleaving depth M, let $p_l(n|M)$ denote the probability that n packets out of the N packets associated with any frame l are either lost by the network or arrive after their playout deadline, where $n = 0, 1, ..., N$. To compute $p_l(n|M)$, we define an N-dimensional vector random process $\mathbf{I}_l$ as follows: for each frame l, $\mathbf{I}_l$ is an N-dimensional random vector $\mathbf{I}_l = (I_{l,1}, I_{l,2}, ..., I_{l,N})$ where $I_{l,k}$, $1 \leq k \leq N$, is the indicator function of the availability of packet $P_{l,k}$ by its playout deadline, *i.e.*,

$$I_{l,k} = \begin{cases} 0 & \text{if } P_{l,k} \text{ arrives before } T^d_{l,k}, \\ 1 & \text{otherwise.} \end{cases} \tag{2.4}$$

Also, define the set $\Omega = \{0, 1\}$ and $\mathscr{F} = \Omega^N$, *i.e.*, for each $\mathbf{w} \in \mathscr{F}$, $\mathbf{w}$ is an

N-tuple $\mathbf{w} = (w_1, w_2, ..., w_N)$, where $w_i = 0$ or 1, $1 \leq i \leq N$. Hence, $p_l(n|M)$ can be computed as

$$p_l(n|M) = \sum_{\mathbf{w} \in \mathscr{D}_n} Prob\{\mathbf{I}_l = \mathbf{w}\}, \tag{2.5}$$

where $\mathscr{D}_n = \{\mathbf{w} : \mathbf{w} \in \mathscr{F}, and \sum_{i=1}^{N} w_i = n\}$. To compute $Prob\{\mathbf{I}_l = \mathbf{w}\}$, for each packet $P_{l,k}$ we define the loss indicator function $LI_{l,k}$ as

$$LI_{l,k} = \begin{cases} 0 & \text{if } P_{l,k} \text{ is received,} \\ 1 & \text{if } P_{l,k} \text{ is lost by the network.} \end{cases} \tag{2.6}$$

Hence, the N-dimensional random vector $\mathbf{LI}_l = (LI_{l,1},\ LI_{l,2},\ ...,\ LI_{l,N})$ represents the packet loss process associated with the frame l. Furthermore, we employ an N-dimensional random vector $\mathbf{DI}_l = (DI_{l,1},\ DI_{l,2},\ ...,\ DI_{l,N})$ to represent the packet delay process in the loss-free case, and the corresponding delay indicator function $DI_{l,k}$ is defined as

$$DI_{l,k} = \begin{cases} 0 & \text{if } P_{l,k} \text{ has been received before } T^d_{l,k}, \\ 1 & \text{if } P_{l,k} \text{ has not been received before } T^d_{l,k}. \end{cases} \tag{2.7}$$

Therefore, we have

$$Prob\{\mathbf{I}_{l,k} = \mathbf{w}\} = \sum_{\mathbf{u} \bigvee \mathbf{v} = \mathbf{w}} Prob\{\bigcap_{k=1}^{N} (LI_{l,k} = u_k, DI_{l,k} = v_k)\}\ , \tag{2.8}$$

where $\mathbf{u} = (u_1, u_2, ..., u_N) \in \mathscr{F}$ and $\mathbf{v} = (v_1, v_2, ..., v_N) \in \mathscr{F}$. The "$\bigvee$" represents the logical OR operation which follows from the fact that $I_{l,k} = 1$ if $LI_{l,k} = 1$ or $DI_{l,k} = 1$. With the assumption that the packet loss process and the packet delay process are independent, we have then

$$\begin{aligned} & Prob\{\bigcap_{k=1}^{N} (LI_{l,k} = u_k, DI_{l,k} = v_k)\} \\ = & Prob\{\bigcap_{k=1}^{N} (LI_{l,k} = u_k)\} \cdot Prob\{\bigcap_{k=1}^{N} (DI_{l,k} = v_k)\}\ . \end{aligned} \tag{2.9}$$

The quantity $Prob\{\bigcap_{k=1}^{N}(LI_{l,k} = u_k)\}$ can be computed from the Gilbert loss model parameters as follows:

$$lProb\{\bigcap_{k=1}^{N}(LI_{l,k} = u_k)\} = Prob\{LI_{l,1} = u_1\} \cdot \prod_{k=1}^{N-1} P_{u_k,u_{k+1}}\,, \tag{2.10}$$

where $P_{u_k,u_{k+1}}$ denotes the corresponding element of the M-step transition probability matrix

$$\mathbf{P^{(M)}} = \begin{pmatrix} 1-p_M & p_M \\ q_M & 1-q_M \end{pmatrix}, \tag{2.11}$$

where p_M and q_M are given by equation (2.2) and (2.3), respectively.

Using the assumption that the packet delay process is *i.i.d.*, $Prob\{\bigcap_{k=1}^{N}(DI_{l,k} = v_k)\}$ can be computed as follows:

$$\begin{aligned} &Prob\{\bigcap_{k=1}^{N}(DI_{l,k} = v_k)\} \\ &= \prod_{k=1}^{N} \big(v_k(1 - F(T^d_{l,k} - T^s_{l,k}|received)) \\ &\quad +(1-v_k)F(T^d_{l,k} - T^s_{l,k}|received)\big)\,. \end{aligned} \tag{2.12}$$

We define the average probability that n packets out of the N packets associated with an arbitrary frame are lost or arrive after their playout deadline given the interleaving depth is M, as $p(n|M)$. The quantity $p(n|M)$ can be computed from:

$$p(n|M) = \frac{1}{L}\sum_{l=1}^{L} p_l(n|M),\ \ where\ n = 0, 1, ..., N. \tag{2.13}$$

2.3.3 Joint Optimization of the Interleaver, the MCP Video Source Coder, and the ULP Allocation

Following the methodology presented in [2], we define the *cdf* $F(x|M)$ as:

$$F(x|M) = \sum_{n=0}^{x} p(n|M);\ \ x = 0, 1, 2, ..., N. \tag{2.14}$$

The quantity $F(f_i|M)$ then is the average probability that no more than f_i packets are lost out of N packets of a frame given the interleaving depth M. Because there are f_i FEC symbols in stream i, stream i can be decoded with the probability $F(f_i|M)$. Hence, the total number of decodable symbols is $M(i, \mathbf{f})$. Let $PSNR_{MCP}[M(i, \mathbf{f})]$ denote the mean PSNR value of the reconstructed video if a total of $M(i, \mathbf{f})$ symbols are decoded given the MCP point is set at MCP. The incremental PSNR if stream i can be decoded is defined as:

$$g_i(\mathbf{f}, MCP) = PSNR_{MCP}[M(i, \mathbf{f})] - PSNR_{MCP}[M(i-1, \mathbf{f})]. \tag{2.15}$$

The quantity $g_i(\mathbf{f}, MCP)$ is the amount by which the PSNR increases when the receiver decodes stream i, given that all streams prior to i have already been correctly decoded. Recall, MCP is the bit count of the EL-MCP for each frame. We set $g_1(\mathbf{f}, MCP)$ to be the difference in PSNR between the case in which the first stream is received and the case in which no information is received.

Given M, $\mathbf{f}$,and MCP, the expected PSNR value can then be found as

$$E[PSNR(M, \mathbf{f}, MCP)] = \sum_{i=1}^{S} F(f_i|M) \cdot g_i(\mathbf{f}, MCP). \tag{2.16}$$

We can jointly optimize M, MCP, and UEP to maximize the expected video PSNR by solving for

$$M^*, MCP^*, \mathbf{f}^* = argmax_{M,MCP,\mathbf{f}} E[PSNR(M, MCP, \mathbf{f})]. \tag{2.17}$$

Since there is no closed-form solution to (2.17), numerical methods are used in this Chapter. The UEP allocation is optimized by the local search algorithm proposed in [2]. The interleaving depth and MCP point selection are optimized by global search.

2.4 Simulations

2.4.1 Simulation Setup

We used an H.26L-MC-FGS codec, comprised of an H.264 TML9 base-layer codec and an enhancement-layer codec employing the MC-FGS scheme. The test video sequences used are the QCIF "Foreman" and "Susie" sequences, which are encoded at 30 fps, with GOP structure I-P-P-P..., and a GOP length of 30. The quantization parameter is set to 27. For the Foreman sequence, the average rate of the BL is 42.84 kbps and the mean PSNR of the BL is 29.16 dB. For the Susie sequence, the average rate of the BL is 32.97 kbps and the mean PSNR of the BL is 30.92 dB. Figures 2.5 and 2.6 show the corresponding Rate-PSNR functions with 8 different MCP points, chosen as 2,400, 4,800, 7,200, 9,600, 12,000, 14,400, 16,800, and 19,200 bits. These curves are measured by decoding the entire BL and the EL bitstreams truncated at the different points as indicated. From Fig.'s 2.5 and 2.6, it can be seen that if the number of decodable EL bits is less than the MCP point, the PSNR drops rapidly because of the drift effect; otherwise, the PSNR rises more slowly because the lost bits in the Extra-EL do not affect the following frame and there is no drift.

In the simulations, the number of packets that are generated within each frame period is $N = 16$. We used RS codes for error protection, and there is 1 byte per RS symbol. The average payload length of packets is 150 bytes. The packet delay is modeled as a right-shifted Gamma distribution whose probability density function (pdf) is:

$$f(d|received) = \frac{\alpha}{\Gamma(n)}(\alpha(d-\kappa))^{n-1}e^{-\alpha(d-\kappa)}, \tag{2.18}$$

where $\alpha = 33.33,\ n = 3,\ \kappa = 0.01$, yielding a mean packet delay of 0.1 second. The cdf of the transport delay can be found as $F(\tau|received) = \int_0^\tau f(t|received)dt$.

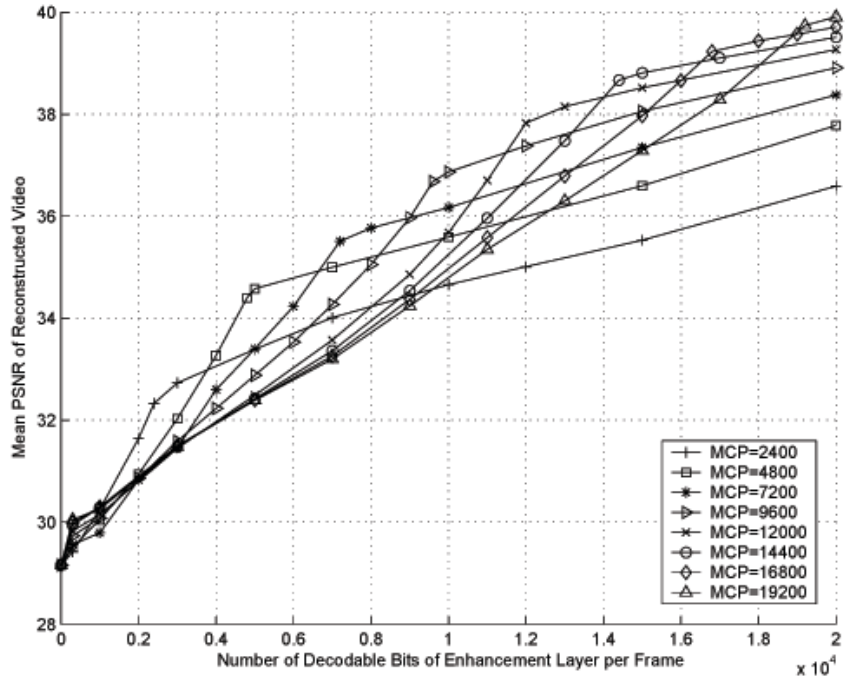

Figure 2.5: Rate-PSNR functions of the QCIF "Foreman" video sequence.

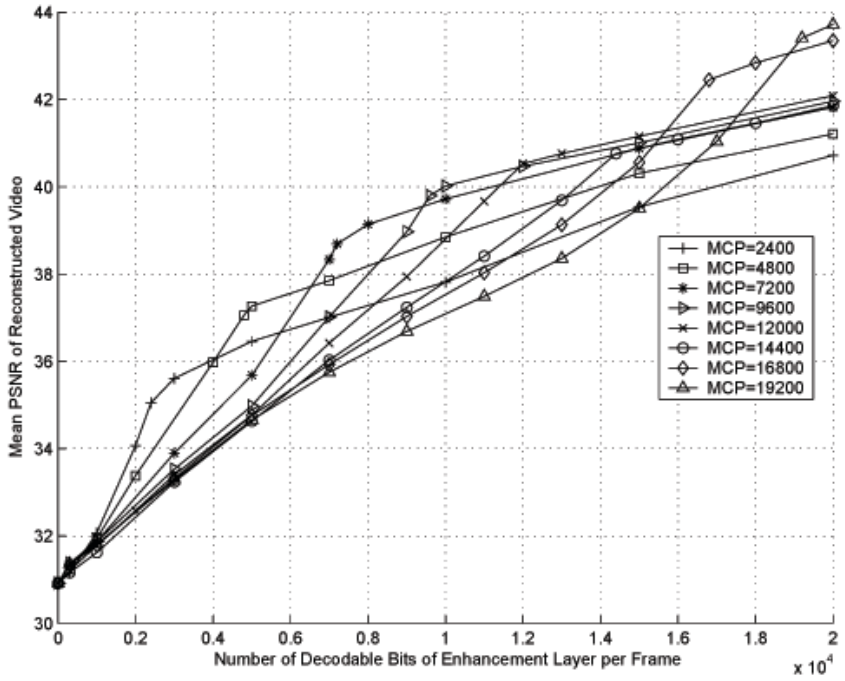

Figure 2.6: Rate-PSNR functions of the QCIF "Susie" video sequence.

2.4.2 Simulation Results and Discussion

We compare the performance of the proposed transport system with that of a comparison system without interleaving under different average burst lengths and delay constraints. In all of these simulations, PL is fixed at 10%. The ABL is varied from 1.11 to 5 and the delay constraint D is varied from 0.2 sec to 0.35 sec. Because the ABL of the Bernoulli loss process with $PL = 10\%$ is 1.11, the Gilbert loss model actually reduces to the Bernoulli loss model when ABL is 1.11. In this case, the interleaving scheme provides no advantage. Hence, as expected, the optimal interleaving depth M is chosen as 1 in this case. As shown in Fig.'s 2.7 and 2.8, the solid surface illustrates the performance of the proposed system with optimal interleaving and the dashed surface depicts the performance of the comparison system without interleaving. The solid surface is always above the dashed surface, which indicates, as expected, that the proposed optimized system is superior to the comparison system without interleaving. It can be observed that the improvement in expected PSNR achieved by this proposed system increases as the average burst length of network losses increases or the delay constraint is relaxed. In the case of $BL = 5$ and $D = 0.35$ sec, the proposed system outperforms the comparison system by over 1.3 dB for the QCIF Foreman video and by 1.2 dB for the QCIF Susie video in expected PSNR.

More quantitative analysis for the Foreman sequence is presented in Fig.'s 2.9 and 2.10. Each of these figures has two parts: Part (a), compares the system performances in expected PSNR and Part (b) indicates the optimal interleaving depth. The system without interleaving corresponds to the case of $M = 1$, which is also plotted in Part (b). In Fig 2.9, the ABL is fixed at 3 and the delay constraint D is varied from 0.2 to 0.35 sec. When D is 0.2 sec, the optimal interleaving depth is indicated as 1

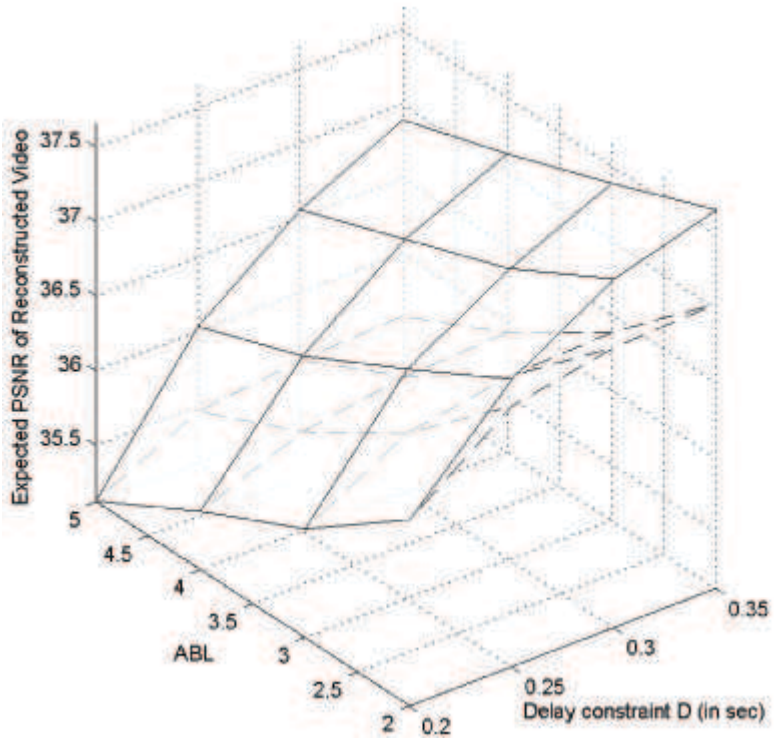

Figure 2.7: Performance comparison for the QCIF "Foreman" video sequence.

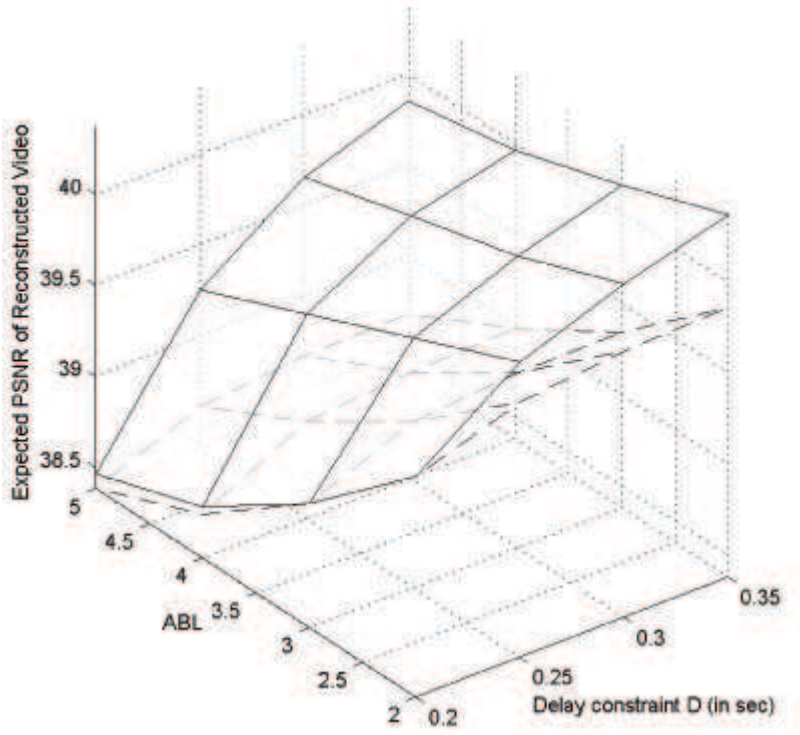

Figure 2.8: Performance comparison for the QCIF "Susie" video sequence.

because the delay constraint is too stringent, relative to the average network delay of 0.1 sec, and all packets need to be forwarded to the network as soon as possible. As the delay constraint is relaxed, larger interleaving depths can be used because the losses due to late arrivals decrease compared with the case of a more stringent delay constraint. For example, when D is relaxed to $D = 0.25$ sec, the optimum interleaving depth increases to $M = 2$ and the sending time of the first packet in the interleaver is then postponed for $T_F = 0.033$ sec. When D is further relaxed by an additional 0.05 sec from 0.25 to 0.30, the optimum interleaving depth increases from 2 to 3 without significant effects due to late arrivals. However, when D is relaxed even further from 0.30 sec to 0.35 sec, the optimal interleaving depth remains at 3. This is because, for this choice of parameters, the corresponding latency associated with increasing M would have resulted in a significant increase in dropped packets due to missed playout deadlines. Hence, by optimization, M is kept at 3 when D increases from 0.3 to 0.35. For further increases in the allowable delay, D, the optimum value of M would, of course, increase.

In Fig 2.10, D is fixed at 0.3 sec and the ABL is varied from 1.11 to 5. As indicated previously, the case of $ABL = 1.11$ corresponds to the Bernoulli loss process since $PL = 10\%$. The interleaving scheme is useless in the Bernoulli loss case, so the optimum M is 1. Larger interleaving depth can cause more losses due to late arrivals; however, compared with the benefit of reducing video distortion caused by the bursty losses, larger interleaving depth can be used as the ABL increases. For example, the optimum M is 3 when ABL is 3. But when ABL is increased 4 to 5, the optimal interleaving depth remains at 3 because of the relatively stringent delay constraint.

Fig. 2.11 illustrates a trace corresponding to $ABL = 3$, $PL = 10\%$ and $D = 0.3$ sec. In this case, the optimal interleaving depth is 3. Part (a) compares the

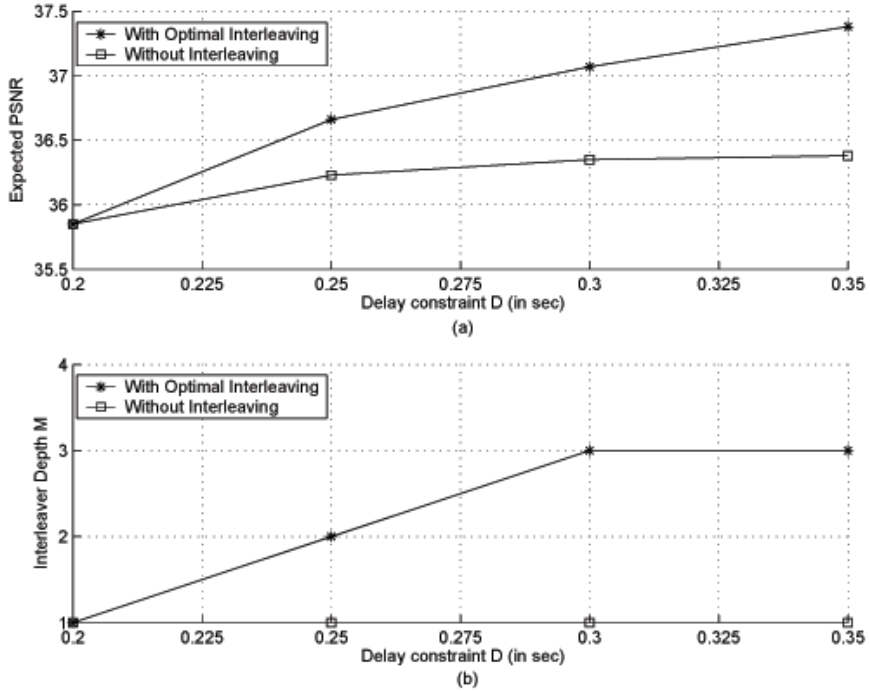

Figure 2.9: System performance comparison under different delay constraints D with fixed $ABL = 3$. (a) PSNR *vs.* D, (b) Interleaving depth *vs.* D.

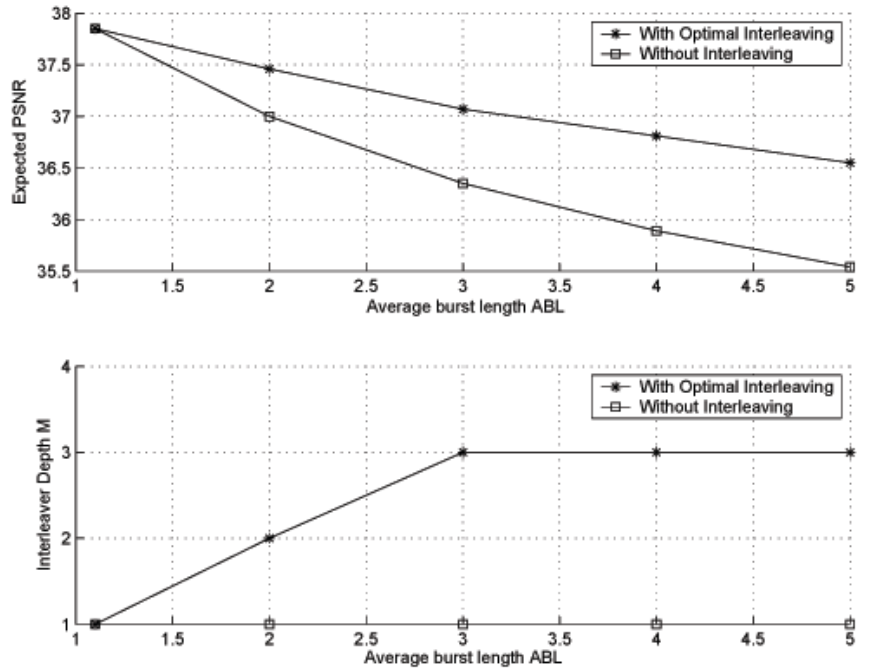

Figure 2.10: System performance comparison under different ABL with fixed $D = 0.3sec$. (a) PSNR *vs.* ABL, (b) Interleaving depth *vs.* ABL.

frame-by-frame PSNR of the Foreman sequence with optimal interleaving and without interleaving. The proposed transport system outperforms the comparison system by 1.9 dB in mean PSNR. Part (b) shows the number of received packets on a frame-by-frame basis. There are 223 packets lost if the optimal interleaving scheme is employed, but only 188 packets are lost without interleaving. We can see that even if more packets are lost, the performance of the proposed system is still superior to that of the system without interleaving because the optimal interleaving scheme can effectively reduce the video distortion cause by bursty losses.

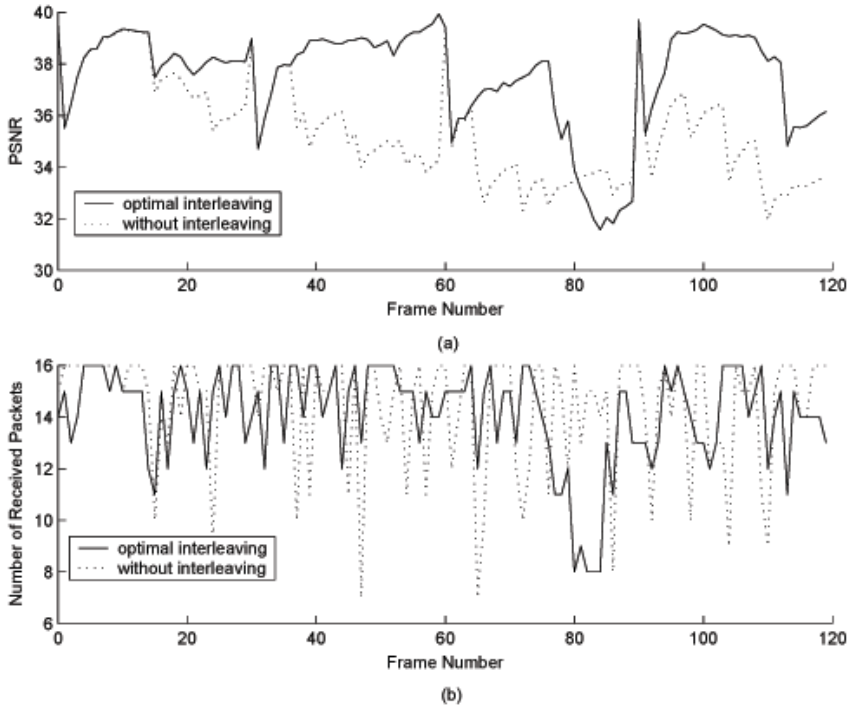

Figure 2.11: System performance comparison with $ABL = 3,\ D = 0.3sec.$ (a) PSNR *vs.* frame number, (b) number of received packets *vs.* frame number.

2.5 Conclusion

In this Chapter, we proposed a video transport system which integrates MC-FGS video source coding, ULP and interleaving schemes. Under an end-to-end delay

constraint, the source coder, ULP allocation and interleaver are jointly optimized. The simulation results demonstrate that the improvement in mean PSNR achieved by this proposed system increases as the average burst length of the network losses increases or the delay constraint is relaxed. Furthermore, it is also shown that even if interleaving does lead to more packet losses due to late arrivals, the performance of the proposed transport system is still superior to the comparison system without interleaving. This is because optimal interleaving can effectively reduce the video distortion through mitigating the effects of the otherwise bursty losses.

CHAPTER 3

Network-Adaptive Transport of Motion-Compensated Fine Granularity Scalability Video Over Multiple Asymmetric Paths

3.1 Motivation

In the previous Chapter, we introduced an optimal interleaving scheme which can be used to combat bursty losses. In this Chapter, we will study a promising technique for real-time video transport, multiple path transport (MPT). As the term indicated, instead of using only one single transmission path, multiple independent paths are employed to transport the multiple bitstreams in MPT schemes. By taking advantage of the uncorrelated nature of the loss processes on these different network paths, the virtual average path exhibits a smaller variability in communication quality [70]. Recently, multiple description (MD) coding has emerged as a powerful technique in video transport, which is particularly efficient when combined with MPT. In MD coding, the source is encoded into multiple bitstreams (descriptions) of approximately equal importance such that each description can be decoded independently. Moreover, the loss of a particular description would not jeopardize the decoding process while the end-quality improves as the number of description received increases.

Unfortunately, most works utilizing MPT assume that the network characteristics of all the routing paths are similar or even identical. These statistically similar or identical paths are commonly referred to as symmetric paths. However, in the Internet, different routing paths generally exhibit different statistical behaviors. This kind of routing paths with different characteristics are generally referred to as asymmetric paths. In the Internet, different routing paths may suffer from different packet loss patterns. The different number of hops and different physical media for each individual path may also lead to different transport delays. It should be noted that the different characteristics of different transport paths not only have significant impacts on the traffic allocation. For example, one may want to put more traffic on the more reliable transport path. As well shall see, from our discussion and studies in this Chapter, it can also affect the construction of the MD codes.

Previous works on multiple description coding and delivery over asymmetric multiple paths include [87]: in which the authors proposed adjusting the frame rate for each description based on the available bandwidth of each individual path. Unfortunately, the author did not consider the packet loss characteristics and there is also no simple mechanism to incorporate such information into the proposed framework. In [88], Kim *et al.* studied the construction of MD codes for the asymmetric paths using the 3D set partitioning in hierarchical trees (3D-SPIHT) video coder [103]. However, while the proposed scheme took into consideration the packet loss probability of each individual path, the video coder, 3D-SPIHT, studied does not incorporate motion-compensated prediction (MCP), which is found in all the video coding standards due to its high compression efficiency. An MCP video coder combines differential coding along an estimated motion trajectory of the picture contents with intraframe-encoding of the prediction residues. Unfortunately, while an improved compression efficiency can

be achieved by MCP which effectively exploits the temporal redundancy, the use of MCP also renders the compressed video extremely sensitive to channel impairments. Sometimes a single bit error or erasure can cause precipitous degradation due to error propagation as a result of predictor-mismatch between the encoder and decoder. This is the so-called "drift" problem. Generally, if a higher fidelity reference frame is used to predict the subsequent frame, a higher compression efficiency can be achieved, at an expense of worse error resilience properties resulting in greater drops in the received quality when there are predictor mismatches. Hence, in order to achieve a better balance between the compression efficiency and error-resilience property, the amount of information passed to the encoder predictor should be carefully controlled.

In this Chapter, we study the transport of a motion-compensated compressed video bitstream using asymmetric paths over the Internet. In particular, we propose a unified approach incorporating adaptive motion-compensated prediction, multiple description coding and unequal multiple path allocation so as to improve both the robustness and error resilience property of the video coding and transmission system, while simultaneously improving the delivered video quality.

The rest of this Chapter is organized as follows. In Section 3.2, we briefly describe some technical preliminaries, including details of the video encoder used in this work, the n-channel motion-compensated symmetric MD coding scheme and the system model. In Section 3.3, we provide a description of the channel models we used, including both a Gilbert-Elliot model and an $M/D/1/K$ queue model. In Section 3.4, we address how to jointly optimize the video encoder, construction of MDs and traffic allocation based on the characteristics of multiple routing paths. The Section 3.5, we present some simulation results and discussion. Finally, Section 3.6 provides a summary and conclusion of the Chapter.

3.2 Preliminaries

3.2.1 MC-FGS Video Coder and n-channel Symmetric Motion-Compensated MD Coding

The MC-FGS coder introduced in Chapter 2 is also employed in this Chapter. As described previously, the MC-FGS video coder partitions the compressed bitstream into a base layer (BL) and an enhancement layer (EL). The BL carries the most important information, such as coding modes, motion vectors and low frequency DCT coefficients. The EL is progressively encoded by bit-plane technique so that it can be truncated at any point. The EL is divided into two portions. The first portion, denoted as EL-MCP, is used for the prediction of the subsequent frames through MCP. The bit count of this EL-MCP is defined as the MCP point. The second portion, denoted as EL-Extra, is used only for enhancing the video quality and is not used for predicting the subsequent frames. If the decoder receives the entire EL-MCP, no error propagation (drift) will occur. But if part of the EL-MCP is not received, drift will occur and errors will propagate until the next Intra-coded frame. Generally, the higher the MCP point is, the larger is the likelihood that drift occurs. On the other hand, as the MCP point increases, the coding efficiency will also increase since the quality of the reference picture improves. Hence, it is necessary for the encoder to trade off the coding efficacy against the risk of drift by properly choosing the MCP point based on the network conditions.

An n-channel motion-compensated MD coding approach was first proposed by Chan *et al.* [1] based on the FEC-based multiple description coding techniques in [73]. The particular technique utilized progressive source coding combined with unequal loss protection (ULP) using maximum distance separable (MDS) Reed-Solomon codes. Using appropriately chosen channel coding rates, the information of different

layers or bitplanes associated with a video frame are effectively represented by different descriptions. As a result, the quality of reconstructed video frame depends only on the number of the descriptions successively received. Since, MC-FGS is employed as the source coding stage, the motion compensation is seamlessly incorporated with the FEC-based multiple description coding. So, this scheme is called as n-channel motion-compensated MD coding.

For more details on MC-FGS and ULP including parameter notations, please see the preliminary section of Chapter 2.

3.2.2 System Model

Fig. 3.1 illustrates the end-to-end video transport system using L multiple routing paths. On the sender side, the video source generates frames with a constant frame rate. We define the frame period as T_F, which is time interval between any two successive frames. These frames are fed into the MC-FGS encoder which compresses each frame in to a non-scalable BL and a progressively coded EL. The compressed bitstream is transcoded into n descriptions by the FEC-MD scheme. Each description is packetized into a network packet. Then n descriptions are then sent through the L distinct routing paths to the destination. These L paths may have different packet loss patterns. Hence, besides optimization of the MCP rate and the ULP allocation, another important issue is how to optimally allocate these n descriptions into the L individual paths to minimize the end-to-end video distortion.

Define L-dimension vector $\mathbf{n} = (n_1, n_2, ..., n_l, ..., n_L)$ as the traffic allocation vector over the L paths, where n_l denotes the number of the descriptions sent through path l, $1 \leq l \leq L$. Obviously, $n = \sum_{l=1}^{l} n_l$, since n is the total number of descriptions. It should be noted that due to unequal allocation of these n descriptions into the L

paths, the time interval between adjacent packets or transmission rate in different paths would be different.

Based on states of network paths, the sender can jointly optimized the MCP point of MC-FGS source coder, the FEC parity vector $\mathbf{f} = (f_1, f_2, ..., f_S)$ and the description allocation vector $\mathbf{n} = (n_1, n_2, ..., n_L)$ to minimize the end-to-end video distortion.

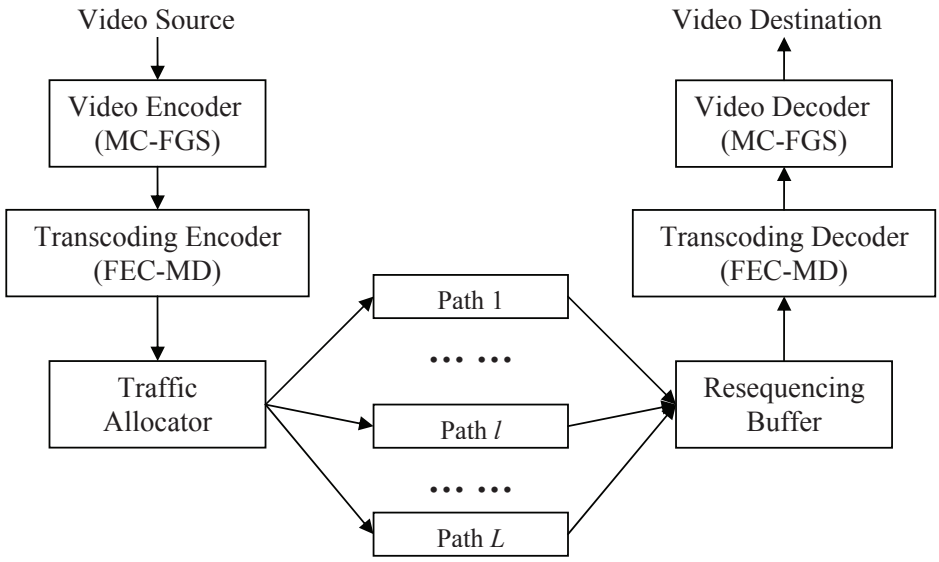

Figure 3.1: End-to-end System Architecture

On the receiver side, the packets correctly delivered are first buffered and rearranged according to their RTP packet sequence numbers. Then the FEC-MD decoder converts these descriptions into a MC-FGS bitstream. After decompression by the source decoder, these frames are ready to be rendered.

3.3 Network Model and Analysis

In this section, two channel models, Gilbert-Elliot model and $M/D/1/K$ queue model are described and analyzed.

3.3.1 Gilbert-Elliot model

Each individual path is modeled as a packet-loss channel using the discrete-time Gilbert-Elliot model [101]. Fig. 3.2 illustrates the Gilbert-Elliot model associated with path l, where $1 \leq l \leq L$. The two states of this model are denoted as G(good) and B(bad). In state G, packets are received successively, and in state B, packets are lost. The holding times of both the good state and bad state are assumed independent and distributed exponentially [104, 105]. The transition probabilities of path l from good state to bad state and from bad state to good state are denoted as p_l and q_l, respectively.

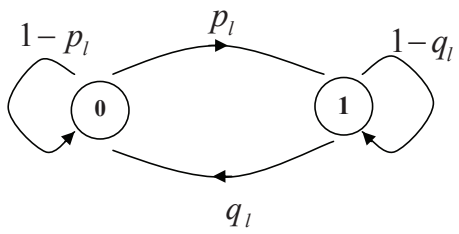

Figure 3.2: Gilbert-Elliott Loss Model

Assume that continuous time is divided into intervals with very short duration Δ and the average packet loss rate PL_l and average burst time ABT_l of path l can be obtained through feedback of control packets. The transition probabilities between any two successive time intervals for path l can be obtained by:

$$p_l = \frac{PL_l}{1 - PL_l} \frac{1}{ABT_l} \; ; \; q_l = \frac{1}{ABT_l} \, . \tag{3.1}$$

Assume there are n_l packets sent along path l with a constant rate within each frame period. Then the packet sending period in path l is $T_{p,l} = T_F/n_l$, where T_F denotes the frame period. So, every packet is sent at the $t_l = T_{p,l}/\Delta$'th time interval after the previous one. As a result, the transition probability of reception or loss of a

packet given reception or loss of the packet t_l time intervals before, can be computed as t_l-step transition probability from:

$$p_{t,l} = p_l \cdot \frac{1-(1-p_l-q_l)^{t_l}}{p_l+q_l}, \tag{3.2}$$

and

$$q_{t,l} = q_l \cdot \frac{1-(1-p_l-q_l)^{t_l}}{p_l+q_l}. \tag{3.3}$$

The limiting form of (3.2) for small Δ can be obtained by substituting from (3.1) with the result:

$$\begin{aligned} p_{t,l} &= PL_l[1-(1-\frac{\Delta}{(1-PL_l)ABT_l})^{\frac{T_{P,l}}{\Delta}}] \\ &\rightarrow PL_l[1-e^{-\frac{1}{1-PL_l}\frac{T_{P,l}}{ABT_l}}] \; as \; \Delta \rightarrow 0, \end{aligned} \tag{3.4}$$

and similarly,

$$p_{t,l} \rightarrow (1-PL_l)[1-e^{-\frac{1}{1-PL_l}\frac{T_{P,l}}{ABT_l}}] \; as \; \Delta \rightarrow 0. \tag{3.5}$$

Given these limiting transition probabilities, the probability that a particular number of packets are lost out of any given number of consecutive packets transmitted through path l can be derived using the method presented in [106]. For completeness, it is repeated here. Let $g_l(\nu)$ denote the probability of receiving $\nu-1$ packets consecutively on path l, and $G_l(\nu)$ denote the probability of receiving more than $\nu-1$ video packets consecutively on path l. We can then obtain:

$$g_l(\nu) = \begin{cases} 1-q_{t,l}, & \nu = 0 \\ q_{t,l} \cdot (1-p_{t,l})^{\nu-2} \cdot p_{t,l}, & \nu > 1\,, \end{cases} \tag{3.6}$$

and

$$G_l(\nu) = \begin{cases} 1, & \nu = 1 \\ q_{t,l} \cdot (1-p_{t,l})^{\nu-2}, & \nu > 1\,, \end{cases} \tag{3.7}$$

Let $R_l(i,m)$ be the probability of $i-1$ packet losses within the next $m-1$ packets following a lost packet on path l. It can be found recursively according to:

$$R_l(i,m) = \begin{cases} G_l(m) & i = 1 \\ \sum_{\nu=1}^{m-i+1} g_l(\nu) \cdot R_l(i-1, m-\nu) & 2 \le i \le m\,. \end{cases} \tag{3.8}$$

Then the PMF that indicates the probabilities of j lost packets out of n_l packets on path l , $p_l(j,n_l)$, is:

$$p_l(j,n_l) = \begin{cases} \sum_{\nu=1}^{n_l-j_l+1} PL_l \cdot G_l(\nu) \cdot R_l(j-1, n_l-\nu+1) & 1 \le j \le n_l \\ 1 - sum_{j=1}^{n_l} p_l(j,n_l) & j = 0\,. \end{cases} \tag{3.9}$$

Following the notation in [106], $p_l(j,n_l)$ is denoted as the Block Error Density Function (BEDF) of the path l.

3.3.2 M/D/1/K queue

As also described in the previous section, the length of packets generated with the n-channel motion-compensated MD coding are all fixed at the same size. Thus, the $M/D/1/K$ queue is an appropriate model to represent network paths. Fig. 3.3 represents an $M/D/1/K$ queue modeling network path l. As depicted, the packet arrival process is assumed a Poisson process with rate λ_l, and the system can hold up to K_l packets. The service rate is a deterministic value μ_l and the service time is $D_l = 1/\mu_l$ time units per packet. Then, the network load of path l is $\rho_l = \lambda_l/\mu_l$. We assume a FIFO service policy is employed in this queueing system.

Appendix B presents a general approach to evaluate $P(j,n)$, the probability of losing j packets in a block of n packets, for any $M/D/1/K$ queue with known K and ρ. Using this approach, the corresponding block error density function $p_l(j,n_l)$ for

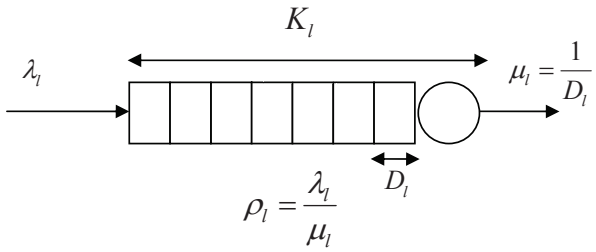

Figure 3.3: $M/D/1/K$ queue system

path l can be easily computed.

3.4 Cross-Layer Optimization: MCP Selection, ULP and Traffic Allocation Choice

As discussed above, we assume that L independent routing paths are employed in this video transport system. Let $p_l(j, n_l)$ denote the PMF representing the probability that j packets are lost out of n_l packets sent along path l, where $0 \leq j \leq n_l$ and $\sum_{l=1}^{L} n_l = n$. By convolving the PMFs of all the paths, we can compute the PMF $p_n(m|\mathbf{n})$, $m = 0, 1, 2, ..., n$, which is the probability that m packets are lost out of the n packets of each frame given the traffic allocation vector $\mathbf{n} = (n_1, n_2, ..., n_L)$. Specifically,

$$p_n(m|\mathbf{n}) = p_1(m, n_1) * p_2(m, n_2) * ... * p_L(m, n_L). \tag{3.10}$$

Following the methodology of [2], the corresponding cumulative distribution function is defined as:

$$F_n(x|\mathbf{n}) = \sum_{m=0}^{x} p_n(m|\mathbf{n}); \; x = 0, 1, 2, ..., n. \tag{3.11}$$

Because there are f_i FEC symbols in stream i, the stream i can be correctly decoded with probability $F_n(f_i|\mathbf{n})$. Recall here, there are $M(i, \mathbf{f}) = \sum_{j=1}^{i} m_j$ information symbols that can be recovered if the i'th stream can be successfully decoded.

Let $PSNR_{MCP}[M(i,\mathbf{f})]$ denote the mean PSNR value of the video if a total of $M(i,\mathbf{f})$ symbols are decoded given the MCP point MCP. The PSNR increment if stream i can be decoded is defined as:

$$g_i(\mathbf{f}, MCP) = PSNR_{MCP}[M(i,\mathbf{f})] - PSNR_{MCP}[M(i-1,\mathbf{f})]. \tag{3.12}$$

The quantity $g_i(\mathbf{f}, MCP)$ is the amount by which the PSNR increases when the receiver decodes stream i, given that all streams prior to i have already been correctly decoded.

Given the FEC parity vector $\mathbf{f}$, the traffic allocation vector $\mathbf{n}$ and the MCP point, the expected PSNR value can be found:

$$E[PSNR(MCP,\mathbf{f},\mathbf{n})] = \sum_{i=1}^{S} F_n(f_i|\mathbf{n}) \cdot g_i(\mathbf{f}, MCP) \tag{3.13}$$

So, we can jointly optimize the MCP point, the FEC parity assignment vector and the traffic allocation vector to maximize the expected video PSNR with the result

$$MCP^*, \mathbf{f}^*, \mathbf{n}^* = argmax_{MCP,\mathbf{f},\mathbf{n}} E[PSNR(MCP,\mathbf{f},\mathbf{n})] \tag{3.14}$$

Since there is no closed-form solution to (3.14), numerical methods are used to solve this optimization problem in this Chapter. The ULP allocation is optimized by the local search algorithm proposed in [2]. MCP point selection and traffic allocation are optimized by global search.

3.5 Simulations

3.5.1 Simulation setup

We used an H.26L-MC-FGS codec, comprised of an H.264 TML9 base layer codec and an enhancement layer codec with the MC-FGS scheme. The test sequence is the

QCIF "Foreman" video encoded at 10 frames per second, with GOP structure I-P-P-P..., and a GOP length 30. The quantization parameters for the BL and ELs are set to QP=27. The average rate of the BL is 23.05 kbps and the mean PSNR of the BL is 28.92 dB. Typical PSNR results vs. the MCP point are illustrated in Fig. 3.4. These curves are measured by decoding the entire BL and a part of EL truncated at different points. The horizontal axis represents the decodable bit count of the EL. With the FEC-MD transcoding scheme described in Section 2.2, each frame is encoded into 16 network packets with payload length 200 bytes. Including the IP/UDP/RTP header, the total transmission rate is 307.2 kbps.

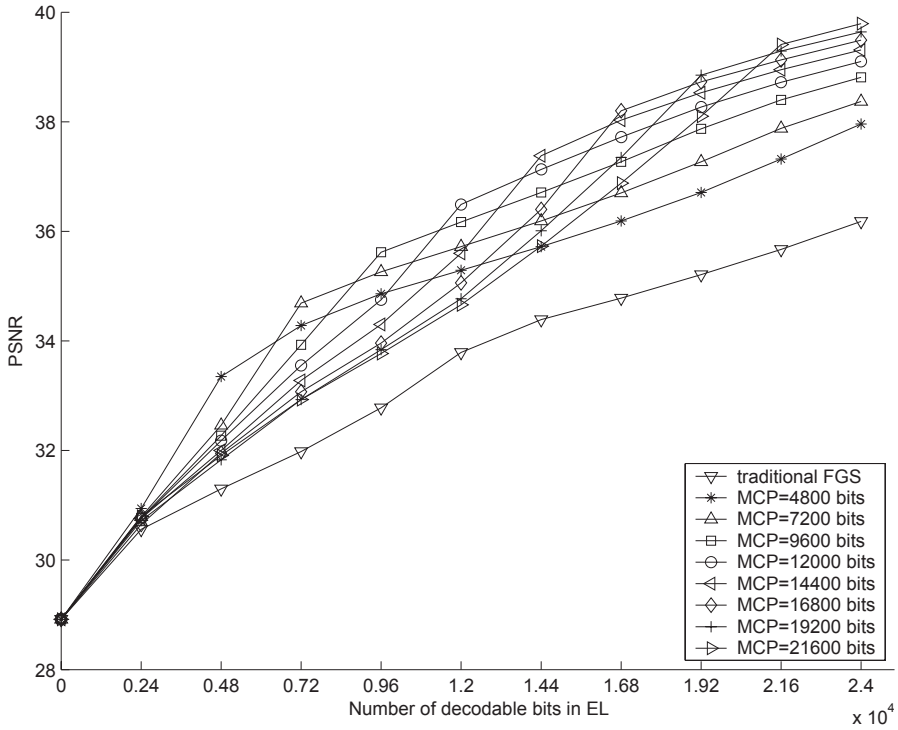

Figure 3.4: PSNR vs. Rate performance for different MCP points.

Two independent paths are employed in the simulation. The proposed transport scheme with optimal traffic allocation is compared with a reference transport scheme with equal traffic allocation. In both of these schemes, the MCP point selection and

the FEC parity assignment are jointly optimized.

The simulations are conducted on the two channel models described previously, including the Gilbert-Elliot model and the $M/D/1/K$ queue model.

3.5.2 Simulation based on Gilbert-Elliot Model

In this subsection, the simulation results and discussions based on the Gilbert-Elliot model are presented. The employed paths are denoted as $Path_1$ and $Path_2$. PL_1 and ABT_1 denote the packet loss rate and average bursty time of $Path_1$. Likewise, PL_2 and ABT_2 denote the packet loss rate and average bursty time of $Path_2$.

Performance Comparison on Asymmetric Paths with Different PL and Equal ABT

In this part, we compare this proposed transport scheme with the reference scheme using two paths with equal ABTs and different PLs.

Fig. 3.5, 3.6 and 3.7 present the comparison between different transport schemes. In all these figures, PL_2 is set to 10% while PL_1 is varied from 5% to 35%. The average bursty time ABT_1 and ABT_2 are identical and set to 0.0125, 0.025 and 0.05 second in these figures, respectively. Each of these figures consists of three subplots. The top plot compares performance in PSNR of these two transport schemes. The middle plot compares the number of packets sent to $Path_1$ in each frame period, and the bottom plot compares the MCP point selection.

In the comparison showed in the top plots, the performance gain due to optimal traffic allocation can be seen. As expected, if PL_1 and PL_2 are relatively similar, the improvement using the optimal traffic allocation is marginal. This is because equal traffic allocation is the optimum choice when the employed paths are with same transport characteristic. But as the two paths become more asymmetric, increased

performance improvement can be achieved by the proposed transport scheme. For example, in Fig. 3.7, when PL1 is 35% and PL2 is 10%, the proposed transport scheme improves the performance by 0.6 dB. As expected, more packets will be sent along the better path with optimal traffic allocation. In the previous example, only 1 packet is sent along $Path_1$ per frame period since the packet loss rate of $Path_1$ is relatively high.

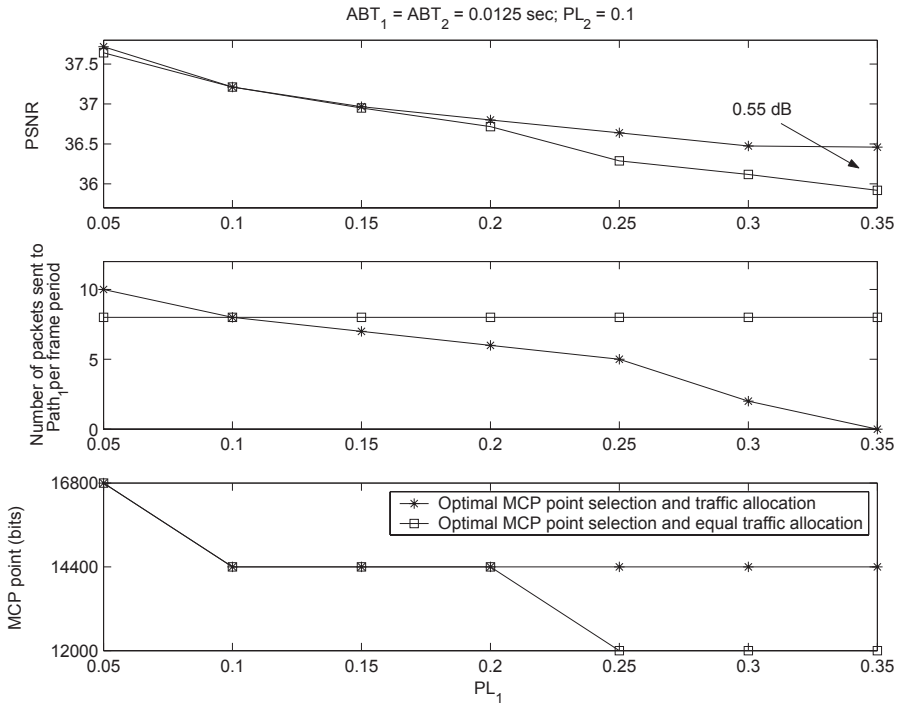

Figure 3.5: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with Different PLs and Equal ABTs; $PL_1 = 0.1$ and $ABT_1 = ABT_2 = 0.0125$ sec.

Similar behaviors can be observed in Fig. 3.5 and 3.6. More packets are sent to the better path and the optimized MCP point of the transport scheme with optimal traffic allocation is always higher than that of the reference transport scheme.

An interesting fact is that the traffic allocation strategy also affect the construction of the MD codes. For example, in Fig. 3.5, when $PL_1 = 0.3$, the MCP point is optimized to 14400 bits/frame when optimal traffic allocation is employed while the

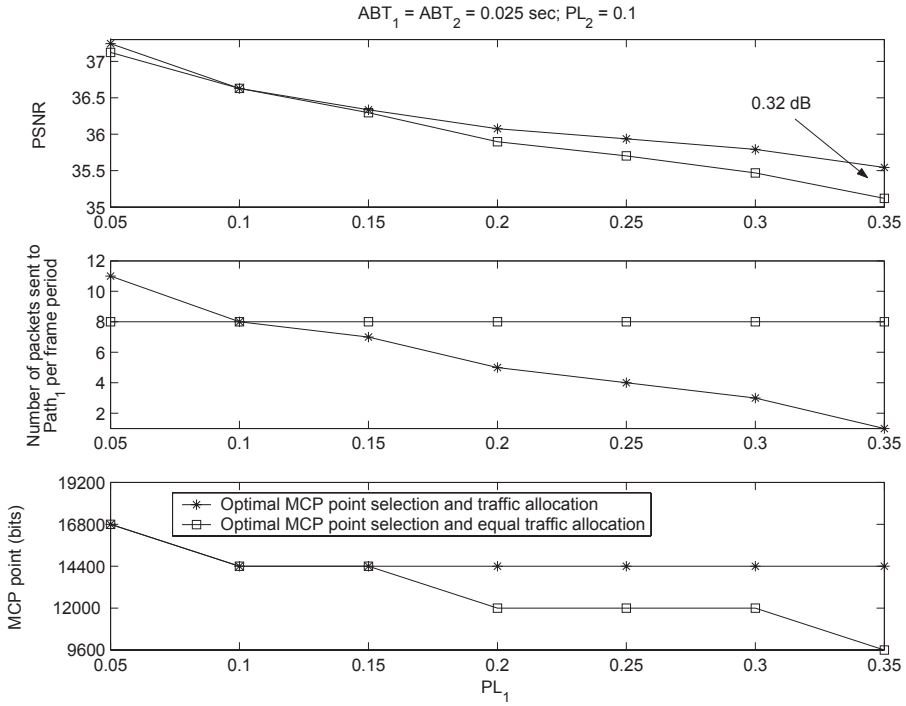

Figure 3.6: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with Different PLs and Equal ABTs; $PL_1 = 0.1$ and $ABT_1 = ABT_2 = 0.025$ sec.

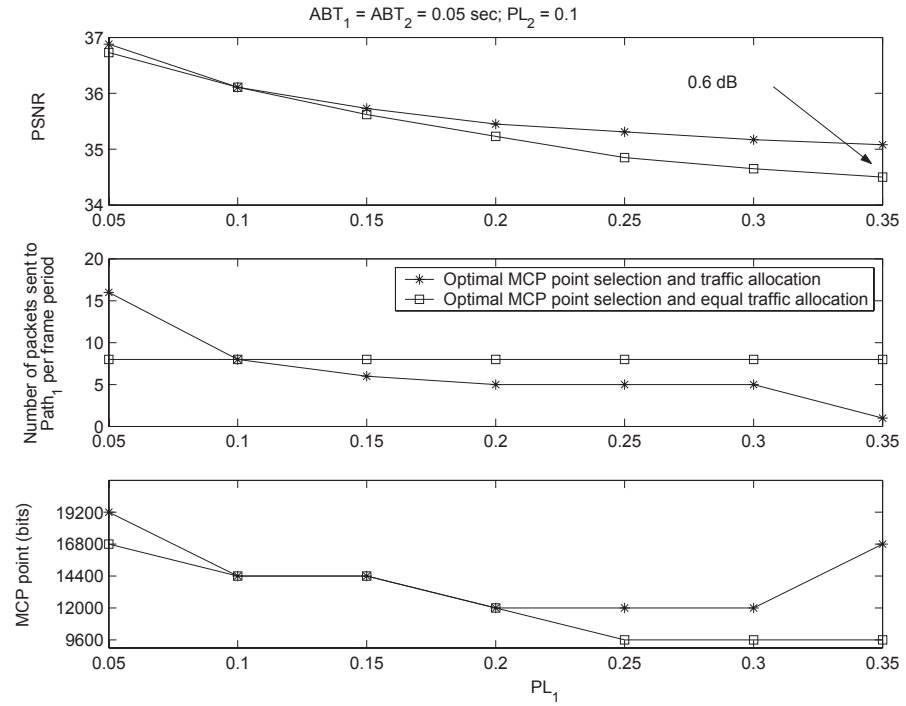

Figure 3.7: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with Different PLs and Equal ABTs; $PL_1 = 0.1$ and $ABT_1 = ABT_2 = 0.05$ sec.

optimal MCP point is 12000 bits/frame in the second reference scheme.

The performance improvement achieved by optimized allocation strategy compared to equal allocation strategy is attributed to optimization of the time interval between adjacent video packets. When the packet loss is independent, all packets will be sent along the better path. But since the loss process over the Internet is usually bursty, some packets still need to be sent along the worse path. One example is in Fig. 3.7 one of the 16 packets is still sent to $Path_1$ when $PL_1 = 35\%$ and $PL_2 = 10\%$.

So it can be concluded that the traffic allocator tries to send all packets along the better path, but due to the characteristic of bursty loss nature of the Internet, some packets will still be sent along the worse path.

Performance Comparison on Asymmetric Paths with Different ABT and Equal PL

In this part, we compare this proposed transport scheme with the reference scheme using two paths with different ABTs and equal PLs. Fig. 3.8, 3.9 and 3.10 present the comparison between different transport schemes. As described previously, in each figure the top, middle and bottom plot compares performance in PSNR of these two transport schemes, the number of packets sent to $Path_1$ in each frame period, and the MCP point selection, respectively. ABT_2 is fixed at 0.0125 second while ABT_1 is varied from 0.0125 to 0.125 second in all these figures. PL_1 and PL_2 are equal and set to 10%, 20% and 30% in these figures, respectively.

It can be observed that the performance of the transport scheme with optimal traffic allocation outperforms the reference scheme in all the cases. Similar to the preceding part, the performance gain is marginal when these two paths are relatively symmetric while the performance gain increases as these two paths become more asymmetric. The performance gain is 0.4dB, 0.8dB and 0.84dB in Fig. 3.8, 3.9

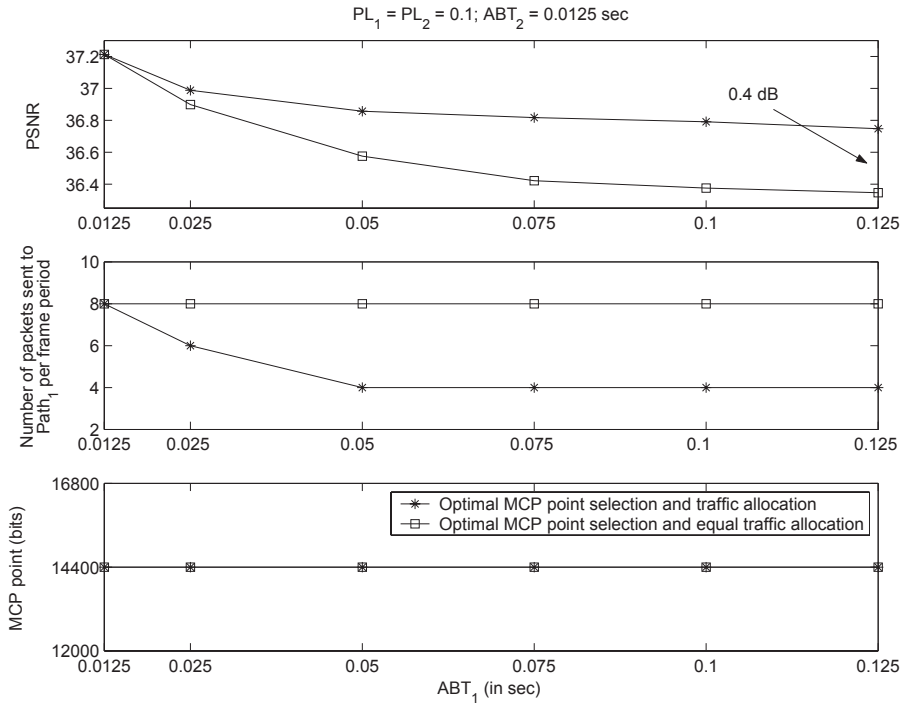

Figure 3.8: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with equal PLs and different ABTs; $PL_1 = PL_2 = 0.1$ and $ABT_2 = 0.0125$ sec.

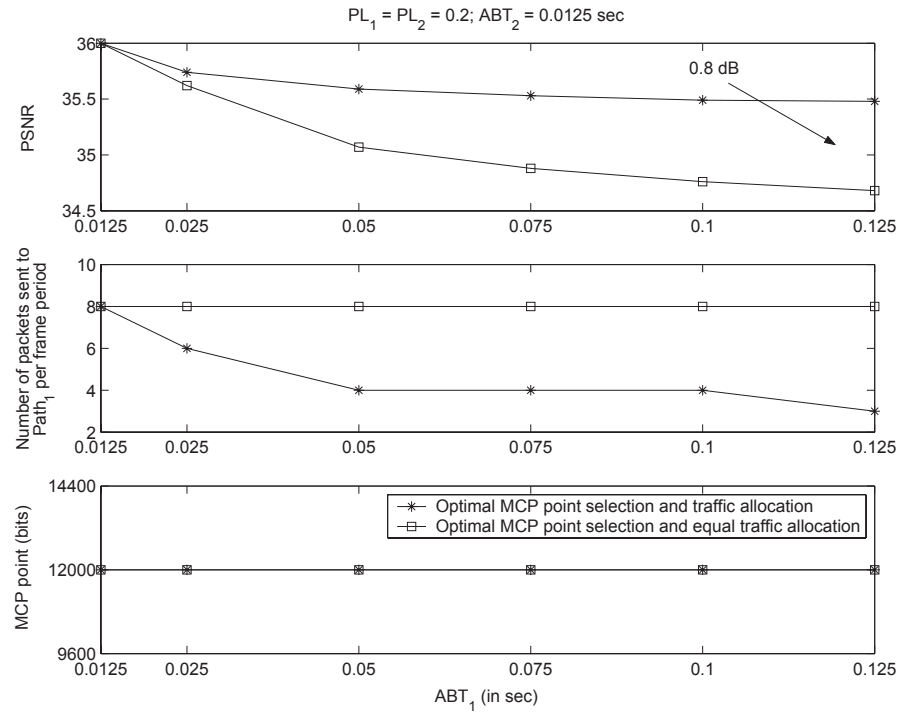

Figure 3.9: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with equal PLs and different ABTs; $PL_1 = PL_2 = 0.2$ and $ABT_2 = 0.0125$ sec.

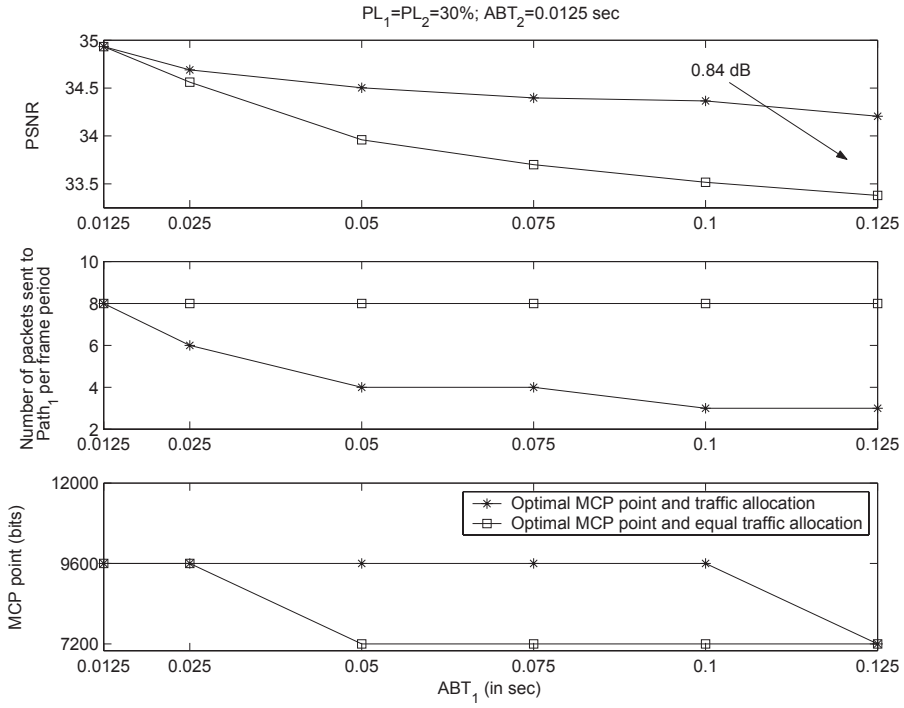

Figure 3.10: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with equal PLs and different ABTs; $PL_1 = PL_2 = 0.3$ and $ABT_2 = 0.0125$ sec.

and 3.10, respectively. It is interesting to observe that, the PSNR improvement will increase as the packet loss rate of both paths increases. This is because if the average burst times of both paths are fixed, the loss processes of both paths become more similar to an independent loss process as their packet loss rate increase. Hence, the more their packet loss rates are, the more PSNR improvement can be achieved.

It also can be seen that fewer packets are sent to $Path_1$ by optimal traffic allocation as the ABT_1 increases. When ABT_1 is relatively large, for example 0.125 second, only 3 packets are sent along $Path_1$ to achieve 0.8dB performance gain.

Based on this simulation, it can be concluded that even using paths with similar packet loss rate, the packets should not be allocated to these paths equally if their average burst durations are not similar.

3.5.3 Simulation based on M/D/1/K queue model

In this subsection, the simulation results and discussions based on the $M/D/1/K$ queue model are presented. As described previously, the employed paths are denoted as $Path_1$ and $Path_2$. ρ_1 and K_1 denote the network load and maximum number of packets that the queue system can hold of $Path_1$. Likewise, ρ_2 and K_2 denote the network load and maximum number of packets that the queue system can hold of $Path_2$.

Performance Comparison on Asymmetric Paths with Different ρ and Equal K

In this part, the transport scheme with optimal traffic allocation and the reference scheme is compared on two asymmetric paths with different ρ and equal K. ρ_2 is fixed at 1.0 while ρ_1 is varied from 0.25 to 2.0, and K_1 and K_2 are identical and set to 3, 5 and 7 in Fig. 3.11, 3.12 and 3.13, respectively. In each figure, top, middle and bottom plot compares performance in PSNR, the number of packets sent to $Path_1$ in each frame period, and the MCP point selection, respectively.

In these figures, we can see: if ρ_1 and ρ_2 are relatively similar, the improvement using the proposed transport scheme is marginal, as might be expected. But as the two paths become more asymmetric, increased performance improvement can be achieved by the proposed transport scheme. And as expected,

For example, in Fig. 3.11 for $K_1 = K_2 = 3$, when $\rho_1 = 0.25$, the transport scheme with optimal traffic allocation improves the performance by 0.6 dB. As expected, more packets will be sent along the better path with optimal traffic allocation. In this example, 13 packets will be sent along $Path_1$ in the optimal traffic allocation scheme while 3 packets will be sent along $Path_2$. And it also can be observed that the traffic allocation strategy affect the construction of the MD codes. Again, in

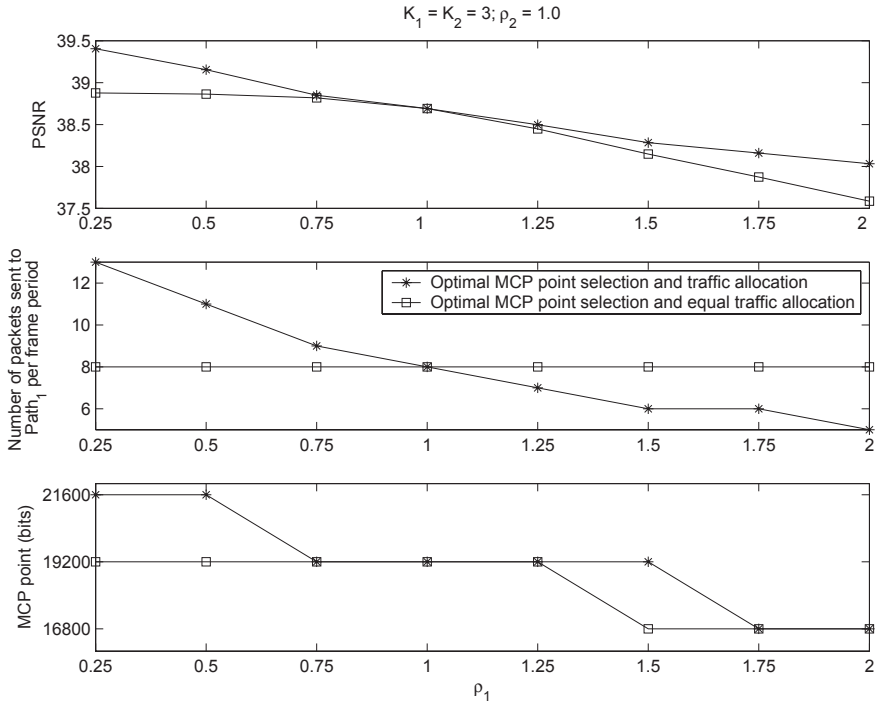

Figure 3.11: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with Different ρ and Equal K; $K_1 = K_2 = 3$ and $\rho_2 = 1.0$.

the same example, the MCP point is optimized to 21600 bits/frame when optimal traffic allocation is employed while the optimized MCP point is 19200 bits/frame in the scheme with equal traffic allocation. When ρ_1 and ρ_2 are relatively similar, the performance gain is marginal and MCP points are identical.

K_1 and K_2 are still identical and increase to 5 and 7 in Fig. 3.12 and 3.13, respectively. Similar behaviors can be observed. More packets are sent to the better path and optimized MCP points in the transport scheme with optimal traffic allocation is always higher than the MCP points in the transport scheme with equal traffic allocation. It can be seen that the performance improvement increases as K_1 and K_2 increase. For example, when $\rho_1 = 2.0$, the performance gain is 0.45dB, 0.74dB and 1.1dB for $K_1 = K_2 = 3$, 5 and 7, respectively. This is because the paths are more potential when K_1 and K_2 increases and optimal traffic allocation can make full use

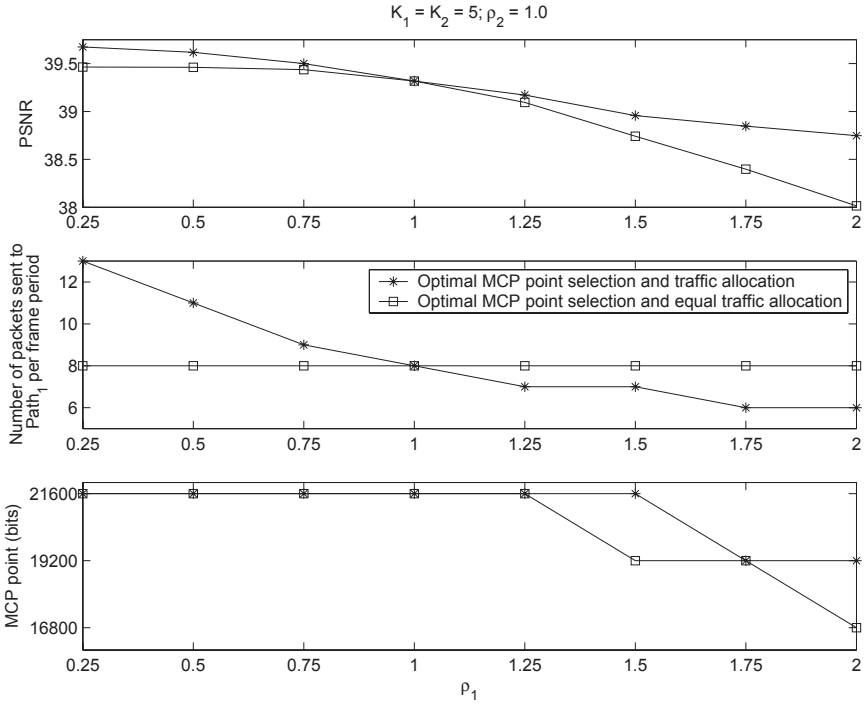

Figure 3.12: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with Different ρ and Equal K; $K_1 = K_2 = 5$ and $\rho_2 = 1.0$.

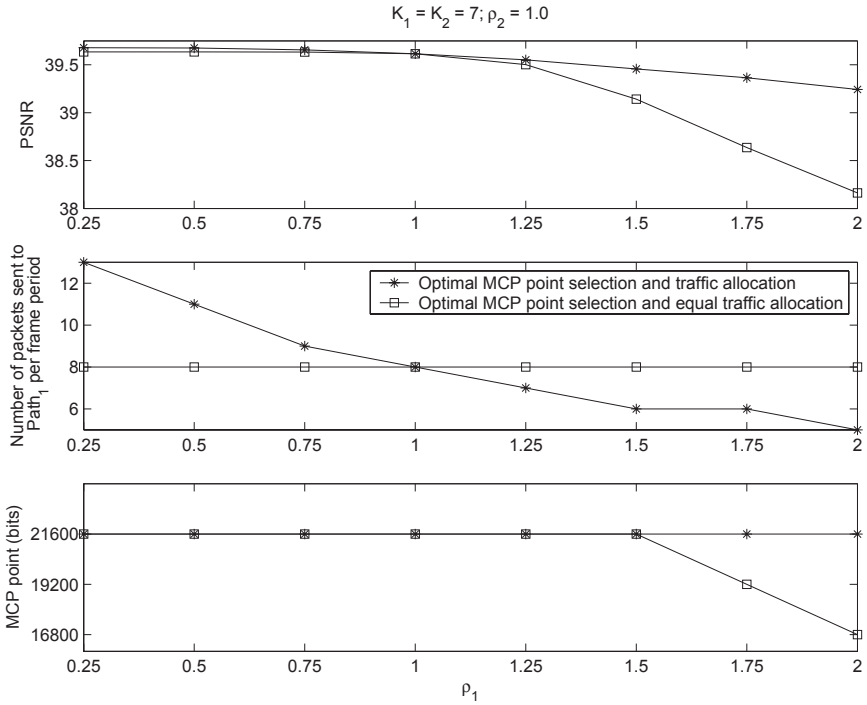

Figure 3.13: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with Different ρ and Equal K; $K_1 = K_2 = 7$ and $\rho_2 = 1.0$.

of this potential.

Performance Comparison on Asymmetric Paths with equal ρ and different K

Likewise, the performance of the transport scheme with optimal traffic allocation and the one with equal allocation is compared on two asymmetric paths with equal ρ and different K. $\rho_1 = \rho_2$ and they are both set to 0.5, 1.0 and 1.5 in Fig. 3.14, 3.15 and 3.16, respectively. $K_2 = 3$ and K_1 is varied from 2 to 12 in all figures. Similarly, each of these figures consists of three parts. The top plot compares performance in of these two transport schemes. The middle plot compares the number of packets sent to $Path_1$ in each frame period, and the bottom plot compares the MCP point selection.

Again, it can be observed that the improvement using the proposed transport scheme is marginal if K_1 and K_2 are relatively similar. But as the two paths become more asymmetric, increased performance improvement can be achieved by the proposed transport scheme.

For example, in Fig. 3.15 for $\rho_1 = \rho_2 = 1.0$, when $K_1 = 12$, the transport scheme with optimal traffic allocation improves the performance by 0.6 dB. In the optimal traffic allocation scheme, 12 packets are sent along $Path_1$ while 4 packets are sent along $Path_2$. The MCP point is optimized to 21600 bits/frame when optimal traffic allocation is employed while the optimized MCP point is 19200 bits/frame in the scheme with equal traffic allocation. When K_1 and K_2 are relatively similar, the performance gain is marginal and MCP points selections are identical.

It can be observed that when network load of both paths is relatively light, packets are allocated into these two paths unevenly. For example, all of the 16 packets are sent to $Path_1$ when $K_1 = 12$ and $\rho_1 = \rho_2 = 0.5$ in Fig. 3.14. This is because sending more

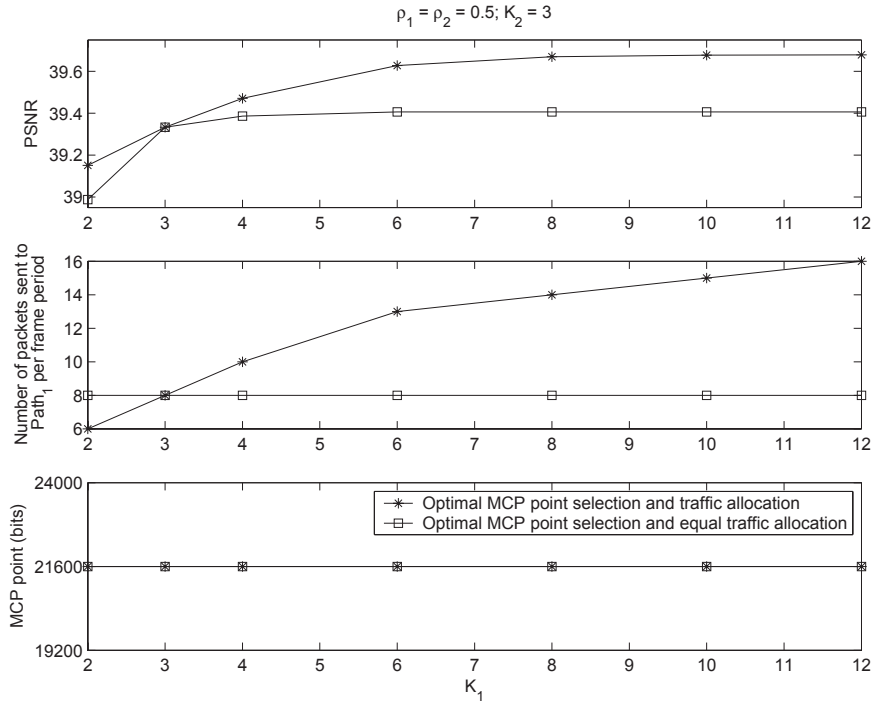

Figure 3.14: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with Equal ρ and Different K; $\rho_1 = \rho_2 = 0.5$ and $K_2 = 3$.

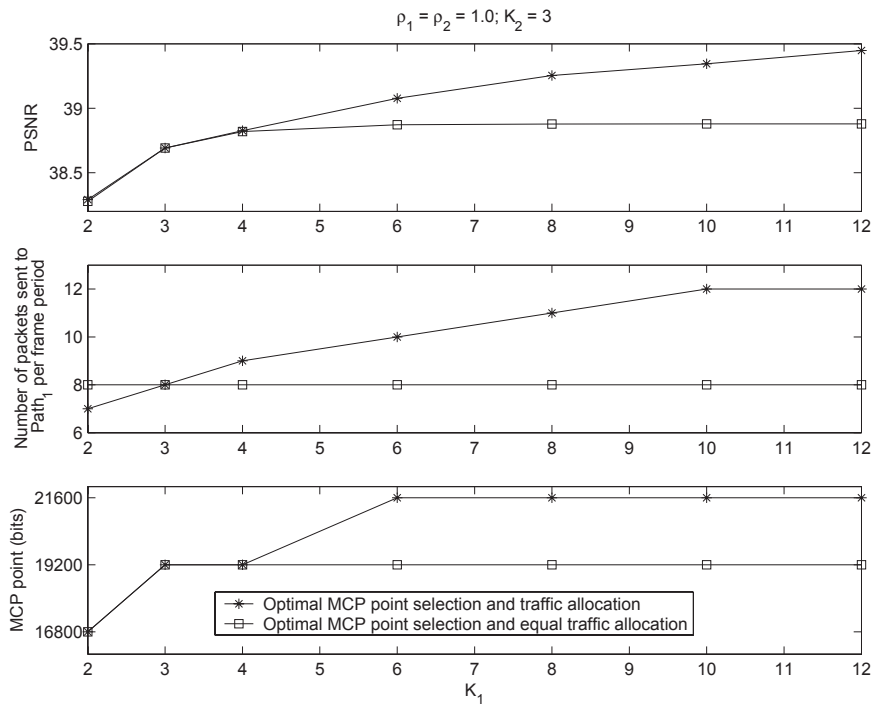

Figure 3.15: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with Equal ρ and Different K; $\rho_1 = \rho_2 = 1.0$ and $K_2 = 3$.

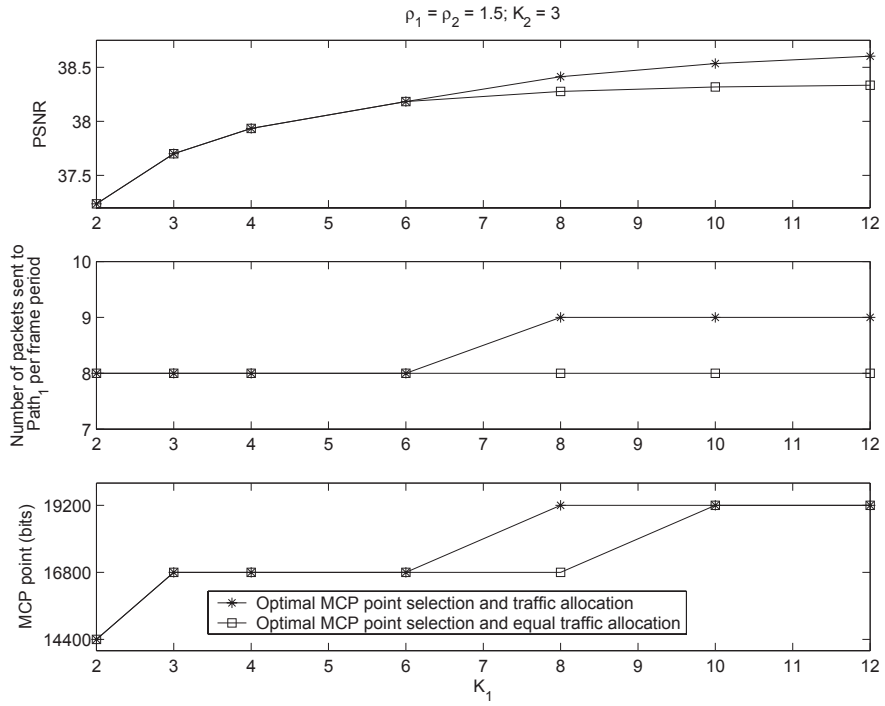

Figure 3.16: Optimized Traffic Allocation Vs. Equal Traffic Allocation on Paths with Equal ρ and Different K; $\rho_1 = \rho_2 = 1.5$ and $K_2 = 3$.

packets will not deteriorate the congestion too much when the original network load is light. But, when network load of both paths is relative high, packets are allocated into these paths more evenly. In another example, only 9 packets are sent to $Path_1$ when $K = 12$ and $\rho_1 = \rho_2 = 1.5$ in Fig. 3.16. This is because that the network load is already very high and sending more packets will deteriorate the congestion and furthermore weaken the transport performance. So, it can be concluded that more packets will be allocated into the better path when network load of both paths is relatively low while optimal traffic allocation will converge to equal traffic allocation as the network load of both paths increases.

3.6 Conclusion

In this Chapter, we proposed a real-time video transport system which combines MC-FGS source coding, FEC-MD transcoding, and traffic allocation strategy. Based on the network states of employed paths, MC-FGS video source coder, the allocation of ULP and the traffic allocation on different paths are jointly optimized to minimize the end-to-end video distortion. The optimal traffic allocation strategy is studied emphatically. We compared the transport scheme with optimal traffic allocation strategy with a reference scheme with equal traffic allocation strategy. The simulation results show that: the proposed scheme can result in substantial performance improvement. Based on the simulations using the Gilbert-Elliot model, some of the packets should still be sent along the path with higher packet loss rate because employing multiple paths can effectively reduce the effects of bursty losses; even using paths with similar packet loss rate, the packets should not be allocated to these paths equally if their average burst durations are not similar. Based on the simulations using $M/D/1/K$ queue, the performance gain improves as K increases; more packets are allocated to the better path when the network load of both paths is light while optimal traffic allocation converges to equal traffic allocation as the network load increases.

CHAPTER 4

A Model-Based Approach to Evaluation of the Efficacy of Progressive Transport of Scalably Encoded Sources in Congested Networks

4.1 Motivation

In previous Chapters, we studied optimal interleaving and multiple path transport schemes in video transport. In this Chapter, we will investigate the efficacy of different transport schemes, such as ULP, FEC and best-effort (BEF) transport schemes, of progressively or scalably encoded sources. As discussed previously, progressive or scalable source coding, particularly for image and video transport, can be effective in coping with dynamic network conditions and heterogeneous terminals [49, 50]. In progressive coding schemes, the source is compressed into a single bitstream, which can be truncated at different points and decoded at different rates. The resulting reconstruction quality of the source depends upon how much of the bitstream is actually successfully decoded. However, at the same time, it makes the compressed bitstreams highly structured and prioritized [107, 108] so that the decoded reconstruction quality of the source generally depends on the position of the first loss. Specifically, the decoder can only decode the part before the first loss and the part after this first loss

are essentially useless.

Forward error correction[3] (FEC) is frequently used to mitigate the impact of packet losses in network transport, but sending redundant packets can also increase the network load and possibly lead to a higher packet loss rate due to the resulting increased congestion. In [105], the efficacy of best-effort (BEF) and FEC transport schemes for scalable video coding were investigated based on both a Gilbert packet loss model and a renewal packet loss model. The results show that for these channels, using the proper amount of FEC can improve the quality of received video substantially. However, the efficacy of using FEC to protect a scalably encoded source is still highly limited because the decoded quality of the source is still dependent on the position of the first unreconstructed loss [105]. An unequal loss protection (ULP) scheme is described in [2], and later extended in [73], which assigns different amounts of protection to the different layers or parts of a bitstream based on their relative importance. In this way, a scalable, prioritized bitstream can be converted into an arbitrary number of packets with approximately equal importance [2, 73], which is considered to be better matched to the best-effort transport properties of current networks. Using this approach, the reconstructed quality of the source only depends on the actual number of received packets and not the relative temporal positions of the losses. Efficacy comparisons between FEC transport and ULP transport schemes can be found in [2, 109–111], among others, but most of these works are based on use of either the independent loss model or the Gilbert loss model, neither of which provide an accurate model of packet loss in congested networks.

To analytically investigate the efficacy of error resilient transport schemes for progressively coded sources, including ULP, BEF and FEC transport schemes for

[3]By FEC we mean here fixed protection of each packet regardless of its importance in distinction to unequal loss protection (ULP) where the protection is tailored to the importance of a packet.

more general loss models, we first develop evaluation and optimization approaches for these transport schemes. In these evaluation approaches, three important quantities are defined: $P(j,n)$, which is the probability of losing j packets in a block of n packets; $P_{BEF}(i,n)$, the probability of packet i being the first lost packet in a block of n packets, and $P_{FEC}(i,n,k)$ the probability of packet i being the first lost and unrecoverable packet in k information packets with FEC protection using block size n. Each of these quantities provides the key to evaluation of the efficacy of the corresponding transport scheme. The analysis of block error density $P(j,n)$ for different channel models has been presented in many prior works. For example, in [112] the quantity $P(j,n)$ is derived by modeling the channel loss process as a Gilbert model. Cidon et al. [113] provided a recursive algorithm to evaluate $P(j,n)$ based on the $M/M/1/K$ queueing model. In [114], analytical formulas for $P(j,n)$ for the $M/M/1/K$ queue are derived using multi-dimensional probability generating functions. Dan et al. [115] provide an analysis of $P(j,n)$ using an $M/D/1/K$ queueing model. However, far less work has been reported on the evaluation of the latter two quantities, $P_{BEF}(i,n)$ and $P_{FEC}(i,n,k)$.

In this Chapter, we provide a comprehensive queueing analysis to evaluate these three crucial quantities based on an $M/D/1/K$ queue [4] model. Then, armed with these results, the efficacy of these three transport schemes for progressively encoded sources are investigated and compared. The major contributions of this Chapter are twofold. Firstly, we provide explicit queueing analysis approaches to evaluate the corresponding end-to-end performance of these three transport schemes. Secondly, we quantitatively analyze and compare the efficacy of these three schemes within a common modeling framework under both idealized source modeling assumptions as

[4] In this Chapter, we focus on the case of fixed packet sizes so that previous work for the $M/M/1/K$ queue model, assuming exponentially distributed packet sizes, are inappropriate.

well as for real-world video sources.

The reminder of this Chapter is organized as follows. Section 4.2 describes each of the three transport schemes under study and formulates appropriate optimization approaches over the corresponding parameter choices. In Section 4.3, we provide a queueing analysis for these three transport schemes under an $M/D/1/K$ queue assumption. In Section 4.4, we quantitatively demonstrate and compare the resulting end-to-end performances through some selected numerical examples. Finally, Section 4.5 provides a summary and conclusions.

4.2 System Description

In this section, we describe the ULP, BEF and optimal FEC transport schemes for progressively encoded bitstreams.

4.2.1 Unequal Loss Protection (ULP) for Progressive Source Coding

A source is assumed progressively coded into a single bitstream. As illustrated in Fig. 2.2, this bitstream is arranged in order of decreasing importance into an array with L rows and n_u columns [5]. Each column represents the payload of a packet which is subsequently sent to the transport network and each row is defined as a stream. Each stream constitutes a channel coding block of length n_u symbols. The symbols to the left of the indicated boundary in Fig. 4.1 are information symbols, those to the right of the boundary are FEC parity symbols. In each stream i, $1 \leq i \leq L$, there are k_i information symbols and $f_i = n_u - k_i$ parity FEC symbols. Therefore, with a Reed-Solomon $RS(n_u, k_i)$ FEC code, up to f_i packet losses can be recovered in the $i'th$

[5] We let n_u, n_b and n_f represent the block size n for the ULP, BEF and FEC transport schemes, respectively, to emphasize that different block sizes may be used for the different schemes.

stream. Because the earlier parts are more important than the later subsequent parts, we assume the f_i are nonincreasing with layer index i, such that $f_{i+1} \leq f_i$. Under this assumption, it follows that if stream i can be decoded then all prior streams j, for $j \leq i$ can likewise be successfully decoded. With this ULP scheme, a progressively coded bitstream is transcoded into n_u packets with roughly equal importance.

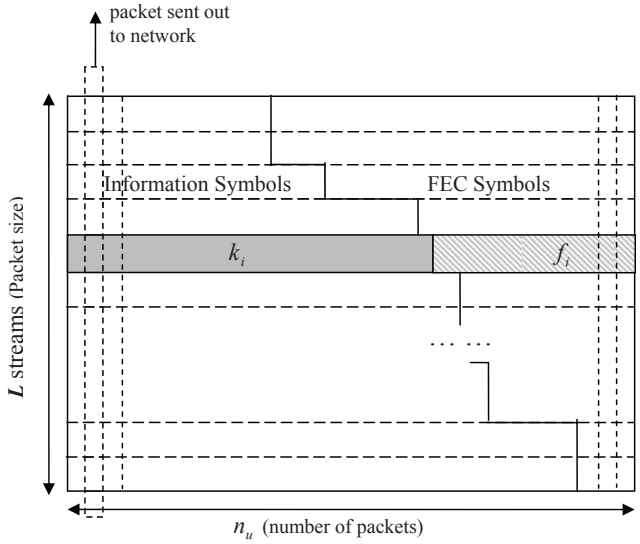

Figure 4.1: The unequal loss protection (ULP) scheme for a scalably encoded source.

Let the $L-$vector $\mathbf{f} = (f_1, f_2, ..., f_L)$ represent the FEC parity assignment across streams. If j, $1 \leq j \leq n_u$, packets are received, it follows that all the layers with $k_i \leq j$ can be successfully decoded. For a given $\mathbf{f}$ and i, we define $M(i|\mathbf{f}) = \sum_{l=1}^{i} k_l$ as the number of information symbols that can be recovered if the i'th stream can be successfully decoded.

Let $P(j, n_u)$, $n_u \geq 1, 0 \leq j \leq n_u$, denote the probability that j packets are lost in a block of n_u consecutive packets. It follows that $M(i|\mathbf{f})$ information symbols can be decoded with probability $C(f_i)$, given by

$$C(f_i) = \sum_{j=0}^{f_i} P(j, n_u), \ i = 1, 2, ..., L. \tag{4.1}$$

We assume the distortion-rate function $D(R)$ for a known source and a specified distortion metric are given and the distortion can be determined as long as the decoding rate R is available. Similar to [2, 77, 110], define the decrease in distortion, given parity assignment $\mathbf{f}$, resulting from the decoding of stream i as

$$G_i(\mathbf{f}) = D(M((i-1)|\mathbf{f})) - D(M(i|\mathbf{f})),\ i = 1, 2, ..., L, \tag{4.2}$$

where $D(M(i|\mathbf{f}))$ denotes the distortion when the decoder decodes the first $M(i|\mathbf{f})$ information symbols, and we assume $D(M(0|\mathbf{f}))$, the distortion if no information symbols are successfully decoded, is known.

Hence, using the distortion-rate (DR) function $D(\cdot)$ and block error density function $P(j, n_u)$, we can then minimize the expected distortion by choosing the FEC assignment accordingly to achieve

$$\overline{D}^*_{ULP} = min_{\{\mathbf{f}\}}\{\sum_{i=1}^{L} C(f_i)G_i(\mathbf{f})\}. \tag{4.3}$$

There are several methods available for adjusting the FEC allocation to achieve $\overline{D}^*_{ULP}$ while maintaining the total size of the array, $n_u \cdot L$ symbols, in Fig. 2.2 fixed. In what follows, we make use of the greedy search algorithm described in [2]. Specifically, starting from an initial configuration, a new allocation of FEC symbols are selectively added to successive streams and the displaced information symbols are then moved to the next sequential stream. This causes a cascade of information symbols to move down the streams until information symbols from the last stream are discarded. In this way, the payload length of each of the n_u packets sent to the network is maintained at L symbols. This determines the packet arrival rate, in appropriate units, at a network service facility and allows evaluation of the corresponding network load if the service rate of the network facility is known.

4.2.2 Best-Effort (BEF) Transport for Progressive Source Coding

In the BEF transport scheme, the progressively encoded bitstream is packetized directly and then sent out to the network without any FEC protection. Because the later parts cannot be decoded without the earlier parts of the progressively coded bitstream, only the packets before the first loss are decodable and the packets after the first loss are useless and are discarded. We assume that the progressive bitstream is encapsulated in n_b packets, each with a payload of L symbols. Here, to provide a basis for comparison with the ULP transport scheme, the packet length L is kept at the same value as in the previous subsection, although the number of packets transmitted, n_b, may be different. Define $P_{BEF}(i, n_b)$, $i = 1, 2, ..., n_b$, as the probability that packet i is the first lost packet in n_b successive packets. Obviously, with probability $P_{BEF}(i, n_b)$, there are then a total of $(i-1)L$ information symbols decoded. Given the DR function $D(\cdot)$, decoding of $(i-1)L$ information symbols will lead to a distortion of $D((i-1)L)$ and, with probability $1 - \sum_{i=1}^{n_b} P_{BEF}(i, n_b)$, these n_b packets are all successively decoded, which will lead to a distortion of $D(n_b L)$. Hence, the expected distortion for the BEF transport scheme is:

$$\overline{D}_{BEF} = \sum_{i=1}^{n_b} P_{BEF}(i, n_b) D((i-1)L) + (1 - \sum_{i=1}^{n_b} P_{BEF}(i, n_b)) D(n_b L). \tag{4.4}$$

4.2.3 Optimal Forward Error Correction (FEC) Transport for Progressive Source Coding

In the FEC transport scheme, a fixed number of FEC parity packets are sent following the information packets to mitigate packet losses. However, sending FEC parity packets will also increase the network load, which can aggravate packet losses. Hence, it is necessary to optimize the FEC coding rate to minimize the overall end-to-end distortion. In a typical systematic FEC coding scheme, such as use of systematic

$RS(n_f, k)$ codes, among a block of size n_f, the first k packets are information packets and the following $n_f - k$ packets are FEC packets. Hence, the channel coding rate is $R_c = k/n_f$. With this scheme, as long as the number of lost packets is less than $n_f - k$, all the information packets can be successfully received or reconstructed. To compare with the two previous transport schemes, payloads of the information packets and parity packets are likewise fixed at L symbols. By defining $P_{FEC}(i, n_f, k)$ as the probability that packet i is the first lost and unreconstructable packet in k information packets using an $RS(n_f, k)$ code, the expected distortion can be computed as:

$$\overline{D}_{FEC} = \sum_{i=1}^{k} P_{FEC}(i, n_f, k) D((i-1)L) + (1 - \sum_{i=1}^{k} P_{FEC}(i, n_f, k)) D(kL). \quad (4.5)$$

Then we can minimize the expected distortion by choosing an appropriate block size n_f and the number of information packets k, $k \leq n_f$, to achieve

$$\overline{D}^*_{FEC} = \min_{\{(n_f,k)\in\mathcal{D}\}} \{\sum_{i=1}^{k} P_{FEC}(i, n_f, k) D((i-1)L) + (1 - \sum_{i=1}^{k} P_{FEC}(i, n_f, k)) D(kL)\}, \quad (4.6)$$

where $\mathcal{D}$ denotes the domain of (n_f, k).

4.3 Queueing Analysis of The $M/D/1/K$ Queue

4.3.1 M/D/1/K Queue

As described in the previous subsection, $P(j, n_u)$, $P_{BEF}(i, n_b)$ and $P_{FEC}(i, n_f, k)$ are the fundamental quantities required to optimize and evaluate the corresponding transport schemes for progressively encoded sources. In this section, we will provide an explicit queueing analysis for each of these crucial quantities.

As also described in the previous section, the length of packets generated with the ULP, BEF and optimal FEC schemes are all fixed at the same size of L symbols. Thus, the $M/D/1/K$ queue is an appropriate model to represent the transport network. As

depicted in Fig. 4.2, the packet arrival process is assumed a Poisson process with rate λ, and the system can hold up to K packets. With time units defined appropriately, and with $\lambda = n$ packets per unit time together with the service rate as μ packets per similar time unit, the network load is $\rho = n/\mu$. For the $M/D/1/K$ queue system, the service rate is a deterministic value and the service time is $D = 1/\mu$ time units per packet. If the packet arrival rates for the ULP, BEF and FEC transport schemes are n_u, n_b and n_f packets per unit time, respectively, the corresponding network loads of these three transport schemes are $\rho_u = n_u/\mu$, $\rho_b = n_b/\mu$ and $\rho_f = n_f/\mu$. We assume a FIFO service policy is employed in this queueing system.

Because the service time for each packet is a deterministic value D, the remaining workload can be measured in time units. Let x denote the remaining workload in the system, which is represented by the shadowed part in Fig. 4.2. The remaining workload x means that all the workload of the system will be cleared after time x if there are no additional packet arrivals during this period. Since the service time of a packet is D, the arrival of a packet in the absence of a departure will increase the workload x by D. The maximum workload of the queueing system is KD since the system can hold up to K packets. A necessary condition for receiving a packet is that the workload just before the arrival of this packet is no larger than $(K-1)D$, otherwise the packet will be lost. Define $V(x)$ as the probability distribution of the remaining workload in the queueing system. As described in what follows, $V(x)$ can be approximated by the steady-state distribution of the $M/E_r/1/K$ queue [6] under appropriate scaling as $r \to \infty$. This follows since the service facility of the $M/E_r/1/K$ queue can be considered to consist of a cascade of r stages, with each having identical exponential service time distribution. To keep the mean service rate constant at the

[6] Here, E_r represents the Erlang-r distribution for packet service time.

value $\mu = 1/D$, the service rate of each stage is scaled by a factor of r to $r\mu$. As $r \to \infty$, the service time of the $M/E_r/1/K$ queue then converges to a constant value D. More details on the $M/E_r/1/K$ queue can be found in [116]. The determination of $V(x)$ using this approach is provided in Appendix A.

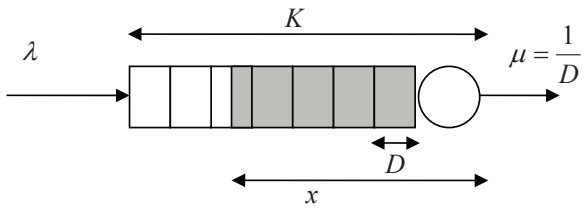

Figure 4.2: $M/D/1/K$ queue system

4.3.2 The Block Error Density Function, $P(j, n_u)$, for the M/D/1/K Queue

Recall, $P(j, n_u)$, $n_u \geq 1, 0 \leq j \leq n_u$, denotes the probability that j packets are lost in a block of n_u consecutive packets, which is a key quantity to optimize allocation of unequal loss protection. In [115], an analysis of $P(j, n)$ is presented for an $MMPP/D/1/K$ queueing system, where $MMPP$ denotes the Markov-modulated Poisson arrival process. In Appendix B, we simplify the arrival process to a single Poisson process and provide a simplified analysis of the $M/D/1/K$ queue. Using this analysis, the quantity $P(j, n_u)$ is evaluated and can furthermore be used to optimize allocation of unequal loss protection. To be noticed, since it is a general approach to compute the block error density function on any $M/D/1/K$ queue, the block error density function is denoted as $P(j, n)$ in Appendix B, which is the probability of losing j packets in a block of n consecutive packets.

An important quantity defined to analyze $P(j, n)$ in Appendix B is $P_x^a(j, n)$, which is the probability of j losses in a block of n packets given that the remaining workload

in the system is x just before the arrival of the first packet of this block. In Appendix B, $P(j,n)$ is first expressed as an expectation value of $P_x^a(j,n)$. Then, a recursive algorithm is formulated to compute $P_x^a(j,n)$. This $P_x^a(j,n)$ is also used to compute $P_{FEC}(i,n_f,k)$ in subsection 4.3.4.

4.3.3 Analysis of $P_{BEF}(i,n_b)$ for the $M/D/1/K$ Queue with Best-Effort Transport

To evaluate the best-effort transport system, analysis of $P_{BEF}(i,n_b)$, which is the probability that packet i is the first lost packet in a block of n_b packets, is necessary. The simplest method to compute $P_{BEF}(i,n_b)$ is $P_{BEF}(i,n_b) = P(0,i-1) - P(0,i)$. Here, $P(0,i-1)$ is the probability of no loss in the block consisting of the first $i-1$ packets whatever the loss pattern in the following block of $n_b - i + 1$ packets is. Similarly, $P(0,i)$ is the probability of no loss in the block of the first i packets whatever the loss pattern in the block of the following $n_b - i$ packets is. So, $P(0,i-1) - P(0,i)$ is then the probability that packet i is the first lost packet in this block. Unfortunately, this evaluation method for $P_{BEF}(i,n_b)$, cannot provide an analysis of the conditional probability distribution of the workload x just before the arrival of packet i given that packet i is lost, which is crucial to evaluate $P_{FEC}(i,n_f,k)$. Hence, this evaluation method using the block error density cannot be extended to evaluate $P_{FEC}(i,n_f,k)$. To lay a solid basis for the analysis of $P_{FEC}(i,n_f,k)$ provided in the next subsection, in this subsection we provide an alternative analytical method to evaluate the quantity $P_{BEF}(i,n_b)$. First, we define:

$$V_j(x) = Prob(\textit{the workload of the system is } x \textit{ just before the arrival of packet } j \textit{ and packets } 1,\ 2, ...,\ j-1 \textit{ are successfully received}) \quad (4.7)$$

The quantity $V_j(x)$ can likewise be computed recursively. The recursion is initiated

with the following relations:

$$V_1(x) = V(x), \tag{4.8}$$

and

$$\begin{aligned} V_j(0) = & \int_0^{(K-1)D} \int_{y+D}^{\infty} f(t) V_{j-1}(y)\, dt\, dy \\ = & \int_0^{(K-1)D} [\int_{y+D}^{\infty} f(t)\, dt] V_{j-1}(y)\, dy \\ = & \int_0^{(K-1)D} e^{-\lambda(y+D)} V_{j-1}(y)\, dy. \end{aligned} \tag{4.9}$$

The first of these two expressions is merely a statement of the fact that the distribution of workload x just before the arrival of packet 1 is the stationary distribution, $V(x)$, of the workload. In the second expression, y denotes the workload just before the arrival of packet $j-1$. Because each arrival packet will increase the workload of the queueing system by D, the workload will be $y+D$ just after the arrival of packet $j-1$. As a result, it will take time $y+D$ to clear the workload. In other words, the workload will be 0 after time $y+D$ if no additional packets arrive. Hence, if the workload is y and 0 just before the arrivals of packets $j-1$ and j, respectively, the interarrival time between packets $j-1$ and j is no less than $y+D$. This is the reason why the inner integral on interarrival time t in (4.9) is from $y+D$ to ∞. Likewise, the integral on y is from 0 to $(K-1)D$ because y must be less than $(K-1)D$ to ensure there is enough room in the buffer to successively receive packet $j-1$.

The resulting recursive expression for computing $V_j(x)$ is then:

$$V_j(x) = \int_{max(0,D-x)}^{KD-x} f(t) V_{j-1}(x-D+t)\, dt. \tag{4.10}$$

This follows from the fact: if the workload just before the arrival of packet j is x and the interarrival time between packets $j-1$ and j is t, the workload just after the arrival of packet $j-1$ must be $x+t$. So, the workload just before the arrival of packet

$j-1$ must be $x-D+t$. Because the workload $x-D+t$ is nonnegative, the lower limit of the integral is $max(0, D-x)$. This makes sense because the minimum workload just before the arrival of packet $j-1$ is 0, increased by D due to the reception of packet $j-1$, the minimum workload just after reception of packet $j-1$ is D. Thus, the minimum workload just before the arrival of packet j is $max(0, D-t)$. Since $x-D+t$ is nonnegative, it follows that $t > D-x$. And, likewise, because t is nonnegative, $t > max(0, D-x)$, which is the lower limit of the integral. The upper limit of the integration is $KD-x$ so that the workload $x-D+t \leq (K-1)D$, which guarantees the reception of packet $j-1$.

With the analysis above, the probability $P_{BEF}(i, n_b)$ that packet i is the first lost packet in a block of n_b packets can be computed from:

$$P_{BEF}(1, n_b) = \int_{(K-1)D}^{KD} V_1(x)\, dx, \tag{4.11}$$

and

$$\begin{aligned} P_{BEF}(i, n_b) = & \int_{(K-2)D}^{(K-1)D} [\int_0^{x-(K-2)D} f(t)\, dt] V_{i-1}(x)\, dx \\ = & \int_{(K-2)D}^{(K-1)D} \{1 - e^{-\rho[x-(K-2)D]}\} V_{i-1}(x)\, dx, \ i > 1, \end{aligned} \tag{4.12}$$

where t denotes the interarrival time between packet i and packet $i-1$, and $V_{i-1}(x)$ denotes the distribution of the workload just before the arrival of packet $i-1$. The first expression is based on the simple fact that a necessary condition for receiving a packet is that the workload just before the arrival of this packet is no larger than $(K-1)D$, otherwise the packet will be lost. Hence, the first packet will be lost if $(K-1)D < x \leq KD$. The explanation for the second expression is: since the workload just before the arrival of packet $i-1$ is x, the workload just after the reception of packet $i-1$ is $x+D$. After the interarrival time t, the workload just before the arrival of packet i is $x+D-t$. Because packet i is lost, $x+D-t > (K-1)D$.

Hence, it follows that $t < x - (K-2)D$. Since the interarrival time $t \geq 0$, the inner integral on t is from 0 to $x - (K-2)D$. For the outer integral on x, because the packet i is lost, $x + D - t$, the workload just before the arrival of this packet i, must be larger than $(K-1)D$. Because $t \geq 0$, $x + D > (K-1)D$. Then, $x > (K-2)D$. On the other hand, because the packet $i-1$ is assumed successively received, the workload x just before the arrival of packet $i-1$ must be no larger than $(K-1)D$ to ensure the reception of packet $i-1$. Thus, the outer integral on x is from $(K-2)D$ to $(K-1)D$.

4.3.4 Analysis of $P_{FEC}(i, n_f, k)$ for the $M/D/1/K$ Queue with FEC

In this subsection, a fixed Reed-Solomon coding scheme, denoted $RS(n_f, k)$, is employed to protect the progressively encoded packets so that the channel coding rate is $R_c = k/n_f$. We assume that the first k packets are information packets and the following $n_f - k$ packets are FEC packets. Let $P_{FEC}(i, n_f, k)$ denote the probability that packet i is the first lost and unreconstructable packet in a block of size n_f, where $1 \leq i \leq k$. Hence,

$$P_{FEC}(i, n_f, k) = P_{BEF}(i, n_f) \cdot Prob(packet\ i\ can\ not \\ be\ reconstructed | packet\ i\ is\ the\ first\ loss). \tag{4.13}$$

As defined previously, $P_{BEF}(i, n_f)$ is the probability that packet i is the first lost packet in the block of length n_f packets. Using the $RS(n_f, k)$ scheme, as long as the number of received packets is greater than or equal to k, all lost packets are reconstructible. Hence, if packet i is the first lost and unreconstructable packet, there must be at least $n_f - k$ packets lost among the following packets $i+1$, $i+2$, ..., n_f. Let $F_i(l, n_f - i)$ denote the probability that l packets are lost among the $n_f - i$ packets

following packet i, given packet i is the first loss. Hence,

$$Prob(packet\ i\ can\ not\ be\ reconstructed|packet\ i\ is\ the\ first\ loss) = \sum_{l=n_f-m}^{n_f-i} F_i(l, n_f - i). \tag{4.14}$$

The quantity $F_i(l, n_f - i)$ can be evaluated according to:

$$F_i(l, n_f - i) = \int_{(K-1)D}^{KD} V_i(x|i)[\int_0^x f(t) P_{x-t}^a(l, n_f - i)\, dt + \int_x^\infty f(t) P_0^a(l, n_f - i)\, dt]\, dx, \tag{4.15}$$

where t denotes the interarrival time between packet i and $i+1$, and $f(t)$ denotes the probability density function of the packet interarrival-time. Again, x denotes the workload just before the arrival of packet i. The quantity $V_i(x|i)$ denotes the conditional probability of the workload x just before the arrival of packet i, given packet i is lost. As mentioned in subsection 4.3.2, $P_y^a(j, n_f - i)$ is the probability that j packets are lost in a block of $n_f - i$ consecutive packets given that the remaining workload is y just before the arrival of packet $i+1$, the first packet of this block of length $n_f - i$. Using the recursive algorithm formulated in Appendix B, the quantity of $P_y^a(j, n_f - i)$ can be computed. Obviously, $(K-1)D < x \leq KD$ because packet i is lost. And the workload just before the arrival of packet $i+1$ is $x - t$ if $t \leq x$ or 0 if $t > x$. These two cases are described by the two integrals in (4.15) versus t. The conditional probability $V_i(x|i)$ can be evaluated by:

$$V_i(x|i) = \frac{V_i(x)}{\int_{(K-1)D}^{KD} V_i(y)\, dy}, \ (K-1)D < x \leq KD. \tag{4.16}$$

4.3.5 Numerical Evaluation

As described previously, $P(j, n_u)$, $P_{BEF}(i, n_b)$ and $P_{FEC}(i, n_f, k)$ are the crucial quantities to evaluate the efficacy of the corresponding three transport schemes. In

this Chapter numerical methods are employed to compute these three quantities. As indicated previously, the workload distribution $V(x)$ can be approximated by the steady-state distribution of the $M/E_r/1/K$ queue in the limit of $r \to \infty$ [116], an analysis of which is included in Appendix A. First, the service time D is divided into N subintervals of length Δ, with $D = N\Delta$ and $\Delta = 1/(N\mu)$. Let r, the number of stages in the service facility of the $M/E_r/1/K$ queue, be equal to N. As N increases to infinity, the $M/E_r/1/K$ queue turns into an $M/D/1/K$ queue. Hence, If N is sufficiently large, the workload distribution $V(x)$ can be accurately approximated by the equilibrium probability P_i of the $M/E_r/1/K$ queue. This means that the discretized workload x can be represented by the remaining stages in the $M/E_r/1/K$ queue system if N is sufficiently large. So, to keep notation consistent, $V(i)$ is used to denote the equilibrium probability P_i of the $M/E_r/1/K$ queue. As indicated previously, ρ_u, ρ_b and ρ_f denote the network loads of these transport schemes, respectively.

Numerical Evaluation of $P(j, n_u)$

A numerical method to evaluate $P(j, n)$ is presented in Appendix B. Likewise, a method to compute $P(j, n)$ is first formulated based on $V(i)$ and $P_i^a(j, n)$, which are the discretized version of $V(x)$ and $P_x^a(j, n)$, respectively. Then, a recursive algorithm is provided to compute $P_i^a(j, n)$. Using this numerical method, $P(j, n_u)$ is easily computed.

Numerical Evaluation of $P_{BEF}(i, n_b)$

To compute $P_{BEF}(i, n_b)$, the quantity of $V_j(k)$ should be evaluated first, which is the probability that k states remain in the queueing system just before the arrival of packet j, given packets $j-1$, $j-2$, ..., 1 are received. The two initial conditions are:

if $j = 1$,

$$V_1(k) = V(k),\ k = 0,\ 1,\ ...,\ KN, \tag{4.17}$$

while, if $j > 1$,

$$V_j(0) = \sum_{\tau=0}^{(K-1)N} e^{-\frac{\rho_b}{N}(\tau+N)} V_{j-1}(\tau). \tag{4.18}$$

The recursive expression for $V_j(k)$ is then:

$$V_j(k) = \sum_{\tau=max(0,N-k)}^{KN-k} \frac{\rho_b}{N} e^{-\tau\frac{\rho_b}{N}} V_{j-1}(k-N+\tau). \tag{4.19}$$

Then $P_{BEF}(i, n_b)$ can be computed as:

$$\begin{aligned} P_{BEF}(1, n_b) &= \textstyle\sum_{k=(K-1)N+1}^{KN} V_1(k),\ i = 1, \\ P_{BEF}(i, n_b) &= \textstyle\sum_{k=(K-2)N+1}^{(K-1)N} \{1 - e^{-\frac{\rho_b}{N}[k-(K-2)N]}\} V_{i-1}(k),\ i > 1. \end{aligned} \tag{4.20}$$

Numerical Evaluation of $P_{FEC}(i, n_f, k)$

To compute $P_{FEC}(i, n_f, k)$, the same procedure is used as presented in the previous section to find $P_{BEF}(i, n_f)$. Similarly, a numerical method is used to compute the quantity $F_i(l, n_f - i)$ in (4.15) according to

$$\begin{aligned} F_i(l, n_f - i) = \sum_{k=(K-1)N+1}^{KN} V_i(k|i) [& \sum_{\tau=0}^{k} \frac{\rho_f}{N} e^{-\tau\frac{\rho_f}{N}} P_{k-\tau}^a(l, n_f - i) \\ & + e^{-\rho_f \frac{k}{N}} P_0^a(l, n_f - i)] \end{aligned} \tag{4.21}$$

where $P_i^a(j, n_f)$ can be computed by the method described in Appendix B. The quantity $V_i(k|i)$ in (4.16) is computed as:

$$V_i(k|i) = \frac{V_i(k)}{\sum_{(K-1)N+1}^{KN} V_i(k)},\ (K-1)N + 1 \le k \le KN. \tag{4.22}$$

4.4 Numerical Results And Discussion

In this section we present some selected numerical results to compare the efficacy of these three transport schemes for progressively encoded sources under various network conditions. First, several examples are provided to evaluate $P(j, n_u)$, $P_{BEF}(i, n_b)$ and $P_{FEC}(i, n_f, k)$ using the numerical methods described in Section 4.3. Secondly, the BEF, FEC and ULP transport schemes are evaluated and compared based on an idealized progressively encoded Gaussian source model. Finally, the evaluation and comparison of these three transport schemes are conducted on standard video sequences compressed using an MPEG-4 FGS encoder.

Before actually presenting the results, specific assumptions for each transport scheme are described. In the BEF transport scheme, the progressively encoded bit-stream is directly capsulated into $n_b = 16$ packets and then sent out to the network in the order of decreasing importance. In the FEC transport scheme, 16 information packets are generated and sent out to the network in the order of decreasing importance exactly as in the BEF scheme. Then $RS(n_f, 16)$ FEC codes are applied to protect these information packets and the resulting codewords are subsequently sent to the network. In particular, we consider an optimum FEC transport scheme with n_f chosen to minimize the end-to-end distortion. In the ULP transport scheme, n_u is also fixed at 16. The payload of each packet in all transport schemes is set to 400 bytes. Including the IP/UDP/RTP header, the length of a packet is 440 bytes, which is somewhat less than the Internet de facto standard maximum transmission unit (MTU) of 576 Bytes.

The $M/D/1/K$ queue described in Section 4.3 is used to model the network. Again, ρ_u, ρ_b and ρ_f are the corresponding effective network loads in the ULP, BEF and FEC schemes, respectively. In particular, for the assumed operating conditions

we have $\rho_u = \rho_b$ and $\rho_f = \rho_b/R_c$, where $R_c = 16/n_f$. In what follows, we will let ρ denote the common network load in the ULP and BEF transport schemes, and $\rho^{'}$ denote the effective network load in the FEC transport scheme with $\rho^{'} = \rho/R_c$.

4.4.1 $P(j, n_u)$, $P_{BEF}(i, n_b)$ and $P_{FEC}(i, n_f, k)$

Based on the analysis in Section III, important quantities, such as $P(j, n_u)$, $P_{BEF}(i, n_b)$ and $P_{FEC}(i, n_f, k)$, can be numerically evaluated. For example, the block error density function $P(j, n_u)$ is the key to optimized ULP allocation. Figure 4.3 illustrates the block error density function $P(j, n_u)$ for the $M/D/1/K$ queue with $K = 3$. The load ρ is varied from 0.6 to 1.2. From this figure, it can be seen that as ρ increases, more packets are likely to be lost in a block.

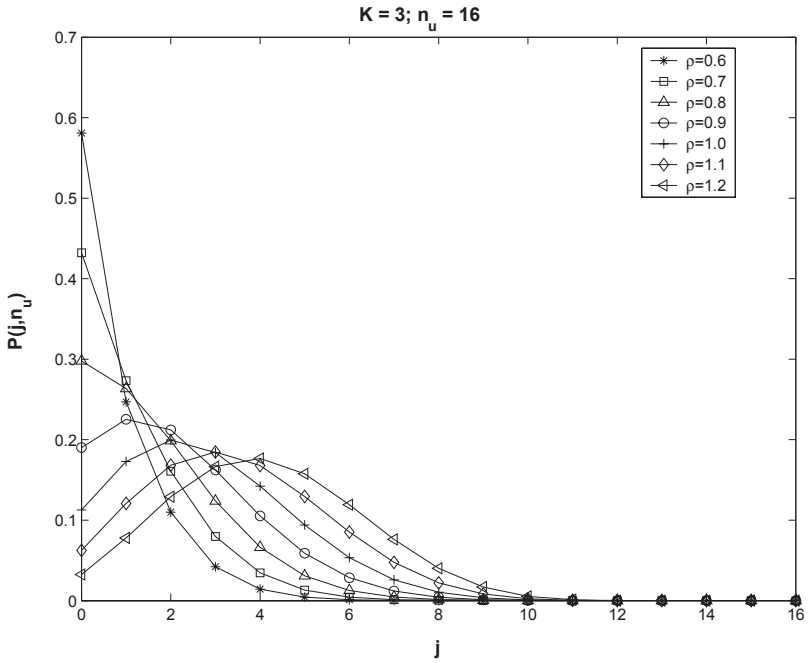

Figure 4.3: $\mathbf{P(j, n_u)}$ The block error density function $P(j, n_u)$ for different network loads ($K = 3, n_u = 16$).

Using the methods presented in Sections III-C and D, we can likewise compute

$P_{BEF}(i, n_b)$ and $P_{FEC}(i, n_f, k)$ and evaluate the performance using either BEF or FEC transport. However, as noted previously, sending FEC parity packets will also increase the network load, which can aggravate packet losses. Figures 4.4 and 4.5 illustrate this phenomenon. In these two figures, $P_{BEF}(i, 16)$ and $P_{FEC}(i, 22, 16)$ are compared for different values of K and ρ. When using FEC with an $RS(22, 16)$ code, the effective network load will increase to $\rho^{'} = (22/16)\rho$. In Fig. 4.4 with $K = 3$ and $\rho = 0.7$, because the load ρ is relatively light, sending FEC packets will not further aggravate packet losses significantly. Hence, the curve of $P_{FEC}(i, 22, 16)$ is below the curve of $P_{BEF}(i, 16)$, which indicates that the $RS(22, 16)$ code can efficiently decrease the end-to-end distortion in this case. However, if the load ρ is relatively high, as in Fig. 5 with $\rho = 1.1$ and $K = 3$, using FEC will degrade the transport performance. By using the $RS(22, 16)$ code, in this case, the probabilities of losing the first several packets increase severely. For example, the probability of losing packet 1 is increased from 15% to 28.3%. Because of the property of progressive coding, the probability that no packet is decodable in this case is 28.3%. Hence, this will lead to a poorer transport efficacy.

4.4.2 Transport of a Scalably Encoded Independent Gaussian Source

Assume a progressively encoded message is generated for a source consisting of 176×144 symbols, which is the size of a QCIF video frame. These symbols are assumed independent and identically distributed (i.i.d.), zero-mean, unit-variance Gaussian random variables. This source is assumed progressively encoded with the cumulative decoding rate denoted as R, which means successful transmission and reception of the first R bits per sample. We assume the achievable distortion-rate

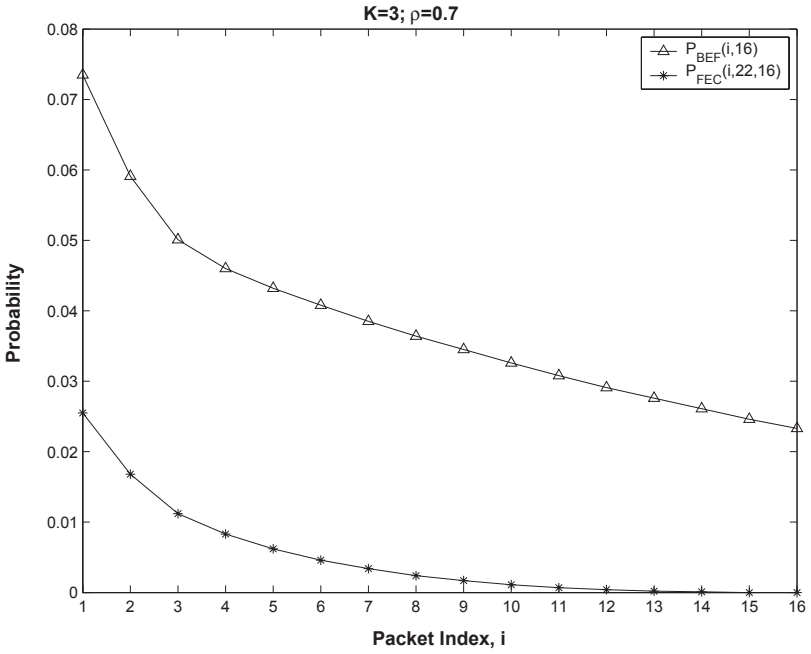

Figure 4.4: Comparison of $P_{BEF}(i, 16)$ and $P_{FEC}(i, 22, 16)$ $(K = 3, \rho = 0.7)$.

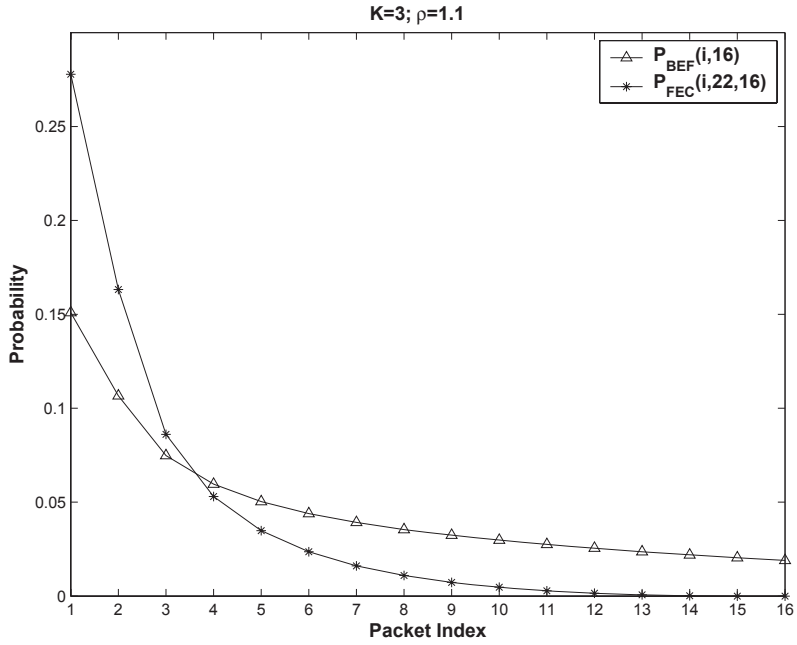

Figure 4.5: Comparison of $P_{BEF}(i, 16)$ and $P_{FEC}(i, 22, 16)$ $(K = 5, \rho = 1.1)$.

region is described by:

$$D \geq 2^{-2R}, \ R \geq 0. \tag{4.23}$$

In what follows, we assume operation on the lower boundary of the distortion-rate (DR) achievable region for evaluation and optimization[7]. We emphasize that this lower bound on performance may not be achievable with existing coding schemes, but the intent of this subsection is not to accurately model the operational DR function but, rather, to evaluate the relative efficacy of the different transport schemes under identical, although idealized conditions. However, the analytic approach presented in the preceding sections is not based on this particular DR function and can be applied to any other appropriate DR function. In Subsection 4.4.3 below, the evaluation and optimization methods presented in Section III are applied to standard video sequences compressed using an MPEG-4 FGS codec.

FEC Transport Scheme

In Sections 4.2.3 and 4.3.4, an approach was formulated to evaluate the efficacy of FEC transport for progressively encoded bitstreams. Based on this approach, the FEC coding can be optimized to minimize the expected distortion. In Fig. 4.6, examples are presented to demonstrate the evaluation and optimization of the FEC transport schemes. As indicated previously, FEC transport fixes the number of information packets and varies the FEC coding block size from 16 to 32. In this figure, K is set to 3 and the network load ρ is fixed at 1.0. It should be recalled, however, the effective network load for the FEC transport scheme is $\rho^{'} = \rho/R_c$. The channel coding rate R_c is indicated along the horizontal axis and, at the same time, the corresponding RS codes are also indicated in the figure.

The distortion performance as a function of R_c in Fig. 4.6 exhibits a U-shaped

[7]In essence, we are assuming the source is successively refinable [117].

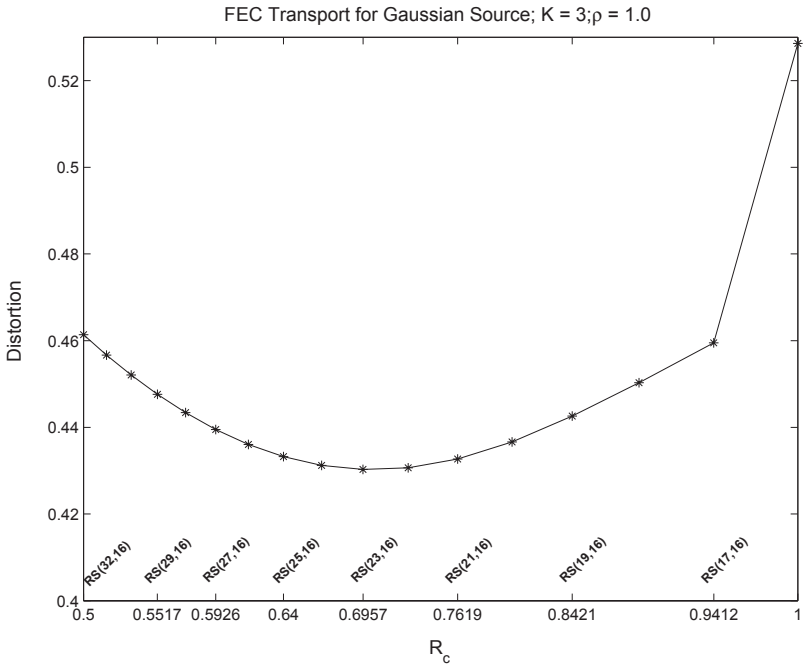

Figure 4.6: Distortion versus R_c for FEC Transport; Gaussian source ($K = 3, \rho = 1.0$).

characteristic, which implies that by carefully selecting FEC channel codes, the expected distortion can be minimized. The reason for the behavior illustrated in Fig. 4.6 is rather obvious. As R_c decreases below the point of minimum distortion, corresponding to increasing FEC overhead, network congestion is exacerbated as a result of the increased effective load $\rho^{'} = \rho/R_c$. The distortion then increases as a result of the associated increased packet losses. Likewise, as R_c increases beyond the point of minimum distortion, corresponding to decreasing FEC overhead, the network congestion is reduced, but the FEC coding is too weak to combat the residual packet losses and the distortion again increases. As indicated in Fig. 4.6, the optimum choice of FEC code in this case is the $RS(23, 16)$ code corresponding to $R_c = 16/23 = 0.6957$. In what follows, the results for the FEC transport scheme are optimized by choosing the RS code that minimizes the distortion.

Performance Comparison of ULP, BEF and Optimal FEC Transport Schemes

In Fig.'s 4.7-4.10, the performance of BEF, optimal FEC and ULP transport schemes are evaluated and compared under different operating conditions corresponding to choice of $K = 3,\ 5,\ 7$ and 10, respectively. In all figures, the network load ρ is varied from 0.6 to 1.2 where, again, the effective load is $\rho^{'} = \rho/R_c$ for FEC transport. For each choice of ρ, the optimum FEC codes for FEC transport are indicated in the figures. In particular, it should be noted that the BEF transport scheme is a special case of the FEC transport scheme with $R_c = 1$, i.e., use in this case of a $RS(16, 16)$ code.

In Fig. 4.7, corresponding to a buffer size $K = 3$, the optimized FEC scheme is clearly always superior to the BEF scheme and, in fact, is superior to the ULP scheme for small network loads up to approximately $\rho = 0.9$. As ρ increases beyond this point, the ULP scheme provides the best performance. As K increases, the advantages of the FEC scheme over BEF diminishes. For example, in Fig. 4.8 corresponding to $K = 5$, the FEC scheme provides improved performance over BEF up to approximately $\rho = 1$ and beyond that point the performance of the two schemes is identical, i.e., the optimum choice of FEC code is the $RS(16, 16)$ code. The reason for this is that increasing K reduces the packet losses sufficiently that the BEF scheme is able to cope with the residual losses whereas use of FEC would only further increase these losses due to the resulting increase in network congestion and this would cause further deterioration in performance. This is further illustrated in Fig.'s 4.9 and 4.10 corresponding to $K = 7$ and $K = 10$, respectively. In Fig. 4.9, the FEC scheme for $K = 7$ offers no advantage over BEF for ρ beyond approximately 0.8 while in Fig. 4.10 for $K = 10$, there is no advantage to FEC over BEF for any value of ρ. In all cases, however, the ULP scheme offers performance advantages over either FEC or

BEF schemes but only for large values of ρ with the crossover point increasing with K.

The reason for the ULP scheme exhibiting poorer performance than either BEF or optimal FEC transport for small values of ρ is a result of the algorithm used to allocate parity symbols while maintaining $n_b = n_u = 16$. Recall that in the ULP allocation scheme parity symbols have to be allocated based on the importance of the different parts or layers for all values of ρ. This allocation results in the pruning of some of the least important information symbols so that the actual transmitted information symbols in less than that in either the optimal FEC or BEF schemes resulting in increased distortion for small packet losses. For small values of ρ, residual packet losses are relatively light and optimum FEC can perform well despite the resulting increase in packet losses due to added congestion. This is illustrated, for example, in Fig. 4.7 for $K = 3$ where the ULP scheme provides improved performance over optimum FEC only for ρ larger than approximately 0.88. Similar behavior is illustrated in Fig. 4.8 for $K = 5$ where now even the BEF scheme provides slightly improved performance over ULP for small ρ. As K increases, the advantage of ULP over either BEF or optimal FEC exists only for relatively large values of ρ. In particular, in Fig. 4.9 for $K = 7$, the crossover point is $\rho = 0.9$ and increases to $\rho = 1.0$ in Fig. 4.10 for $K = 10$.

4.4.3 Transport of Fine Granularity Scalability (FGS) Video

The treatment of an independent Gaussian source under idealized assumptions in the preceding subsection provides a useful context for evaluation of the relative efficacy of the three transport schemes under consideration. In the present subsection we provide corresponding results for a real-world scalable source encoding scheme.

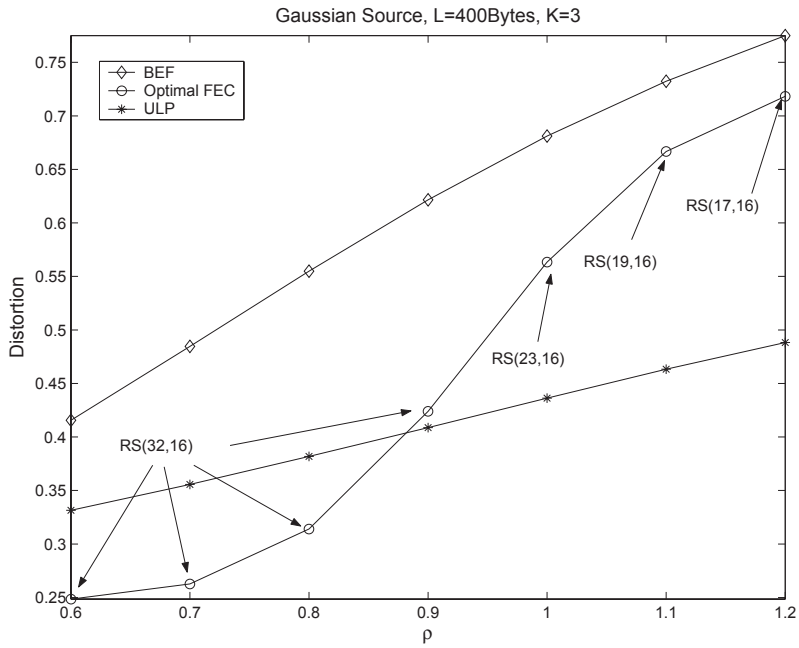

Figure 4.7: Comparison of ULP, best-effort and optimal FEC transport; Guassian source ($K = 3$).

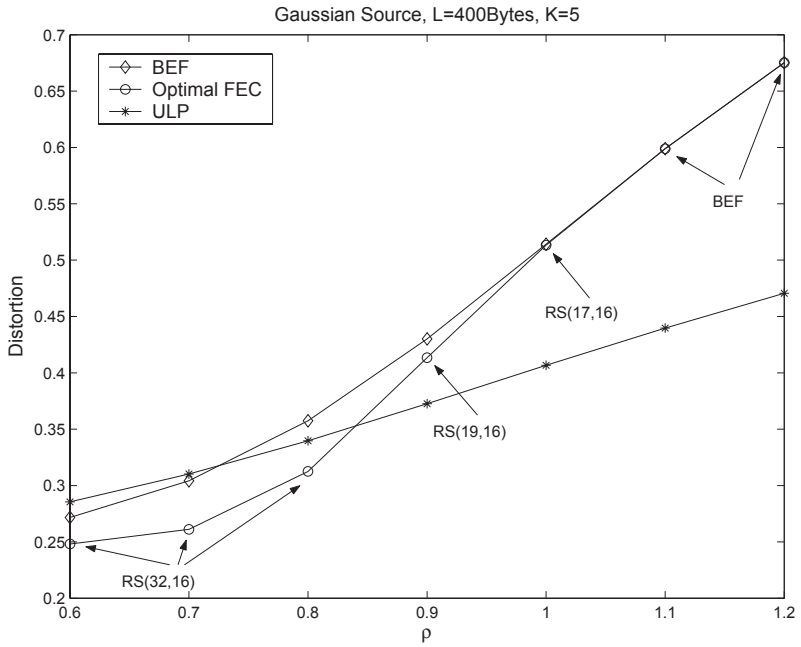

Figure 4.8: Comparison of ULP, best-effort and optimal FEC transport; Gaussian source ($K = 5$).

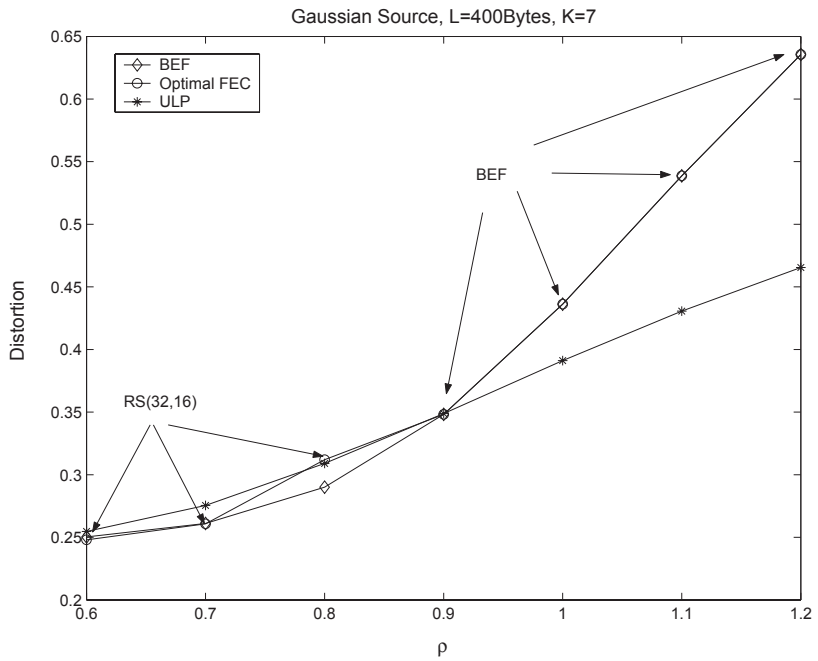

Figure 4.9: Comparison of ULP, best-effort and optimal FEC transport; Gaussian source ($K = 7$).

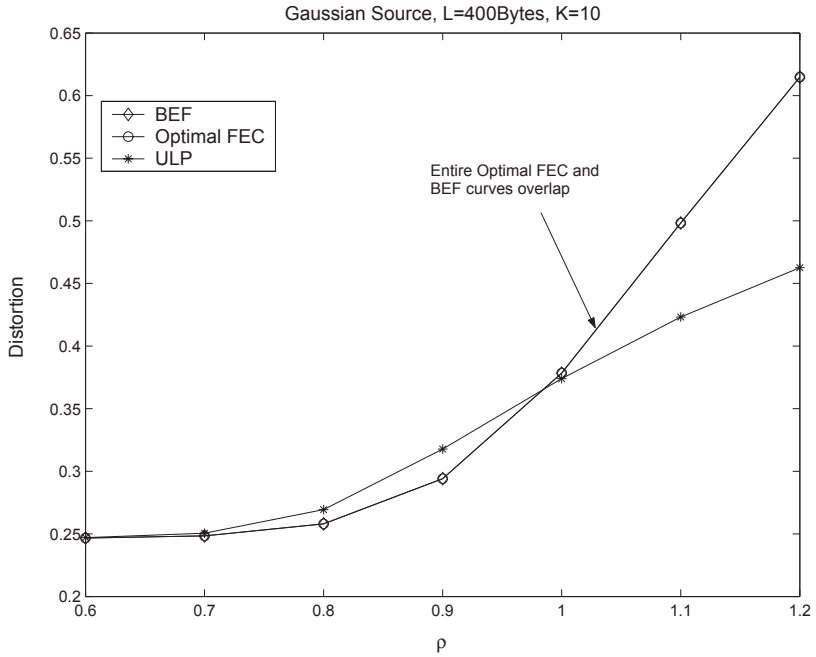

Figure 4.10: Comparison of ULP, best-effort and optimal FEC transport; Gaussian source ($K = 10$).

In particular, we make use of an MPEG-4 FGS encoder with average PSNR as the performance measure and demonstrate the relative efficacy of the three transport schemes through simulation results. In all cases, the MPEG-4 FGS encoder compresses representative video sources into two layers: a base layer (BL) and a single enhancement layer (EL). The test sequences used are the QCIF "Susie", "Container" and "Foreman" sequences at 10 frames per second with GOP structure I-P-P-P ... and a GOP length of 30. These sequences provide representative video sources with increasing levels of relative motion. In all simulation results the quantization parameter (QP) is set to 27 for both the BL and the EL. The operational rate-PSNR characteristics in each case are determined by decoding the bitstreams truncated at different points of the FGS EL.

FEC Transport Scheme

The performance of average PSNR versus channel coding rate R_c for the reconstructed QCIF "Container" sequence is illustrated in Fig. 4.11 with $K = 3$ and $\rho = 0.9$. Again, $RS(n, k)$ codes are used as in the preceding subsection with k fixed at 16 and n varying from 16 to 32. Recall $R_c = 1$, corresponding to the $RS(16, 16)$ code, represents the BEF scheme. As in Fig. 4.6, the corresponding RS codes are illustrated along the R_c axis. The behavior in Fig. 4.11 exhibits a clear inverted U-shaped characteristic, completely analogous to Fig. 4.6, indicating an optimum choice for R_c. In this case, the optimum value occurs at $R_c = 0.6667$, corresponding to the $RS(24, 16)$ code. For reasons provided previously for the idealized Gaussian source, this behavior is entirely to be expected. For R_c below the point of maximum PSNR, the increasing FEC overhead exacerbates the network congestion causing increased packet losses leading to a decrease in PSNR. Likewise, for R_c above this optimum point, the network congestion is reduced but the FEC is too weak to be effective. As in the case of the Gaussian source, in the results which follow the FEC is chosen to optimize the PSNR.

Performance Comparison of ULP, BEF and Optimal FEC Transmission Schemes

In this subsection, the approach formulated in Sections 4.2.3 and 4.3.4 is applied to the three test sequences to evaluate and compare the efficacy of BEF, optimal FEC and ULP transport schemes. In Fig.'s 4.12 and 4.13, the low-motion "Susie" sequence is used to compare these three transport schemes with K set to 3 and 5, respectively. The comparison based on the increased motion "Container" sequence is provided in Fig.'s 4.14, 4.15, 4.16 and 4.17, where K is set to 3, 5, 7 and 10, respectively. Similarly, Fig.'s 4.18 and 4.19 present the comparison based the relatively high-motion

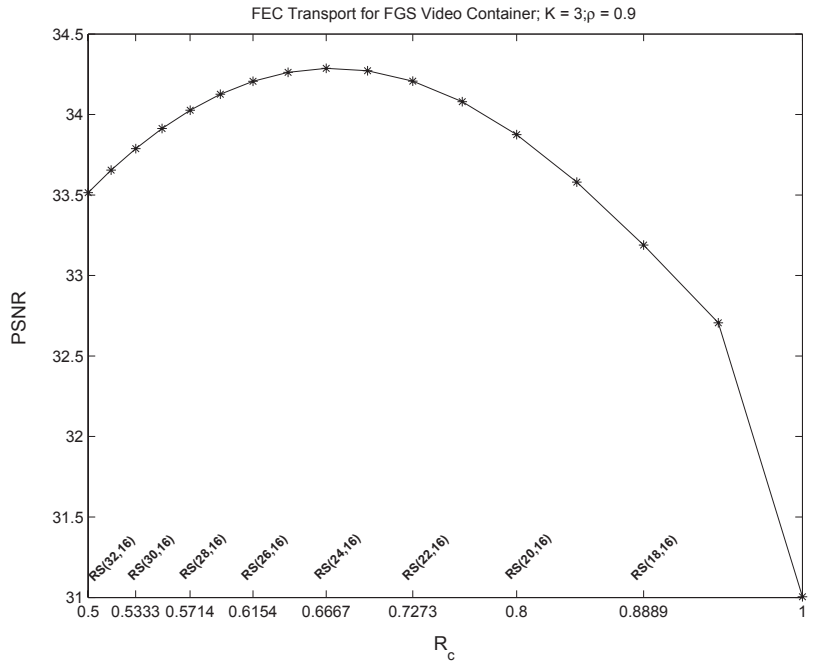

Figure 4.11: PSNR versus R_c for FEC Transport; Container sequence ($K = 3, \rho = 0.9$).

"Foreman" sequence, where K is set to 3 and 5, respectively. The RS codes used in the optimal FEC schemes are labeled on these figures.

As can be observed, the performance of the BEF transport scheme drops rapidly in all cases as the network load ρ increases since no protection is provided in the face of the increasing packet losses. The drop in PSNR is more precipitous as the relative motion in the test sequence increases in going from the "Susie" sequence through the "Foreman" sequence. For example, in Fig. 4.12 for the "Susie" sequence with $K = 3$ the PSNR is 31.54dB at $\rho = 1$. In Fig. 4.14, for the "Container" sequence under the same conditions the PSNR is reduced to 28.84dB while in Fig. 4.18 for the "Foreman" sequence under the same conditions the corresponding PSNR has been reduced even further to 26.93dB. Also to be observed is the improvement in PSNR with increasing K as can best be observed for the "Container" sequence in Fig.'s 4.14 - 4.17 corresponding to $K = 3,\ 5,\ 7$, and 10, respectively. For example, at $\rho = 1.0$

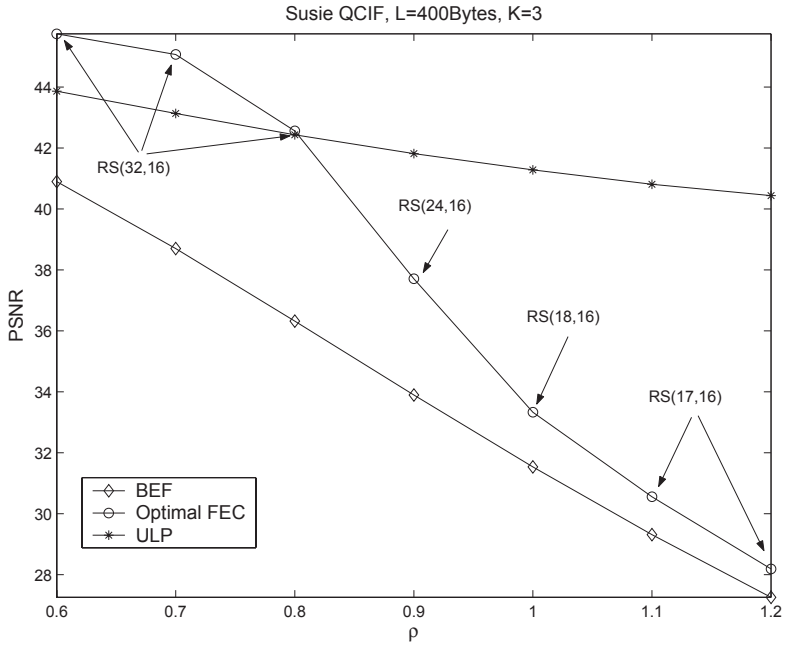

Figure 4.12: Comparison of ULP, best-effort and optimal FEC transport; Susie sequence ($K = 3$).

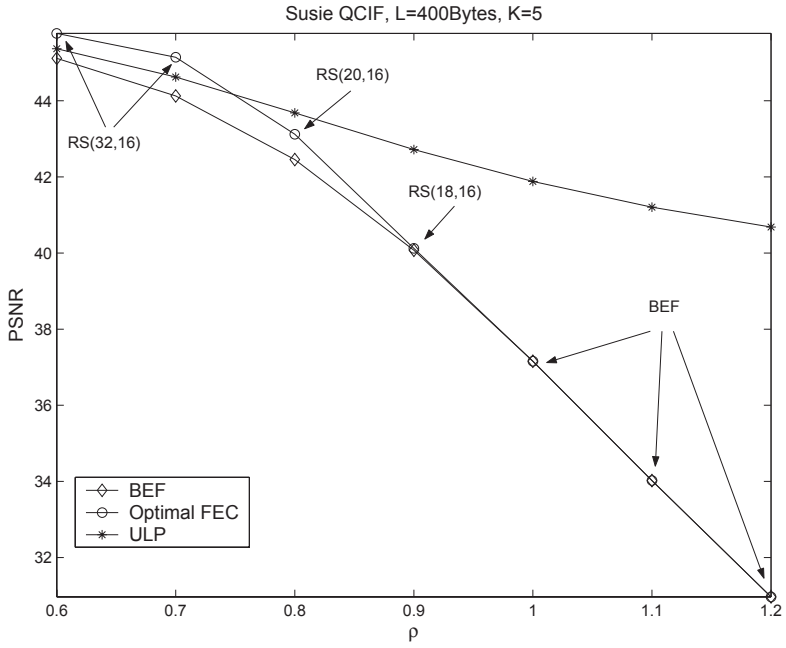

Figure 4.13: Comparison of ULP, best-effort and optimal FEC transport; Susie sequence ($K = 5$).

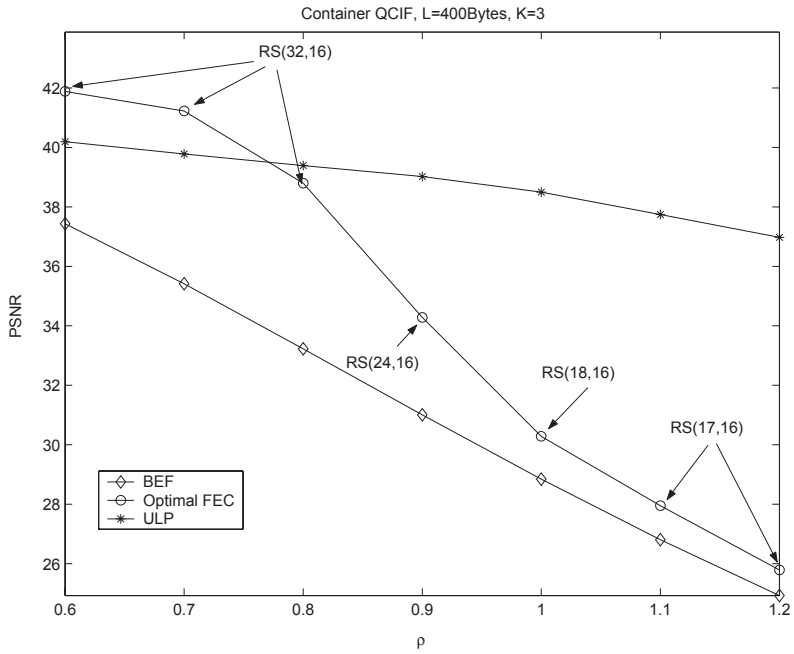

Figure 4.14: Comparison of ULP, best-effort and optimal FEC transport; Container sequence ($K = 3$).

the corresponding PSNR values for BEF transport are 28.84dB, 34dB, 36.33dB and 38.05dB, respectively.

Examination of Fig.'s 4.12 - 4.19 also illustrate similar behavior in the relative performance of the optimum FEC and ULP transport schemes as observed previously for the idealized independent Gaussian source. In particular, in all cases it can be observed that the ULP scheme outperforms the optimum FEC scheme only for larger values of ρ with the crossover point increasing in K. For example, in Fig.'s 4.14 - 4.17 for the "Container" sequence with $K = 3,\ 5,\ 7$, and 10, respectively, the crossover point in Fig. 4.14 is $\rho = 0.78$ for $K = 3$ increasing eventually to approximately $\rho = 0.9$ in Fig. 4.17 for $K = 10$. Likewise, it can be seen from this series that as K increases the optimum FEC scheme offers little or no advantage over the BEF scheme. For example, in Fig. 4.16 with $K = 7$ the optimum FEC scheme provides no advantage over the BEF scheme for ρ larger than approximately 0.8 while in Fig.

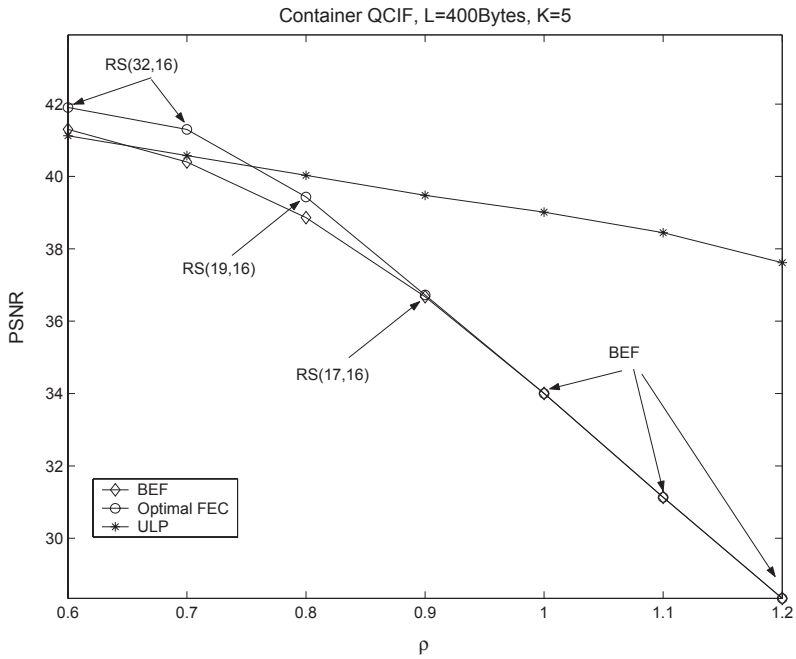

Figure 4.15: Comparison of ULP, best-effort and optimal FEC transport; Container sequence ($K = 5$).

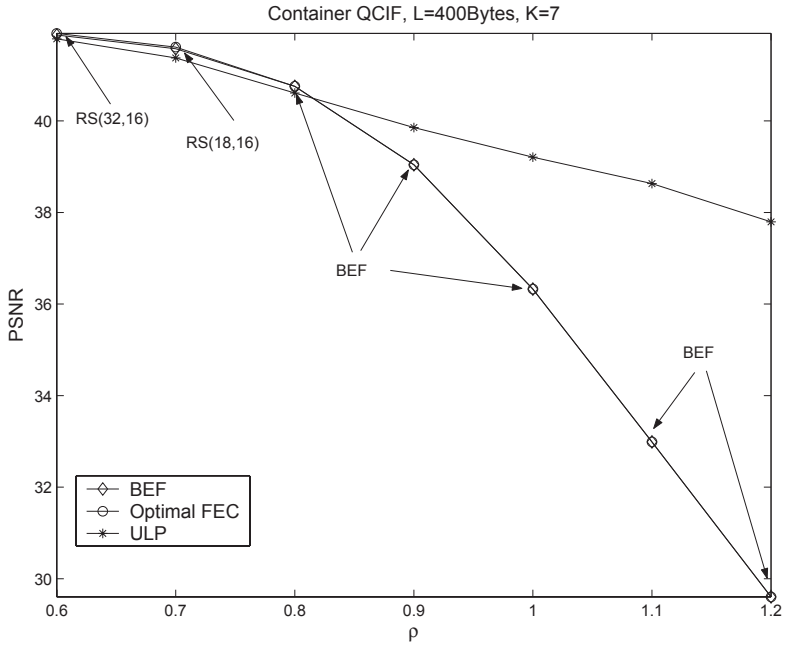

Figure 4.16: Comparison of ULP, best-effort and optimal FEC transport; Container sequence ($K = 7$).

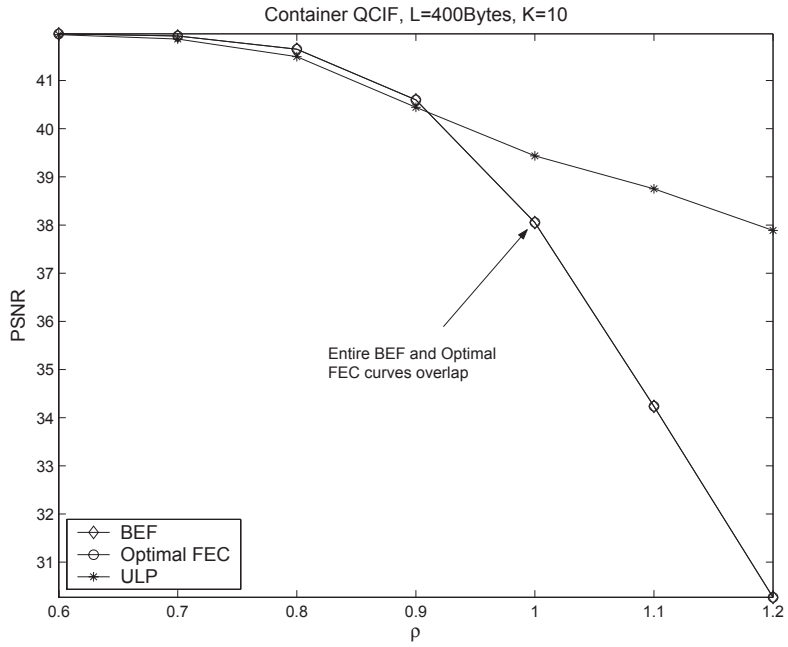

Figure 4.17: Comparison of ULP, best-effort and optimal FEC transport; Container sequence ($K = 10$).

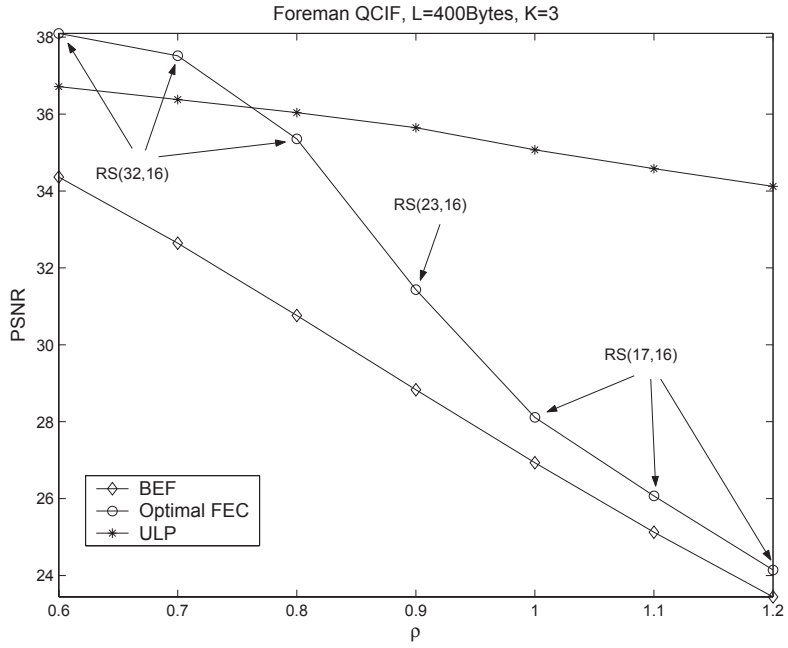

Figure 4.18: Comparison of ULP, best-effort and optimal FEC transport; Foreman sequence ($K = 3$).

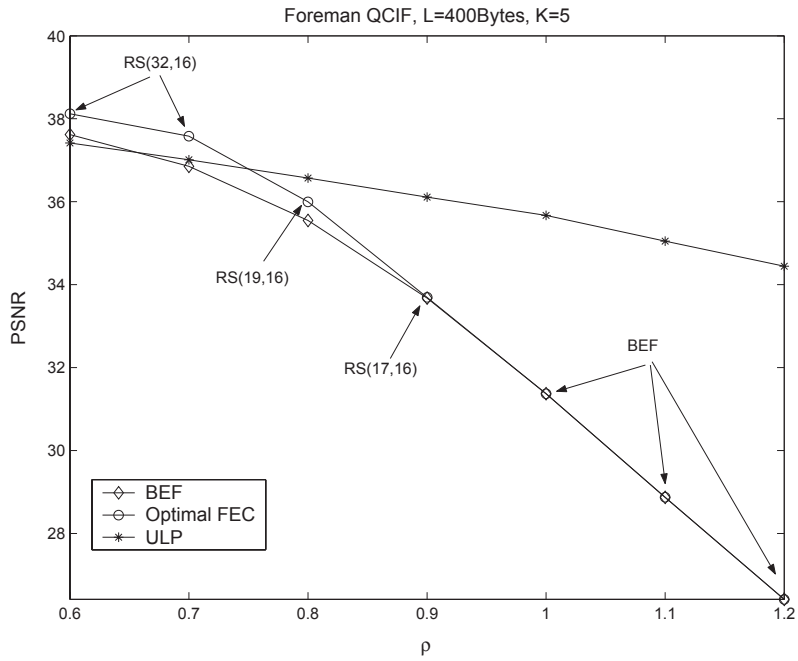

Figure 4.19: Comparison of ULP, best-effort and optimal FEC transport; Foreman sequence ($K = 5$).

4.17 for $K = 10$ there is no advantage of the optimum FEC scheme over the BEF scheme for any value of ρ.

Again, we note that in all cases the ULP scheme provides substantial performance improvement over either the optimum FEC or BEF schemes for increasing values of ρ. Although the absolute values of PSNR decrease with increasing relative motion in the test sequence, the performance improvement of the ULP scheme over either the optimum FEC or BEF transport schemes is relatively insensitive to the scene motion. For example, comparing the performance for $\rho = 1.2$ and $K = 5$ in Fig.'s 4.13, 4.15 and 4.19 for the "Susie", "Container" and "Foreman" sequences, respectively, the performance improvement of the ULP scheme over either the optimum FEC or BEF schemes is approximately 8-9dB. Overall, the results indicate that the ULP scheme provides a relatively robust solution to the transport of scalably encoded sources over congested networks. There is substantial performance advantage at high loads

while, although there is some relative performance disadvantages at light loads, the performance disadvantage is quite small.

4.5 Summary and Conclusions

In this Chapter, we investigated three transport schemes, BEF, optimal FEC and ULP schemes, for transport of progressively encoded bitstreams over congested networks. Firstly, we provide three explicit queueing analysis approaches to evaluate the efficacy of these three transport schemes. Secondly, we quantitatively analyze and compare the efficacy of these three schemes using both an idealized Gaussian source and practical MPEG-4 FGS encoded video sources. The selected numerical results show that under comparable operating conditions the ULP scheme outperforms the other two when the network load is relatively high. However, the optimal FEC scheme is more efficient when the network load is relatively low. The results indicate, among other things, that care should be used in selecting an appropriate transport scheme for progressively encoded, or scalable, sources transmitted over congested networks. In particular, this selection should be made based on knowledge of prevailing network congestion conditions. However, the results suggest that the ULP scheme provides a relatively robust solution in the sense that there is substantial performance advantage at high loads while relatively small performance disadvantage at light loads.

CHAPTER 5

Conclusions and Future Work

In this work, the challenges of video transport over best-effort networks have been addressed using several innovative techniques, such as adaptive MCP, ULP, MDC, MPT and interleaving. The efficacy of ULP transport of scalably encoded sources is evaluated and compared with best-effort and FEC transport.

A video transport system which integrates MC-FGS video source coding, ULP and interleaving schemes is proposed. Under an imposed delay constraint, the source coder, ULP allocation and interleaver are jointly optimized. The simulation results demonstrate that the proposed system can substantially improve the video transport efficacy and the improvement in mean PSNR achieved increases as the average burst length of the network losses increases or the delay constraint is relaxed. It is also shown that even if interleaving does lead to more packet losses due to late arrivals, the performance of the proposed transport system is still superior to the comparison system without interleaving. This is because optimal interleaving can effectively reduce the video distortion through mitigating the effects of the otherwise bursty losses.

Furthermore, a unified approach incorporating adaptive motion compensation prediction, multiple description coding and unequal multiple path allocation is developed to improve both the robustness and error resilience property of the video coding and

transmission system, while simultaneously improving the delivered video quality. The simulation results show: The simulation result shows that: the proposed scheme can result in substantial performance improvement. Based on the simulations using the Gilbert-Elliot model, some of the packets should still be sent along the path with higher packet loss rate because employing multiple paths can effectively reduce the effects of bursty losses; even using paths with similar packet loss rate, the packets should not be allocated to these paths equally if their average burst durations are not similar. Based on the simulations using $M/D/1/K$ queue, the performance gain improves as K increases; more packets are allocated to the better path when the network load of both paths is light while optimal traffic allocation converges to equal traffic allocation as the network load increases.

To analytically investigate the efficacy of error resilient transport schemes for progressively encoded sources, including unequal loss protection, best-effort and FEC transport schemes, we first provide a comprehensive queueing analysis. Armed with the results, these three transport schemes are quantitatively evaluated and compared. The selected numerical results show that under comparable operating conditions the ULP scheme outperforms the other two when the network load is relatively high. However, the optimal FEC scheme is more efficient when the network load is relatively low. The results indicate, among other things, that care should be used in selecting an appropriate transport scheme for progressively encoded, or scalable, sources transmitted over congested networks. In particular, this selection should be made based on knowledge of prevailing network congestion conditions. However, the results suggest that the ULP scheme provides a relatively robust solution in the sense that there is substantial performance advantage at high loads while relatively small performance disadvantage at light loads.

Delay constraint play an important role in the design of real-time video transport systems. In Chapter 2, a optimal interleaving approach is developed under a delay-constraint. A future direction for extending the work presented in previous Chapters can be investigation on the impact of imposed delay constraint in Chapter 4.

APPENDIX A

V_i for the $M/E_r/1/K$ Queue

The distribution of the remaining workload in the $M/D/1/K$ queue, $V(x)$, can be approximated by the steady state distribution of the $M/E_r/1/K$ queue as $r \to \infty$. An $M/E_r/1/K$ queue system is illustrated in Fig. A.1. The packet arrival process is a Poisson process with rate λ, and the system can hold up to K packets. The service time is an r-stage Erlangian distribution, which can be considered as a sum of r exponential random variables. In the figure, the large oval represents the service facility, whose service rate is μ. The internal structure of this service facility is revealed as a series of r small ovals. Each of them represents a single exponential service facility (stage) whose service rate is $r\mu$. In this system, only when a packet departs by exiting from the right side of the large oval service facility, a new packet may then enter from the left side and proceed one stage at a time through the sequence of r stages.

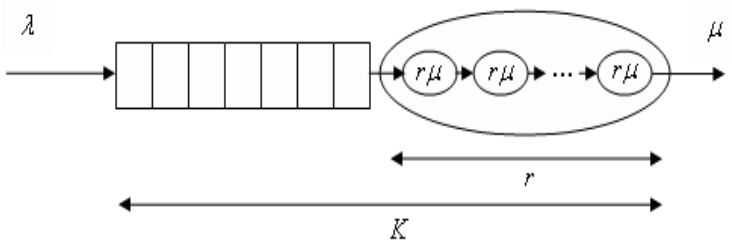

Figure A.1: $M/E_r/1/K$ queue system.

To analyze the $M/E_r/1/K$ queue, define the state variable as the total number of

service stages yet to be completed by all packets in the system at the time the state is described. The state-transition-rate diagram can be depicted as in Fig. A.2. Each oval represents a state. The internally labeled number indicates the number of stages left in the total system. Since a new arrival packet will increase the number of stages yet to be completed by r, the system state will transit to a state which is r positions to its right if a new packet arrives. But a completion of one stage will decrease the stage number by 1 and, as a result, the system state will transit to its left neighbor if a stage is completed.

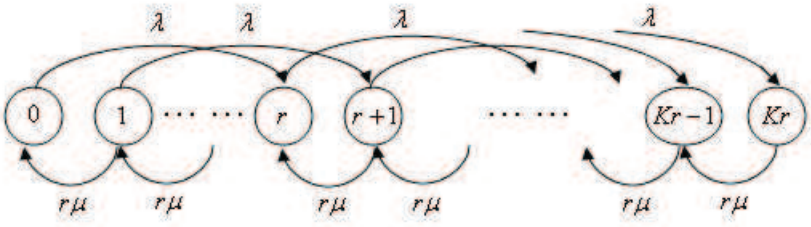

Figure A.2: State-transition-rate diagram for number of stages: $M/E_r/1/K$.

Furthermore, define V_i as the equilibrium probability that a total of i stages are left in the system, which is a discretized version of $V(x)$. Based on Fig. A.2, the system state equations can be derived as:

$$\begin{aligned}
rcl\lambda V_0 &= r\mu V_1, \\
(\lambda + r\mu)V_{i-1} &= r\mu V_i,\ 2 \le i \le r, \\
(\lambda + r\mu)V_{i-1} &= \lambda V_{i-r-1} + r\mu V_i,\ r+1 \le i \le (K-1)r+1, \\
r\mu V_{i-1} &= \lambda V_{i-r-1} + r\mu V_i,\ (K-1)r+2 \le i \le Kr, \\
r\mu V_{Kr} &= \lambda V_{(K-1)r}.
\end{aligned} \tag{A.1}$$

Also, we have

$$\sum_{i=0}^{Kr} V_i = 1. \tag{A.2}$$

Regroup these equations and write them into a matrix form as

$$\begin{pmatrix}
\rho & -r & 0 & & & \cdots & & \cdots & & & 0 \\
0 & \rho+r & -r & 0 & & & & & & & \\
\vdots & \ddots & \ddots & \ddots & \ddots & & & & & & \\
0 & \cdots & 0 & \rho+r & -r & 0 & & & & & \\
-\rho & 0 & \cdots & 0 & \rho+r & -r & 0 & & & & \vdots \\
0 & \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots & & & \vdots \\
 & \ddots & -\rho & & & 0 & \rho+r & -r & 0 & & \\
\vdots & & & -\rho & & & 0 & r & -r & \ddots & \\
 & & & & \ddots & \ddots & & \ddots & \ddots & \ddots & 0 \\
0 & & \cdots & & 0 & -\rho & 0 & \cdots & 0 & r & -r \\
1 & 1 & 1 & \cdots & & & \cdots & & & 1 & 1
\end{pmatrix}
\begin{pmatrix} V_0 \\ V_1 \\ \\ \\ \\ \\ \vdots \\ \vdots \\ \\ \\ V_{Kr} \end{pmatrix}
=
\begin{pmatrix} 0 \\ 0 \\ \\ \\ \\ \\ \vdots \\ \vdots \\ \\ \\ 1 \end{pmatrix}, \tag{A.3}$$

where ρ denotes the system load, which is λ/μ.

Clearly, the equilibrium probability $V_i,\ i = 0, 1, ..., Kr$ can be computed from this matrix form equation. As r increases to infinity under appropriate scaling of the per stage service rate, the Erlangian distribution of service time will go to a unit impulse function, which means the service time is a deterministic number, $1/\mu$. So, in this case, the $M/E_r/1/K$ queue turns into a $M/D/1/K$ queue.

APPENDIX B

$P(j,n)$ for the $M/D/1/K$ Queue

In [115], an analysis of $P(j,n)$ is presented for an $MMPP/D/1/K$ queueing system, where $MMPP$ denotes the Markov-modulated Poisson arrival process. In this Appendix, we restrict the arrival process to be a single Poisson process and provide a simplified analysis of the resulting $M/D/1/K$ queue.

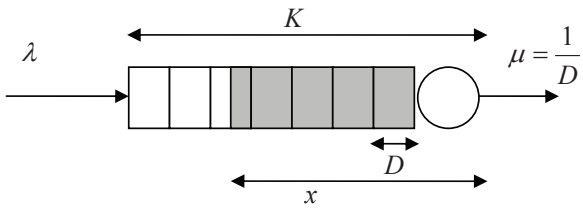

Figure B.1: $M/D/1/K$ queue system

As depicted in Fig. B.1, the packet arrival process is assumed a Poisson process with rate λ, and the system can hold up to K packets. The service rate is a deterministic value μ and the service time is $D = 1/\mu$ time units per packet. The network load is $\rho = n/\mu$. We assume a FIFO service policy is employed in this queueing system.

Because the service time for each packet is a deterministic value D, the remaining workload can be measured in time units. Let x denote the remaining workload in the system. The remaining workload x means that all the workload of the system will be cleared after time x if there are no additional packet arrivals during this period. Since the service time of a packet is D, the arrival of a packet in the absence

of a departure will increase the workload x by D. The maximum workload of the queueing system is KD since the system can hold up to K packets. A necessary condition for receiving a packet is that the workload just before the arrival of this packet is no larger than $(K-1)D$, otherwise the packet will be lost. Define $V(x)$ as the probability distribution of the remaining workload in the queueing system. As described in what follows, $V(x)$ can be approximated by the steady-state distribution of the $M/E_r/1/K$ queue [8] under appropriate scaling as $r \to \infty$. This follows since the service facility of the $M/E_r/1/K$ queue can be considered to consist of a cascade of r stages, with each having identical exponential service time distribution. To keep the mean service rate constant at the value $\mu = 1/D$, the service rate of each stage is scaled by a factor of r to $r\mu$. As $r \to \infty$, the service time of the $M/E_r/1/K$ queue then converges to a constant value D. More details on the $M/E_r/1/K$ queue can be found in [116]. The determination of $V(x)$ using this approach is provided in Appendix A.

B.1 The Block Error Density Function, $P(j,n)$, for the M/D/1/K Queue

$P(j,n)$, $n \geq 1, 0 \leq j \leq n$, denotes the probability that j packets are lost in a block of n consecutive packets. Again, with $V(x)$ the probability distribution of the remaining workload in the queueing system, $P(j,n)$ can be computed as:

$$P(j,n) = \int_0^{KD} V(x) P_x^a(j,n)\, dx, \tag{B.1}$$

where $P_x^a(j,n)$ denotes the probability of j losses in a block of n packets given that the remaining workload in the system is x just before the arrival of the first packet

[8] Here, E_r represents the Erlang-r distribution for packet service time.

of this block.

The quantity $P_x^a(j,n)$ can be computed recursively, starting with the initial conditions for $n = 1$:

i) if $0 \leq x \leq (K-1)D$,

$$P_x^a(j,1) = \begin{cases} 1, & j = 0 \\ 0, & j > 0 \end{cases} \tag{B.2}$$

ii) if $x > (K-1)D$,

$$P_x^a(j,1) = \begin{cases} 0, & j = 0, j \geq 2 \\ 1, & j = 1 \end{cases} \tag{B.3}$$

The justification for the initial conditions represented by (B.2) and (B.3) is relatively straightforward: if the system has space in its buffer, the arriving packet will be held in the buffer, otherwise, the packet will be lost.

For $n \geq 2$ the recursive equations are, as in [115]:

i) if $0 \leq x \leq (K-1)D$,

$$P_x^a(j,n) = \int_0^{x+D} f(t) P_{x+D-t}^a(j,n-1)\, dt + \int_{x+D}^{\infty} f(t) P_0^a(j,n-1)\, dt, \tag{B.4}$$

ii) if $x > (K-1)D$,

$$P_x^a(j,n) = \int_0^{x} f(t) P_{x-t}^a(j-1,n-1)\, dt + \int_{x}^{\infty} f(t) P_0^a(j-1,n-1)\, dt, \tag{B.5}$$

where t denotes the interarrival time between the first packet and the second packet of the block, and $f(t)$ denotes the probability density function of the packet interarrival-time. If $0 \leq x \leq (K-1)D$, as in the first case above, the system has space to hold

the first packet of the block. So, the first packet will be held in the buffer and the workload will increase by D units. To have j lost packets out of the n packets, j packets must be lost out of the following $n-1$ packets starting from the next arrival epoch. If t, the interarrival time between the first packet and the second packet, is less than $x+D$, the remaining workload will be $x+D-t$ just before the arrival of the second packet. Otherwise, the workload will be 0. The second case treats the situation where there is not enough space in the buffer to hold a packet. So, the first packet will be dropped and the workload will not increase. Hence, to have j losses out of the block of size n, there must be $j-1$ losses among the following $n-1$ packets. At the arrival epoch of the second packet in the block, the workload will be $x-t$ if t is less than x. Otherwise, the workload will be 0.

B.2 Numerical Evaluation of $P(j,n)$

The workload distribution $V(x)$ can be approximated by the steady-state distribution of the $M/E_r/1/K$ queue in the limit of $r \to \infty$ [116], an analysis of which is included in Appendix A. First, the service time D is divided into N subintervals of length Δ, with $D = N\Delta$ and $\Delta = 1/(N\mu)$. Let r, the number of stages in the service facility of the $M/E_r/1/K$ queue, be equal to N. As N increases to infinity, the $M/E_r/1/K$ queue turns into an $M/D/1/K$ queue. Hence, If N is sufficiently large, the workload distribution $V(x)$ can be accurately approximated by the equilibrium probability P_i of the $M/E_r/1/K$ queue. This means that the discretized workload x can be represented by the remaining stages in the $M/E_r/1/K$ queue system if N is sufficiently large. So, to keep notation consistent, $V(i)$ is used to denote the equilibrium probability P_i of the $M/E_r/1/K$ queue.

The block error density function $P(j,n)$ is then approximated for large N as

$$P(j,n) = \sum_{i=0}^{KN} V(i) P_i^a(j,n). \tag{B.6}$$

The recursive computation is given below. The initial condition for $n = 1$ are:

i) if $0 \le i \le (K-1)N$,

$$P_i^a(j,1) = \begin{cases} 1, & j = 0 \\ 0, & j > 0, \end{cases} \tag{B.7}$$

and

ii) if $i > (K-1)N$

$$P_i^a(j,1) = \begin{cases} 0, & j = 0, j \ge 2 \\ 1, & j = 1. \end{cases} \tag{B.8}$$

Similarly, for $n \ge 2$, the recursive equations are:

i) if $0 \le i \le (K-1)N$,

$$\begin{aligned} P_i^a(j,n) = & \sum_{\tau=0}^{i+N} f(\tau\Delta)\Delta P_{i+N-\tau}^a(j,n-1) + P_0^a(j,n-1)\int_{i\Delta+D}^{\infty} f(t)\,dt \\ = & \sum_{\tau=0}^{i+N} \frac{\rho}{N} e^{-\tau\rho/N} P_{i+N-\tau}^a(j,n-1) + P_0^a(j,n-1) e^{-\rho(i/N+1)}, \end{aligned} \tag{B.9}$$

while,

ii) if $i > (K-1)N$,

$$\begin{aligned} P_i^a(j,n) = & \sum_{\tau=0}^{i} f(\tau\Delta)\Delta P_{i-\tau}^a(j-1,n-1) + P_0^a(j-1,n-1)\int_{i\Delta}^{\infty} f(t)\,dt \\ = & \sum_{\tau=0}^{i} \frac{\rho}{N} e^{-\tau\rho/N} P_{i-\tau}^a(j-1,n-1) + P_0^a(j-1,n-1) e^{-\rho(i/N+1)}. \end{aligned} \tag{B.10}$$

Bibliography

[1] Chan, Y. S., Cosman, P. C., and Milstein, L. B., "A multiple description coding and delivery scheme for motion compensation fine granularity scalable video," *IEEE Transactions on Image Processing*, Vol. 17, No. 8, Aug. 2008, pp. 1353–1367.

[2] Mohr, A. E., Riskin, E. A., and Ladner, R. E., "Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 6, June 2000, pp. 819–828.

[3] "Video codec for audiovisual services at p x 64 kb/s," *ITU-T Recommendation H.261, ITU-T, Version 1*, Nov 1990.

[4] "Generic coding of moving pictures and associated audio information - part 2: video," *ITU-T and ISO/IEC JTC 1, ITU-T Recommendation H.262 and ISO/IEC 13 818-2 (MPEG-2)*, 1994.

[5] "Video coding for Low bit rate communication," *IITU-T, ITU-T Recommendation H.263 version 1*, 1995.

[6] "Advanced video coding for generic audiovisual services," *ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 AVC) ITU-T and ISO/IEC JTC 1*, May 2003.

[7] "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/spart 2: video," *ISO/IEC 11172-2 (MPEG-1 Video), ISO/IEC JTC 1*, 1993.

[8] "Coding of audio-visual objects – part 2: visual," *ISO/IEC 14 496-2 (MPEG-4 Visual Version 1)*, April 1999.

[9] Symes, P., *Digital video compression*, McGraw-Hill/TAB Electronics, 2003.

[10] Richardson, I., *H.264 and MPEG-4 video compression: video coding for next generation multimedia*, Wiley, 2003.

[11] Girod, B., "The efficiency of motion-compensated predication for hybrid coding of video sequences," *IEEE Journal on Selected Areas in Communications*, Vol. 5, No. 7, Aug. 1987, pp. 1140–1154.

[12] Bolot, J., "End-to-end delay and loss behavior in the Internet," *Proceedings of ACM SIGCOMM 1993*, San Francisco, CA, Sept. 1993, pp. 289–298.

[13] Yajnik, M., Moon, S., Kurose, J., and Towsley, D., "Measurement and modelling of the temporal dependence in packet loss," *In Proceedings on IEEE Infocomm'99*, 1999, pp. 345–352.

[14] Paxson, V., "End-to-end Internet packet dynamics," No. 3, June 1999, pp. 277–292.

[15] Sanghi, D., Agrawala, A., Gudmundsson, O., Gudmundsson, O., and Jain, B. N., "Experimental assessment of end-to-end behavior on Internet," *Proceedings of IEEE INFOCOM 93*, San Francisco, CA, March 1993, p. 867887.

[16] Ramjee, R., Ramjee, R., Kurose, J., Towsley, D., and Schulzrinne, H., "Adaptive playout mechanisms for packetized audio applications in wide-area networks," *Proceedings of IEEE INFOCOM 94*, Montreal, Canada, June 1994, pp. 680–688.

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

[18] Girod, B., Kalman, M., Liang, Y. J., and Zhang, R., "Advances in channel-adaptive video streaming," *Wireless Communications and Mobile Computing*, Vol. 2, No. 6, Sept. 2002, pp. 573–584.

[19] Civanlar, M., A. Luthra, S. W., and (eds), W. Z., "Special issue on streaming video," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 3, March 2001.

[20] Chen, C., Cosman, P., Kingsbury, N., Liang, J., and (eds), J. M., "Special issue on error resilient image and video transmission," *IEEE Journal on Selected Area in Communications*, Vol. 18, No. 6, June 2001.

[21] Conklin, G., Greenbaum, G., Lillevold, K., Lippman, A., and Reznik, Y., "Video coding for streaming media delivery on the Internet," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 3, March 2001, pp. 269–281.

[22] Kurceren, R., *Joint source-channel coding approach to transport of digital video on lossy networks*, Ph.D. thesis, Rensselaer Polytechnic Institute, Troy, New York, May 2001.

[23] Yu, X., *Modeling and analysis of digital video transmission over networks using packet-level FEC*, Ph.D. thesis, University of Miami, Coral Gables, Florida, May 2007.

[24] Liang, Y., *Robust low-latency voice and video communication over best-effort networks*, Ph.D. thesis, Stanford University, Stanford, California, Aug. 2003.

[25] Wang, Y., Wenger, S., Wen, J., and Katsaggelos, A. K., "Review of error resilient coding techniques for real-time video communications," *IEEE Signal Processing Magazine*, Vol. 17, No. 4, July 2000, pp. 61–82.

[26] Wen, J. and Villasenor, J., "Reversible variable length codes for robust image and video transmission," *in Proceedings of the Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1997.

[27] Redmill, D. W. and Kingsbury, N. G., "The EREC: an error resilient technique for coding variable length blocks of data," *IEEE Transactions on Image Processing*, Vol. 5, No. 4, April 1996, pp. 565–574.

[28] Cote, G., Shirani, S., and Kossentini, F., "Optimal mode selection and synchronization for robust video communications over error prone networks," *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 6, June 2000, pp. 952–965.

[29] Moccagatta, I., Soudagar, S., Liang, J., and Chen, H., "Error-resilient coding in JPEG-2000 and MPEG-4," *IEEE Journal on selected areas in communications*, , No. 6, June 2000, pp. 899–914.

[30] Wang, Y., Reibman, A. R., and Lin, S., "Multiple description coding for video delivery," *Proceedings of IEEE*, Vol. 93, No. 1, Jan. 2005, pp. 57–70.

[31] Wang, Y. and fan Zhu, Q., "Error control and concealment for video communication: a review," *Proceedings of the IEEE*, Vol. 86, No. 5, May 1998, pp. 974–997.

[32] Wah, B. W., Su, X., and Lin, D., "A survey of error-concealment schemes for real-time audio and video transmissions over the Internet," *In Proceedings on IEEE International Symposium on Multimedia Software Engineering*, Taipei, Taiwan, Dec. 2000, pp. 17–24.

[33] Suh, J.-W. and Ho, Y.-S., "Error concealment based on directional interpolation," *IEEE Transactions on Consumer Electronics*, , No. 3, Aug. 1997, pp. 295–302.

[34] Wada, M., "Selective recovery of video packet loss using error concealment," *IEEE Journal on Selected Areas in Communications*, , No. 5, June 1989, pp. 807–814.

[35] Stuhlmuller, K., Farber, N., Link, M., and Girod, B., "Analysis of video transmission over lossy channels," *IEEE Journal on Selected Areas in Communication*, Vol. 18, No. 6, June 2000.

[36] Yu, X., Modestino, J. W., Kurceren, R., and Chan, Y. S., "A model-based approach to evaluation of the effcacy of FEC coding in combating network packet losses," *IEEE/ACM Transactions on Networking*, Vol. 16, No. 3, June 2008.

[37] Yu, X., Modestino, J. W., and bajic, I. V., "Performance analysis of the efficacy of packet-level FEC in improving video transport over networks," *IEEE International Conference on Image Processing (ICIP'05)*, Sept. 2005.

[38] McAuley, A. J., "Reliable broadband communications using a burst erasure correcting code," *in Proceedings of ACM SIGCOMM 1990*, Philadelphia, PA, Sept. 1990, p. 287306.

[39] Frossard, P., "FEC performance in multimedia streaming," *IEEE Communication Letter*, Vol. 5, No. 3, May 2001, pp. 122–144.

[40] Nonnenmacher, J., Biersack, E., and Towsley, D., "Parity-based loss recovery for reliable multicast transmission," *in Proceedings of ACM SIGCOM 1997*, Cannes, France, Sept. 1997, p. 289299.

[41] Oguz, N. C. and Ayanoglu, E., "Performance analysis of two-level forward error correction for lost cell recovery in ATM networks," *in Proceedings of IEEE INFOCOM95*, Boston, MA, April 1995, p. 6b.3.16b.3.10.

[42] Ghanbari, M., "Two-layer coding of video signals for VBR networks," *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 5, June 1989, pp. 771–781.

[43] Ramchandran, K., Ortega, A., Uz, K., and Vetterli, M., "Multiresolution broadcast for digital HDTV using joint source/channel coding," *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 1, Jan. 1993, pp. 6–23.

[44] Zhang, Y.-Q., Liu, Y.-J., and Pickholtz, R., "Layered image transmission over cellular radio channels," *IEEE Transactions on Vehicular Technology*, Vol. 43, No. 3, Aug. 1994, pp. 786–794.

[45] Khansari, M. and Vetterli, M., "Layered transmission of signals over power-constrained wireless channels," *in Proc. of the IEEE International Conference on Image Processing (ICIP)*, Washington, DC, Oct. 1995, pp. 380–383.

[46] Aravind, R., Civanlar, M., and Reibman, A., "Packet loss resilience of MPEG-2 scalable video coding algorithms," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 5, Oct. 1996, pp. 426435.

[47] Girod, B., Farber, N., and Horn, U., "Scalable codec architectures for Internet video-on-demand," *in Proc. of the Thirty-First Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1997, pp. 357–361.

[48] Reibman, A. R., Bottou, L., and Basso, A., "Scalable video coding with managed drift," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 2, Feb. 2003, pp. 131– 140.

[49] Schwarz, H., Marpe, D., and wiegand, T., "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Transactions on circuits and Systems for Video Technology*, Vol. 17, No. 9, Sep 2007, pp. 1103–1120.

[50] Li, W., "Overview of fine granularity scalability in MPEG-4 video standard," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 3, March 2001, pp. 301–317.

[51] van der Schaar, M. and Radha, H., "Adaptive motion-compensation fine-granularity scalability (AMC-FGS) for wireless video," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 12, No. 6, June 2002, pp. 360–371.

[52] Wu, F., Li, S., and Zhang, Y. Q., "A framework for efficient progressive fine granularity scalable video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 3, March 2001, pp. 332344.

[53] Yang, X. and Ramchandran, K., "Optimal subband filter banks for multiple description coding," *IEEE Transactions on information theory*, Vol. 46, No. 7, Nov. 2000, pp. 2477–2490.

[54] Goyal, V. K., "Multiple description coding: compression meets the network," *IEEE Signal Processing Magazine*, Vol. 18, No. 5, Sept. 2001, pp. 74–93.

[55] Ozarow, L., "On a source coding problem with two channels and three receivers," *Bell System Technical Journal*, Vol. 59, No. 10, Dec. 1980, pp. 1909–1920.

[56] Wolf, J., Wyner, A., and Ziv, J., "Source coding for multiple descriptions," *Bell System Technical Journal*, Vol. 59, No. 8, Oct. 1980, pp. 1417–1426.

[57] Gamal, A. and Cover, T., "Achievable rates for multiple descriptions," *IEEE Transactions on Information Theory*, Vol. 28, No. 6, Nov. 1982, pp. 851–857.

[58] Ahlswede, R., "The rate-distortion region for multiple descriptions without excess rate," *IEEE Transactions on Information Theory*, Vol. 31, No. 6, Nov. 1985, pp. 721–726.

[59] Vaishampayan, V., "Design of multiple description scalar quantizer," *IEEE Transactions on Information Theory*, Vol. 39, No. 3, May 1993, pp. 821–834.

[60] Zhang, Z. and Berger, T., "Multiple description source coding with no excess marginal rate," *IEEE Transactions on Information Theory*, Vol. 41, No. 2, March 1995, pp. 349–357.

[61] Batllo, J.-C. and Vaishampayan, V., "Asymptotic performance of multiple description transform codes," *IEEE Transactions on Information Theory*, Vol. 43, No. 2, March 1997, pp. 703–707.

[62] Goyal, V., Vetterli, M., and Kovacevic, J., "Multiple description transform coding: robustness to erasures using tight frame expansions," *in Proc. IEEE International Symposium on Information Theory*, Cambridge, MA, Aug. 1998, p. 408.

[63] Chou, P., Mehrotra, S., and Wang, A., "Multiple description decoding of overcomplete expansions using projections onto convex sets," *in Proc. IEEE Conference on Data Compression (DCC)*, Snowbird, UT, March 1999, pp. 72–81.

[64] Goyal, V. and Kovacevic, J., "Generalized multiple description coding with correlating transforms," *IEEE Transactions on Information Theory*, Vol. 47, No. 6, Sept. 2001, pp. 2199–2224.

[65] Alasti, M., Sayrafian-Pour, K., Ephremides, A., and Farvardin, N., "Multiple description coding in networks with congestion problem," *IEEE Transactions on Information Theory*, Vol. 47, No. 3, March 2001, pp. 891902.

[66] Arean, R., Kovacevic, J., and Goyal, V., "Multiple description perceptual audio coding with correlating transforms," *IEEE Transactions on Speech and Audio Processing*, Vol. 8, No. 2, March 2000, pp. 140–145.

[67] Hemami, S., "Reconstruction-optimized lapped orthogonal transforms for robust image transmission," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 2, April 1996, pp. 168–181.

[68] Wang, Y., Orchard, M., Vaishampayan, V., and Reibman, A., "Multiple description coding using pairwise correlating transforms," *IEEE Transactions on Image Processing*, Vol. 10, No. 3, March 2001, pp. 351–366.

[69] Reibmannnn, A. R., Jafarkhani, H., Wang, Y., Orchard, M. T., and Puri, R., "Multiple-description video coding using motion-compensation temporal prediction," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 12, No. 3, March 2002, pp. 193–203.

[70] Apostolopoulos, J., "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," *Visual communications image processing*, Jan. 2001, pp. 392–409.

[71] Franchi, N., Fumagalli, M., Lancini, R., and Tubaro, S., "Multiple video coding for scalable and robust transmission over IP," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, No. 3, March 2005, pp. 321– 334.

[72] Cho, S. and Pearlman, W., "A full-feature, error resilient, scalable wavelet video codec based on the set partitioning in hierarchical trees (SPIHT) algorithm,"

IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No. 3, March 2002, pp. 157–171.

[73] Puri, R., Lee, K. W., and Ramchandran, K., "An integrated source transcoding and congestion control paradigm for video streaming in the Internet," *IEEE Transactions on Multimedia*, Vol. 3, No. 1, March 2001, pp. 18–32.

[74] Puri, R. and Ramchandran, K., "Multiple description source coding using forward error correction codes," *in Proc. 33rd Asilomar Conference on Signal, Systems and Computer*, Pacific Grove, CA, Oct. 1999, p. 342346.

[75] Cheng, L., Zhang, W., and Chen, L., "Rate-distortion optimized unequal loss protection for FGS compressed video," *IEEE transactions on broadcasting*, Vol. 50, No. 2, June 2004, pp. 126131.

[76] van der Schaar, M. and Radha, H., "Unequal packet loss resilience for fine-granular-scalability video," *IEEE transactions on Multimedia*, Vol. 3, No. 4, Dec. 2001, pp. 381394.

[77] Kim, J., Mersereau, R. M., and Altunbasak, Y., "Error-resilient image and video transmission over the Internet using unequal error protection," *IEEE Transcations on Image Processing*, Vol. 12, No. 2, Feb. 2003, pp. 121–131.

[78] Chan, Y. S., Cosman, P. C., and Milstein, L. B., "n-channel symmetric motion-compensated multiple description coding for video communications over OFDM networks," *Proceedings of Military Communications Conference*, Washington D. C., Oct. 2006, pp. 1–7.

[79] Gogate, N., Chung, D.-M., Panwar, S., , and Wang, Y., "Supporting image and video applications in a multihop radio environment using path diversity and multiple description coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 12, No. 9, Sept. 2002, pp. 777–792.

[80] Nguyen, T. and Zakhor, A., "Distributed video streaming over the Internet," *in Proceedings of SPIE Conference on Multimedia Computing and Networking*, San Jose, CA, Jan. 2002.

[81] Nguyen, T. and Zakhor, A., "Distributed video streaming with forward error correction," *in Proceedings of Packet Video Workshop*, Pittsburg, PA, April 2002.

[82] Liang, Y. J., Setton, E., and Girod, B., "Network-adaptive video communication using packet path diversity and rate-distortion optimized reference picture selection," *Journal of VLSI Signal Processing Systems*, Vol. 41, No. 3, Nov. 2005, pp. 345–354.

[83] Apostolopoulos, J. G. and Trott, M. D., "Path diversity for enhanced media streaming," *IEEE Communications Magazine*, Vol. 42, No. 8, Aug. 2004, pp. 80–87.

[84] Nguyen, T. and Zakhor, A., "Path diversity with forward error correction (PDF) system for packet switched networks," *IEEE Transactions on Circuit and System for Video technology*, Vol. 12, No. 6, June 2002, pp. 360–371.

[85] Yu, X., Modestino, J. W., and Bajic, I. V., "Modeling and analysis of multipath video transport over lossy networks using packet-level FEC," *Proceedings of DMS'05*, Banff, AB, Canada, Sept. 2005, pp. 265–270.

[86] Apostolopoulos, J., Tan, W. T., and Wee, S., "Modeling path diversity for multiple description video communication," *Proceedings of ICASSP'02*, Orlando, FL, May 2002, pp. III–2161 – III–2164.

[87] Apostolopoulos, J. and Wee, S., "Unbalanced multiple description video communication using path diversity," *Proceedings of IEEE ICIP'01*, Thessaloniki, Greece, Oct. 2001, pp. 966–969.

[88] Kim, J., Mersereau, R. M., and Altunbasak, Y., "Distributed video streaming using multiple description coding and unequal error protection," *IEEE Transactions on Image Processing*, Vol. 14, No. 7, July 2005, pp. 849–861.

[89] Fan, D., Chan, Y. S., Yu, X., and Modestino, J. W., "Network-adaptive transport of motion-compensated fine granularity scalability video using multiple asymmetric paths," *in Proceedings of ICIP 2008*, San Diego, CA, Oct. 2008, pp. 2320 – 2323.

[90] Savage, S., Collins, A., Hoffman, E., Snell, J., and Anderson, T., "The end-to-end effects of Internet path selection," *Computer Communication Review, ACM SIGCOM '99*, Vol. 29, No. 4, Oct 1999, pp. 289–299.

[91] Begen, A. C., Altunbasak, Y., Ergun, O., and Ammar, M. H., "Multi-path selection for multiple description video streaming over overlay Networks," *EURASIP Signal Processing: Image Communication*, Vol. 20, No. 1, Jan. 2005, pp. 39–60.

[92] Wah, B. W. and Lin, D., "Transformation-based reconstruction for real-time voice transmissions over the Internet," *IEEE Transactions on Multimedia*, Vol. 1, No. 4, Dec. 1999, pp. 342–351.

[93] Shacham, N. and Mckenney, P., "Packet recovery in high-speed networks using coding and buffer management," *Annual Joint Conference of the IEEE Computer and Communications Societies Proceedings (INFOCOM 1990)*, Vol. 1, San Francisco,CA, June 1990, pp. 124–131.

[94] Liang, Y. J., Apostolopoulos, J. G., and Girod, B., "Model-based delay-distortion optimization for video streaming using packet interleaving," *Proceedings 36th Asilomar Conference on Signals, Systems and Computers*, Vol. 2, Pacific Grove, CA, Nov. 2002, pp. 1315–1319.

[95] Gan, T., Gan, L., and Ma, K.-K., "Reducing video-quality fluctuations for streaming scalable video using unequal error protection, retransmission, and interleaving," *IEEE Transactions on Image Processing*, Vol. 15, No. 4, April 2006, pp. 819 – 832.

[96] Fan, D., Modestino, J. W., Chan, Y. S., and Yu, X., "Delay-constrained motion-compensated FGS video transport with optimal interleaving," *in Proceedings of Milcom 2007*, Orlando, FL, Oct. 2007, pp. 1–7.

[97] Zhao, Y., Ahalt, S. C., and Dong, J., "Optimal interleaving for 3-D zerotree-wavelet video packets over bursty lossy channels," *in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Philadelphia, PA, March 2005.

[98] Wu, F., Li, S., and Zhang, Y.-Q., "A framework for efficient progressive fine granularity scalablevideo coding," *IEEE Journal on Circuits and Systems for Video Technology*, Vol. 11, No. 3, March 2001, pp. 332–344.

[99] Huang, H.-C., Wang, C.-N., and Chiang, T., "A robust fine granularity scalability using trellis-based predictive leak," *IEEE Journal on Circuits and Systems for Video Technology*, Vol. 12, No. 6, June 2002, pp. 172–185.

[100] Cai, J. and Li, X., "Layered unequal loss protection with pre-interleaving for fast progressive image transmission over packet-loss channels," *ACM Trans. On Multimedia Computing, Communications and Applications*, Vol. 1, No. 4, Nov. 2005, pp. 338–353.

[101] Kanal, L. N. and Sastry, A. R. K., "Models for channels with memory and their applications to error control," *Proceedings of IEEE*, Vol. 66, No. 7, July 1978, pp. 724–744.

[102] Norris, J., *Markov chains*, Cambridge University Press, 1997.

[103] Kim, B.-J., Xiong, Z., and Pearlman, W. A., "Low bit-rate, scalable video coding with 3D set partitioning in hierarchical trees (3D SPIHT)," *IEEE Transactions on Circuits and System for Video Technology*, Vol. 10, No. 8, Dec. 2000, pp. 1374–1387.

[104] Kalman, M., Steinbach, E., and Girod, B., "Adaptive media playout for low-delay video streaming over error-prone channels," *IEEE Transcations on circuits and Systems for video Technology*, Vol. 14, No. 6, June 2004, pp. 841–851.

[105] Kang, S. R. and Loguinov, D., "Modeling best-effort and FEC streaming of scalable video in lossy network channels," *IEEE Transactions on Networking*, Vol. 15, No. 1, Feb. 2007, pp. 187–200.

[106] Elliott, E. O., "A model of switched telephone network for data communications," *Bell. Syst. Techn. J.*, Vol. 44, Jan. 1965, pp. 89–109.

[107] Said, A. and Pearlman, W., "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 3, Jun 1996, pp. 243–250.

[108] Taubman, D., "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, Vol. 9, No. 7, Jul 2000, pp. 1158–1170.

[109] Horn, U., Stuhmuller, K., Link, M., and Girod, B., "Robust internet video transmission based on scalable coding and unequal error protection," *Signal Processing: Image Communication*, Vol. 15, No. 1-2, Sept. 1999, pp. 77–94.

[110] Chande, V. and Farvardin, N., "Progressive transmission of images over memoryless noisy channels," *IEEE Jounal on Selected Areas in Communication*, Vol. 18, No. 6, Jun 2000, pp. 850–860.

[111] Chande, V. and Farvardin, N., "Joint source-channel coding for progressive transmission of embedded source coders," *Data Compression Conference, 1999. Proceedings. DCC '99*, Snowbird, UT, USA, March 1999, pp. 52–61.

[112] Elliott, E. O., "A model of the switched telephone network for data communications," *Bell Sys. Tech. Journal*, Vol. 44, January 1963, pp. 89–109.

[113] Cidon, I., Khamisy, A., and Sidi, M., "Analysis of packet loss processes in high-speed networks," *IEEE Transactions on Information Theory*, Vol. 39, No. 1, Jan. 1993, pp. 98–108.

[114] Altman, E. and Jean-Marie, A., "Loss probabilities for messages with redundant packets feeding a finite buffer," *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 5, Jun 1998, pp. 778–787.

[115] Dan, G., Fodor, V., and Karlsson, G., "On the effects of the packet size distribution on the packet loss processs," *Telecommunication Systems*, Vol. 32, No. 1, May 2006, pp. 31–53.

[116] Kleinrock, L., *Queueing Systems*, Vol. 1, Wiley, New York, 1976.

[117] Equitz, W. H. R. and Cover, T. M., "Successive refinement of information," *IEEE Transactions on Information Theory*, Vol. 37, No. 2, March 1991, pp. 269–275.