

On the effects of the packet size distribution on the packet loss process

György Dán, Viktória Fodor and Gunnar Karlsson
KTH, Royal Institute of Technology,
Department of Signals, Sensors and Systems
10044 Stockholm, Sweden
{gyuri,vfodor,gk}@s3.kth.se

Abstract

Real-time multimedia applications have to use forward error correction (FEC) and error concealment techniques to cope with losses in today's best-effort Internet. The efficiency of these solutions is known however to depend on the correlation between losses in the media stream. In this paper we investigate how the packet size distribution affects the packet loss process, that is, the distribution of the number of lost packets in a block, the related FEC performance and the average loss run length. We present mathematical models for the loss process of the MMPP+M/D/1/K and the MMPP+M/M/1/K queues; we validate the models via simulations, and compare the results to simulation results with an MPEG-4 coded video trace. We conclude that the deterministic packet size distribution (PSD) not only results in lower stationary loss probability than the exponential one, but also gives a less correlated loss process, both at a particular average link load and at a particular stationary loss probability as seen by the media stream. Our results show that the effects of the PSD on FEC performance are higher in access networks, where a single multimedia stream might affect the multiplexing behavior and thus can improve the queuing performance by decreasing the variance of its PSD.

1 Introduction

For stream-type multimedia communications the number of losses in a block of packets and the average number of consecutively lost packets have a great impact both on the user-perceived visual quality and on the possible ways of improving it, for example by forward error correction and error concealment.

Forward error correction (FEC) has been proposed to recover from information losses in real-time applications, where the latency introduced by retransmission schemes is not acceptable [Biersack, 1992, Dán and Fodor, 2003]. FEC increases the redundancy of the transmitted stream and recovers losses based on the redundant information. There are two main directions of FEC design to recover from packet losses. One solution [Dube and Altman, 2002], proposed by the IETF and implemented in Internet audio tools is to add a redundant copy of the original packet to one of the subsequent

⁰This work has been supported in part by E-NEXT and by the Swedish Foundation for Strategic Research under the program Affordable Wireless Services and Infrastructures.

packets. The other set of solutions, considered in this paper, use block coding schemes based on algebraic coding, e.g. Reed-Solomon coding [Kawahara et al., 1994]. The error correcting capability of RS codes with k data packets and c redundant packets is c if data is lost. Thus the performance of an FEC scheme is largely affected by the characteristics of the loss process, i.e. the probability of losing more than c packets in a block of $k + c$ packets.

In the case of multimedia, errors that could not be corrected have to be detected and concealed. The efficiency of the concealment depends however on the loss pattern, in particular the number of consecutively lost packets has been shown to have an influence. In [Mohamed et al., 2004] the authors concluded that in the case of real-time speech a longer average loss run length results in better perceived quality. In [Liang et al., 2003] it has been shown however that in the case of compressed video the distortion caused by losses is higher if losses occur in bursts than if they occur isolated.

The number of packets lost in a block of packets and the average loss burst length depend on three properties of the traffic: the burstiness of the tagged stream, the burstiness of the background traffic and the packet size distribution (PSD) in the network. While the effects of the burstiness of the tagged stream and the background traffic have been investigated before [Biersack, 1992, Kawahara et al., 1994, Dán and Fodor, 2003, Cidon et al., 1993, Frossard, 2001], the effects of the packet size distribution remain unclear. The literature on the efficiency of FEC can be divided into three groups based on their assumption on the PSD. The first group considers fixed length packets and uses simulations and discrete time queuing systems as they considered the transmission of multimedia over ATM networks [Biersack, 1992, Kawahara et al., 1994, Schulzrinne et al., 1992]. The second group considers exponential PSD due to the simplicity of the resulting analytic models and the assumption that the PSD in the Internet can be approximated with the exponential PSD [Altman and Jean-Marie, 1998, Cidon et al., 1993, Ait-Hellal et al., 1999, Gurewitz et al., 2000]. The third group of works uses channel models to model losses, e.g. the Gilbert model [Gilbert, 1960], and thus does not consider any particular PSD [Frossard, 2001]. Albeit there is extensive literature on modeling correlated losses, a comparison of the results with different PSDs has never been done, and thus it is not clear if and how the PSD affects the loss process, i.e. the number of packets lost in a block of packets and the average loss run length, and hence the related performance of FEC and loss concealment techniques.

In this paper we present two models to analyze the packet loss process of a bursty source, for example VBR video, multiplexed with background traffic in a single multiplexer with a finite queue. We consider exponentially distributed and constant packet sizes. We model the bursty source by an L-state Markov-modulated Poisson process (MMPP) while the background traffic is modeled by a Poisson process. We validate the models via simulations and show results obtained with an MPEG-4 coded video trace. We compare the packet loss process, the related FEC performance and the average loss run length in case of the two packet size distributions.

It is well known that compressed media, primarily VBR video, exhibits a self-similar nature [Beran et al., 1995]. Ryu and Elwalid [Ryu and Elwalid, 1996] showed that short term correlations have dominant influence on the network performance under realistic scenarios of buffer sizes for real-time traffic. Grossglauser and Bolot [Grossglauser and Bolot, 1999] showed that even Markovian models are capable to capture the correlation structure influencing the queuing performance of long range dependent traffic for finite queue sizes, as the impact on loss of the correlation in the arrival process becomes nil beyond a time scale. In [Robert and Le Boudec, 1997]

Robert and Le Boudec used an MMPP to fit the mean and Hurst parameter of pseudo self-similar traffic. In [Andersen and Nielsen, 1998] Andersen and Nielsen used the superposition of two-state MMPPs to model second order self-similar processes, and showed that the model gives an accurate estimate of the queue length distribution. Yoshihara et al. used the superposition of 2-state IPPs to model self-similar traffic in [Yoshihara et al., 2001] and compared the loss probability of the resulting MMPP/D/1/K queue with simulations. They found that the approximation works well under heavy load conditions and it gives an upper bound on the packet loss probabilities. Salvador et al. [Salvador et al., 2003] proposed a parameter fitting method for the superposition of MMPPs to model self-similar traffic, and achieved accurate estimates of the queuing behavior. Klemm et al. [Klemm et al., 1993] used the Batch Markovian Arrival Process to model IP traffic exhibiting long-range dependence, and showed that the model is effective both in terms of statistical properties and queuing performance. Thus the MMPP may be a practical model to derive approximate results for the queuing behavior of LRD traffic such as real-time VBR video, especially in the case of small buffer sizes. Furthermore it is a popular source model for packetized voice [Heffes and Lucantoni, 1986], and hence the results shown in this paper can be applied to voice over IP performance evaluation as well. Recently Cao et al. [Cao et al., 2002] showed that the traffic generated by a large number of sources tends to Poisson as the load increases due to statistical multiplexing and hence justifies the Poisson model for the background traffic.

The paper is organized as follows. Section 2 gives an overview of the previous work on the modeling of the loss process of a single server queue. In Section 3 we describe our models to calculate the loss probability in a block of packets. In Section 4 we validate our models by simulations and compare the results obtained with the two PSDs. In Section 5 we conclude our work.

2 Related Work

In [Cidon et al., 1993], Cidon et al. present an exact analysis of the packet loss process in an M/M/1/K queue, that is the probability of losing j packets in a block of n packets, and show that the distribution of losses may be bursty compared to the assumption of independence. They also consider a discrete time system fed with a Bernoulli arrival process describing the behavior of an ATM multiplexer. In [Gurewitz et al., 2000], Gurewitz et al. present explicit expressions for the above quantities of interest for the M/M/1/K queue. In [Altman and Jean-Marie, 1998] the multidimensional generating function of the probability of j losses in a block of n packets is obtained and an easy-to-calculate asymptotic result is given under the condition that $n \leq K + j + 1$. The model was extended in [Dube et al., 2003] to include random losses.

The above models consider exponentially distributed service times. There has been far less work on the analysis of the loss process however for service time distributions other than exponential. Models with general and deterministic service times have been proposed for calculating various measures of queuing performance, such as the stationary loss probability, the distribution of the busy period and the virtual waiting time. In [Heffes and Lucantoni, 1986] the performance of the MMPP/G/1/K queue was evaluated considering the superposition of multimedia and data traffic at a single server queue, and the corresponding delay distribution was given. The waiting time and queue length distribution of the N/G/1/K queue (N stands for the Neuts process [Neuts, 1979]) was derived in [Blondia, 1989] including the MMPP/G/1/K queue as a

special case.

In [Ait-Hellal et al., 1999], Ait-Hellal et al. present an asymptotic result for a system where the service times and the interarrival times are stationary ergodic, in particular they show that if the block lengths n are large enough, then the uncorrected loss probabilities can be made arbitrarily small for any ratio of redundancy c/n , given that the average packet loss probability is lower than the ratio of redundancy. This result does not apply however to practical systems where the available delay and thus the maximum block length is limited. In [Abramov, 2004] Abramov gives an analysis of the asymptotic behavior of the M/GI/1 finite buffer queue, and studies the message loss probabilities if redundant packets are added to messages composed of a variable number of packets.

The conditional loss probability (CLP), i.e. the probability that a packet is lost given that the preceding packet was lost, is derived for the N*IPP/D/1/K queue in [Schulzrinne et al., 1992] and it is shown that the CLP can be orders of magnitude higher than the loss probability. Kawahara et al. consider a discrete time system fed with an interrupted Bernoulli process and bursty background traffic modeled by a Markov modulated Bernoulli process in [Kawahara et al., 1994]. They use the model to investigate the effects of the burstiness of the tagged source and the buffer size on the effectiveness of FEC in ATM networks.

Despite of the extensive work on modeling the packet loss correlations and the probability of losses in a block, an analysis of the effects of the packet size distribution on the loss process has not been done.

3 Model Description

We consider a single multiplexer fed by two sources, a Markov-modulated Poisson process (MMPP) and a Poisson process, representing the tagged source and the background traffic respectively. We assume that the sources feeding the system are independent. The MMPP is described by the infinitesimal generator Q with elements r_{ij} and the arrival rate matrix $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_L\}$, where λ_i is the average arrival rate while the underlying Markov chain is in state i . The Poisson process modeling the background traffic has average arrival rate λ . The superposition of the two sources can be described by a single MMPP with arrival rate matrix $\hat{\Lambda} = \Lambda \oplus \lambda = \Lambda + \lambda I = \text{diag}\{\hat{\lambda}_1, \dots, \hat{\lambda}_L\}$, and infinitesimal generator $\hat{Q} = Q$, where \oplus is the Kronecker sum.

The packets are stored in a buffer that can host up to K packets, and are served according to a FIFO policy. Every n consecutive packets from the tagged source form a block, and we are interested in the probability distribution of the number of lost packets in a block arriving from the MMPP in the steady state of the system. Thus our purpose is to calculate the probability $P(j, n), n \geq 1, 0 \leq j \leq n$ of j losses in a block of n packets. Throughout the calculations we use notations similar to those in [Cidon et al., 1993].

3.1 Exponential Packet Size Distribution

First we consider a system with exponential service time distribution and average service time $1/\mu$. We define the probability $P_{i,l}^a(j, n), 0 \leq i \leq K, l = 1 \dots L, n \geq 1, 0 \leq j \leq n$ as the probability of j losses in a block of n packets, given that the number of packets in the system is i just before the arrival epoch of the first packet in the block and the first packet of the block is generated in state l of the MMPP. Similarly, we define the

probability $P_{i,l}^s(j,n)$, $0 \leq i \leq K, l = 1 \dots L, 0 \leq j \leq n$ as the probability of j losses in a block of n packets, given that the number of packets in the system is i just before the arrival of a packet from the background traffic and the MMPP is in state l . As the first packet in the block is arbitrary,

$$P(j,n) = \sum_{l=1}^L \sum_{i=0}^K \Pi(i,l) P_{i,l}^a(j,n) \quad (1)$$

$\Pi(i,l)$, the steady state distribution of number of packets in the queue as seen by an arriving packet can be derived from the steady state distribution of the $MMPP/M/1/K$ queue as

$$\Pi(i,l) = \frac{\pi(i,l)\lambda_l}{\sum_{l=1}^L \lambda_l \sum_{i=0}^K \pi(i,l)}, \quad (2)$$

where $\pi(i,l)$, $0 \leq i \leq K, l = 1 \dots L$ denotes the probability that there are i packets in the queue and the MMPP is in state l in the steady state of the $MMPP/M/1/K$ queue [Gross and Harris, 1998].

In the following we give a recursion to calculate the probabilities $P_{i,l}^a(j,n)$. To this end we first define $Q_{i+1}^{l,m}(k)$, the joint conditional probability of the events that out of i packets k leave during an interarrival time and the next arrival occurs in state m of the underlying Markov chain, given that the last arrival occurred in state l . A way to calculate $Q_{i+1}^{l,m}(k)$ is given in Appendix B. Furthermore we introduce the notation $p_m = \frac{\lambda_m}{\lambda_m + \lambda}$ and $\bar{p}_m = \frac{\lambda}{\lambda_m + \lambda}$ for the probability that a packet arrives from the tagged source and from the background traffic respectively when the MMPP is in state m . The recursion is then initiated for $n = 1$ with the following relations

$$\begin{aligned} P_{i,l}^a(j,1) &= \begin{cases} 1 & j=0 \\ 0 & j \geq 1 \end{cases} & i \leq K-1 \\ P_{i,l}^a(j,1) &= \begin{cases} 0 & j=0, j \geq 2 \\ 1 & j=1 \end{cases} & K-1 < i. \end{aligned} \quad (3)$$

For $n \geq 2$ the following equations hold.

$$P_{i,l}^a(j,n) = \sum_{m=1}^L \sum_{k=0}^{i+1} Q_{i+1}^{l,m}(k) \{p_m P_{i+1-k,m}^a(j,n-1) + \bar{p}_m P_{i+1-k,m}^s(j,n-1)\} \quad (4)$$

for $0 \leq i \leq K-1$ and

$$P_{i,l}^a(j,n) = \sum_{m=1}^L \sum_{k=0}^K Q_K^{l,m}(k) \{p_m P_{K-k,m}^a(j-1,n-1) + \bar{p}_m P_{K-k,m}^s(j-1,n-1)\} \quad (5)$$

for $i = K$. $P_{i,l}^s(j,n)$ is given by

$$P_{i,l}^s(j,n) = \sum_{m=1}^L \sum_{k=0}^{i+1} Q_{i+1}^{l,m}(k) \{p_m P_{i+1-k,m}^a(j,n) + \bar{p}_m P_{i+1-k,m}^s(j,n)\} \quad (6)$$

for $0 \leq i \leq K-1$ and

$$P_{i,l}^s(j,n) = \sum_{m=1}^L \sum_{k=0}^K Q_K^{l,m}(k) \{p_m P_{K-k,m}^a(j,n) + \bar{p}_m P_{K-k,m}^s(j,n)\} \quad (7)$$

for $i = K$.

The procedure of computing $P_{i,l}^a(j, n)$ is as follows. First we calculate $P_{i,l}^a(j, 1)$, $i = 0 \dots K$ from the initial conditions (3). Then in iteration k we first calculate $P_{i,l}^s(j, k)$, $k = 1 \dots n - 1$ using equations (6) and (7) and the probabilities $P_{i,l}^a(j, k)$, which have been calculated during iteration $k - 1$. Then we calculate $P_{i,l}^a(j, k + 1)$ using equations (4) and (5).

3.2 Deterministic Packet Size Distribution

Next, we consider a system with constant packet sizes of transmission time D . We define the probability $P_{x,l}^a(j, n)$, $0 \leq x \leq KD$, $l = 1 \dots L$, $n \geq 1$, $0 \leq j \leq n$ as 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 in the block and the first packet of the block is generated in state l of the MMPP. Similarly, we define the probability $P_{x,l}^s(j, n)$, $0 \leq x \leq KD$, $l = 1 \dots L$, $n \geq 1$, $0 \leq j \leq n$ as the probability of j losses in a block of n packets from the tagged source, given that the remaining workload in the system is x just before the arrival of a packet from the background traffic and the MMPP is in state l . As the first packet in the block is arbitrary,

$$P(j, n) = \sum_{l=1}^L \int_{x=0}^{KD} V(x, l) P_{x,l}^a(j, n) dx. \quad (8)$$

An approximation for $V(x, l)$, the workload distribution of the steady state queue as seen by an arriving packet, can be given based on the steady state distribution of the $MMPP/E_r/1/K$ queue as outlined in Appendix A.

In the following we give a recursion to calculate the probabilities $P_{x,l}^a(j, n)$. Let us first denote by $f_{l,m}(t)$ the joint probability density function of the interarrival-time distribution of the joint arrival process and that the next arrival will occur in state m of the MMPP given that the last arrival was in state l of the MMPP. A way to calculate $f_{l,m}(t)$ is given in Appendix B.

The recursion is initiated for $n = 1$ with the following relations

$$\begin{aligned} P_{x,l}^a(j, 1) &= \begin{cases} 1 & j = 0 \\ 0 & j \geq 1 \end{cases} & x \leq (K-1)D, \\ P_{x,l}^a(j, 1) &= \begin{cases} 0 & j = 0, j \geq 2 \\ 1 & j = 1 \end{cases} & (K-1)D < x. \end{aligned} \quad (9)$$

Using the notation $p_m = \frac{\lambda_m}{\lambda_m + \lambda}$ and $\bar{p}_m = \frac{\lambda}{\lambda_m + \lambda}$ as in Subsection 3.1, for $n \geq 2$ the following equations hold.

$$\begin{aligned} P_{x,l}^a(j, n) &= \sum_{m=1}^L \int_0^{x+D} f_{l,m}(t) \{ p_m P_{x+D-t,m}^a(j, n-1) \\ &\quad + \bar{p}_m P_{x+D-t,m}^s(j, n-1) \} dt \\ &\quad + \int_{x+D}^{\infty} f_{l,m}(t) \{ p_m P_{0,m}^a(j, n-1) + \bar{p}_m P_{0,m}^s(j, n-1) \} dt \end{aligned} \quad (10)$$

for $0 \leq x \leq (K-1)D$ and for $(K-1)D < x$

$$P_{x,l}^a(j, n) = \sum_{m=1}^L \int_0^x f_{l,m}(t) \{ p_m P_{x-t,m}^a(j-1, n-1) \} dt \quad (11)$$

$$\begin{aligned}
& + \bar{p}_m P_{x-t,m}^s(j-1, n-1) \} dt \\
& + \int_x^\infty f_{l,m}(t) \{ p_m P_{0,m}^a(j-1, n-1) + \bar{p}_m P_{0,m}^s(j-1, n-1) \} dt.
\end{aligned}$$

$P_{x,l}^s(j, n)$ is given by

$$\begin{aligned}
P_{x,l}^s(j, n) &= \sum_{m=1}^L \int_0^{x+D} f_{l,m}(t) \{ p_m P_{x+D-t,m}^a(j, n) \\
& \quad + \bar{p}_m P_{x+D-t,m}^s(j, n) \} dt \\
& \quad + \int_{x+D}^\infty f_{l,m}(t) \{ p_m P_{0,m}^a(j, n) + \bar{p}_m P_{0,m}^s(j, n) \} dt
\end{aligned} \tag{12}$$

for $0 \leq x \leq (K-1)D$ and for $(K-1)D < x$

$$\begin{aligned}
P_{x,l}^s(j, n) &= \sum_{m=1}^L \int_0^x f_{l,m}(t) \{ p_m P_{x-t,m}^a(j, n) \\
& \quad + \bar{p}_m P_{x-t,m}^s(j, n) \} dt \\
& \quad + \int_x^\infty f_{l,m}(t) \{ p_m P_{0,m}^a(j, n) + \bar{p}_m P_{0,m}^s(j, n) \} dt.
\end{aligned} \tag{13}$$

3.3 Numerical Evaluation

The above set of integral equations can be solved using numerical integration. The finite integrals in equations (10) to (13) are calculated numerically while the infinite integrals - as the integrand only depends on t in $f_{l,m}(t)$ - can be evaluated analytically as shown in Appendix B (26). We introduce Δ the step size for the numerical integration such that $D = N\Delta$, and so instead of equations (9-13) we can write

$$\begin{aligned}
P_{i,l}^a(j, 1) &= \begin{cases} 1 & j=0 \\ 0 & j \geq 1 \end{cases} & i \leq (K-1)N, \\
P_{i,l}^a(j, 1) &= \begin{cases} 0 & j=0, j \geq 2 \\ 1 & j=1 \end{cases} & (K-1)N < i.
\end{aligned} \tag{14}$$

For $n \geq 2$ the following recursive equations hold.

$$\begin{aligned}
P_{i,l}^a(j, n) &= \sum_{m=1}^L \sum_{\tau=0}^{i+N} f_{l,m}(\tau\Delta) c_\tau^{i+N} \{ p_m P_{i+N-\tau,m}^a(j, n-1) \\
& \quad + \bar{p}_m P_{i+N-\tau,m}^s(j, n-1) \} \\
& \quad + \int_{i\Delta+D}^\infty f_{l,m}(t) \{ p_m P_{0,m}^a(j, n-1) + \bar{p}_m P_{0,m}^s(j, n-1) \} dt
\end{aligned} \tag{15}$$

for $0 \leq i \leq (K-1)N$ and for $(K-1)N < i$

$$\begin{aligned}
P_{i,l}^a(j, n) &= \sum_{m=1}^L \sum_{\tau=0}^i f_{l,m}(\tau\Delta) c_\tau^i \{ p_m P_{i-\tau,m}^a(j-1, n-1) \\
& \quad + \bar{p}_m P_{i-\tau,m}^s(j-1, n-1) \} \\
& \quad + \int_{i\Delta}^\infty f_{l,m}(t) \{ p_m P_{0,m}^a(j-1, n-1) + \bar{p}_m P_{0,m}^s(j-1, n-1) \} dt.
\end{aligned} \tag{16}$$

$P_{i,l}^s(j, n)$ is given by

$$\begin{aligned} P_{i,l}^s(j, n) = & \sum_{m=1}^L \sum_{\tau=0}^{i+N} f_{l,m}(\tau\Delta) c_{\tau}^{i+N} \{p_m P_{i+N-\tau,m}^a(j, n) \\ & + \bar{p}_m P_{i+N-\tau,m}^s(j, n)\} \\ & + \int_{i\Delta+D}^{\infty} f_{l,m}(t) \{p_m P_{0,m}^a(j, n) + \bar{p}_m P_{0,m}^s(j, n)\} dt \end{aligned} \quad (17)$$

for $0 \leq i \leq (K-1)N$ and for $(K-1)N < i$

$$\begin{aligned} P_{i,l}^s(j, n) = & \sum_{m=1}^L \sum_{\tau=0}^i f_{l,m}(\tau\Delta) c_{\tau}^i \{p_m P_{i-\tau,m}^a(j, n) \\ & + \bar{p}_m P_{i-\tau,m}^s(j, n)\} \\ & + \int_{i\Delta}^{\infty} f_{l,m}(t) \{p_m P_{0,m}^a(j, n) + \bar{p}_m P_{0,m}^s(j, n)\} dt, \end{aligned} \quad (18)$$

where the coefficient c_{τ}^i is the τ^{th} weighting coefficient in the i point numerical integration using the Newton-Cotes formulas. The Newton-Cotes rules in (15)-(18) can be built up of rules with $\lfloor \frac{N}{2} \rfloor$ or more points. By choosing $N \geq 10$ the remainder term can be written as $-\kappa(N) \left(\frac{1}{N}\right)^7$, where $\kappa(N)$ is a decreasing function of N and depends on D and the sixth or higher order derivatives of the integrand [Abramowitz and Stegun, 1972].

The procedure of computing $P_{i,l}^a(j, n)$ is as follows. First we calculate $P_{i,l}^a(j, 1)$, $i = 0 \dots KN$ from the initial conditions (14). Then in iteration k we first calculate $P_{i,l}^s(j, k)$, $k = 1 \dots n-1$ using equations (17) and (18) and the probabilities $P_{i,l}^a(j, k)$, which have been calculated during iteration $k-1$. Then we calculate $P_{i,l}^a(j, k+1)$ using equations (15) and (16).

4 Performance Analysis

In this section we validate the two models described in Section 3 via simulations and use them to evaluate the effects of the PSD on the distribution of the number of lost packets in a block of packets, the related FEC performance and the average loss run length. The average packet length of both the tagged and the background traffic is set to 188 bytes, as given for the transport stream in the MPEG-2 standard [Hoffman and Fernando, 1998]. Note that increasing the average packet length is equivalent to decreasing the link speed, and thus the particular fixed value of the average packet length does not limit the generality of the results presented here. The considered link speeds are 10 Mbps, 22.5 Mbps and 45 Mbps. The queue lengths are set to 5, 10 and 20 packets depending on the link speed, which results in a maximum average delay of around 0.5 ms in all cases. We performed simulations with an MPEG-4 coded talk-show video trace. The PSD of the video trace was in all cases deterministic, while for the PSD of the background traffic we used the exponential or the deterministic distribution depending on the considered scenario. In the analytical models and in the simulations validating the models we considered a 3 state MMPP, with an average bitrate of 540 kbps, arrival intensities $\lambda_1 = 83/s$, $\lambda_2 = 367/s$, $\lambda_3 = 661/s$ and transition rates $r_{12} = 40$, $r_{13} = 1$, $r_{21} = 10$, $r_{23} = 1.3$, $r_{31} = 6$, $r_{32} = 0.1$. These values were derived from the MPEG-4 encoded video trace to match the queuing performance on a 10 Mbps link with exponential PSD. In the case of the analytical model and the corresponding simulations the PSD of both the tagged source and the background traffic

were set to exponential or deterministic depending on the considered scenario. The simulations were performed in ns-2, the results have a 5 percent margin of error at a 95 percent level of confidence.

We use three measures to compare the packet loss process. The first one is a commonly used measure of closeness, the Kullback-Leibler distance [Kullback, 1959] defined for two distributions as

$$d(\mathcal{P}_1, \mathcal{P}_2) = \sum_{j=0}^n P_1(j, n) \log_2 \frac{P_1(j, n)}{P_2(j, n)}, \quad (19)$$

The Kullback-Leibler distance is the same as the relative entropy of \mathcal{P}_1 with respect to \mathcal{P}_2 . It is not a true metric, as it is not symmetric and does not satisfy the triangle inequality, but it is always non-negative and equals zero only if $\mathcal{P}_1 = \mathcal{P}_2$.

The second measure is based on the gain that can be achieved by using FEC. We denote by FEC(k,c+k) an FEC scheme using Reed-Solomon coding with k data packets and c redundant packets. Given the probabilities $P(j, n)$ the uncorrected loss probability for such a scheme can be calculated as

$$P_{loss}^{k,c+k} = \frac{1}{c+k} \sum_{j=c+1}^{c+k} jP(j, c+k). \quad (20)$$

Based on the uncorrected packet loss probability we define the FEC gain as the ratio of the average loss probability without the use of FEC and the uncorrected loss probability when using FEC: $f(k, c+k) = P_{loss} / P_{loss}^{k,c+k}$.

The third measure is the average loss run length, also called average loss burst size, i.e. the average number of consecutively lost packets. Given the probabilities $P(j, n)$ the average loss run length can be calculated as

$$E[B] = \sum_{n=1}^{\infty} P(B \geq n) / P(1, 1) = \sum_{n=1}^{\infty} P(n, n) / P(1, 1). \quad (21)$$

A higher value of the average loss run length results in fewer loss bursts given a particular stationary loss probability.

4.1 Constant average load

In this section we consider the case when the average load in the network is constant and compare the packet loss process, the efficiency of FEC and the average loss run length for the two PSDs. Figures 1-3 show the uncorrected loss probability without using FEC (denoted by FEC(1,1)) and using two different FEC schemes for the three considered link speeds as a function of the average load. The figures show simulation results using the MMPP, denoted by subscript s, and analytical results using the models presented in the paper, denoted by subscript m. The capital letter D stands for the deterministic, the capital letter M for the exponential PSD. The figures show that the models fit perfectly with the simulations. We can also observe that at a given average load the difference between the results with exponential and deterministic packet sizes is several orders of magnitude, not only in terms of stationary loss probability, but also in terms of the efficiency of FEC to reduce the uncorrected loss probability. As the link speed increases, the efficiency of FEC increases at a given average load level.

Figures 4-6 show similar scenarios to those in Figures 1-3, but show simulation results obtained with the video trace, denoted by subscript t, and analytical results

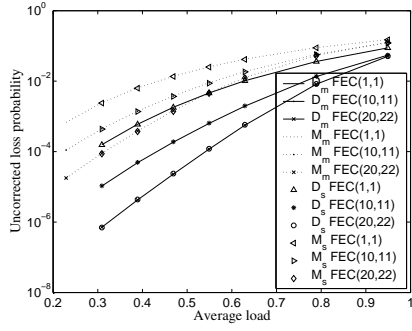


Figure 1: Uncorrected loss probability vs. average load on a 10 Mbps link using various FEC schemes for the MMPP model.

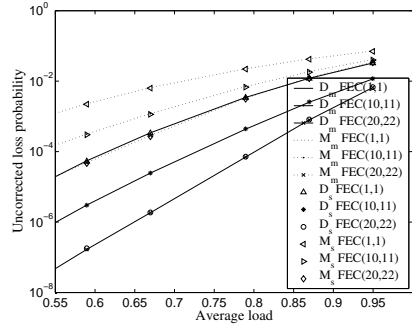


Figure 2: Uncorrected loss probability vs. average load on a 22.5 Mbps link using various FEC schemes for the MMPP model.

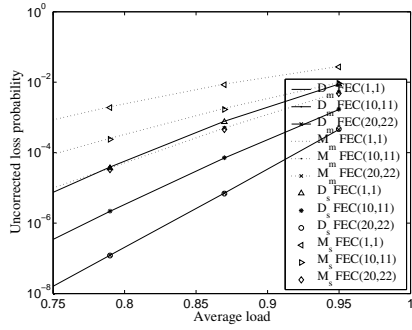


Figure 3: Uncorrected loss probability vs. average load on a 45 Mbps link using various FEC schemes for the MMPP model.

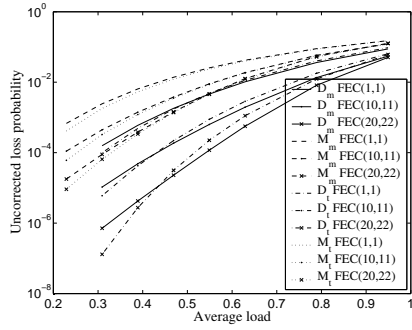


Figure 4: Uncorrected loss probability vs. average load on a 10 Mbps link using various FEC schemes (MMPP model vs. trace).

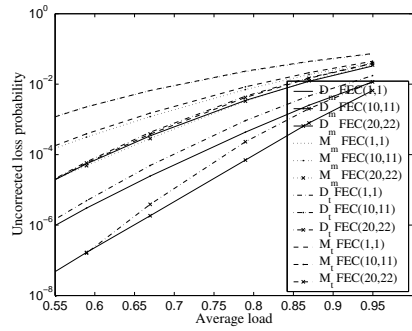


Figure 5: Uncorrected loss probability vs. average load on a 22.5 Mbps link using various FEC schemes (MMPP model vs. trace).

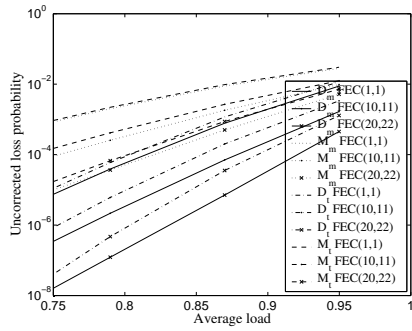


Figure 6: Uncorrected loss probability vs. average load on a 45 Mbps link using various FEC schemes (MMPP model vs. trace).

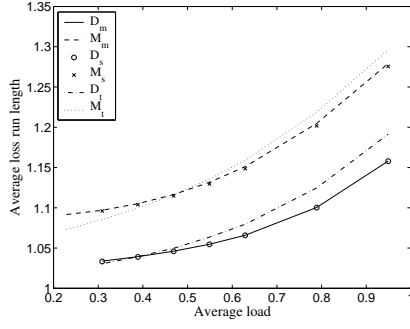


Figure 7: Average loss run length vs. average load on a 10 Mbps link.

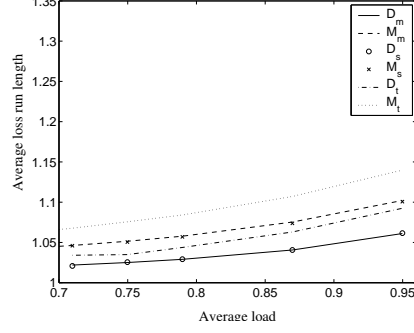


Figure 8: Average loss run length vs. average load on a 45 Mbps link.

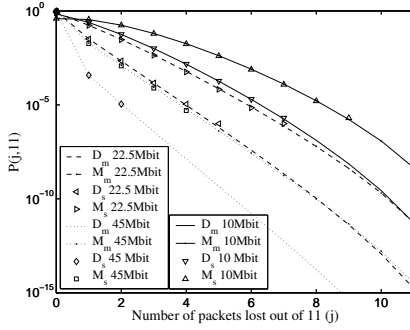


Figure 9: Probability of losing j packets in a block of 11 packets at average load level $\rho = 0.8$.

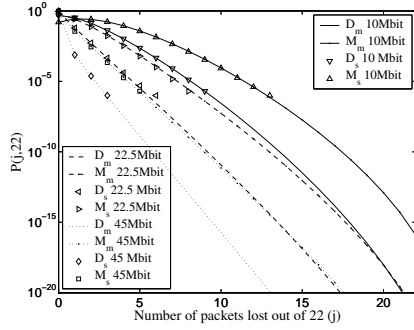


Figure 10: Probability of losing j packets in a block of 22 packets at average load level $\rho = 0.8$.

using the models presented in the paper, denoted by subscript m. The figures show that the MMPP approximation gives a rather accurate estimate of the video trace in terms of queuing performance despite of the low number of states of the model. The difference between the results with the model and the traces increases as the link speed increases.

Figures 7-8 show the average loss run length for two considered link speeds as a function of the average load. Comparing the results given by the analytical models and the simulations shows that the models presented give accurate results. We see that the average loss run length is higher for the exponential PSD, which is partly due to the higher stationary loss probability when using the exponential PSD compared to when using the deterministic PSD. Comparing the figures we see that the difference between the results with the different PSDs decreases as the link speed increases. Comparing results with the video trace and the MMPP we see that the model is accurate at a link speed of 10 Mbps, and underestimates the loss correlations for other link speeds. The relationship between results with the different PSDs is similar however, the average loss run length is lower in case of the deterministic PSD and the difference between the results decreases as the link speed increases.

Figures 9 and 10 show the probability of losing j packets in a block of 11 and 22

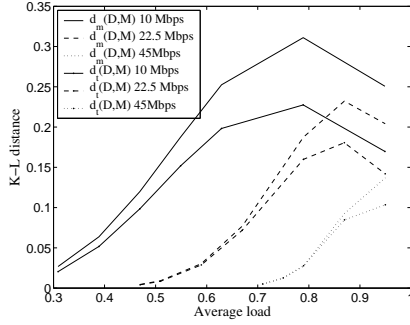


Figure 11: Kullback-Leibler distance vs. average load for $P(j,11)$.

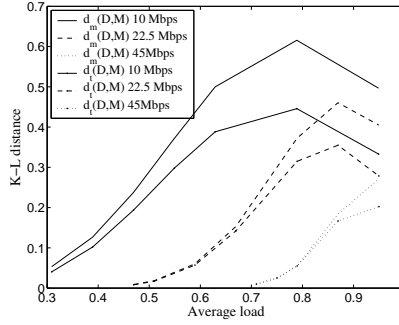


Figure 12: Kullback-Leibler distance vs. average load for $P(j,22)$.

packets respectively at an average load of $\rho = 0.8$ for the three considered link speeds. Comparing the results for the different PSDs we see that the probability of a large number of losses in a block is lower in the case of the deterministic PSD, which will result in better performance for FEC. The curves for the the different link speeds have different shapes, which is due to the different stationary loss probabilities.

Figures 11 and 12 show the Kullback-Leibler distance of $P(j,11)$ and $P(j,22)$ respectively. The figures show that the difference between the deterministic and exponential cases decreases as the link speed increases. At the same time results obtained for the different link speeds show the same properties and thus where possible we will only show results for the 10 Mbps link. The results obtained via simulations with the video trace show similar behavior to those obtained with the MMPP model. Comparing the results for the two different block lengths shows that the difference between the distributions increases as the block length increases. This observation is in accordance with Figures 9 and 10.

Figures 13 and 14 show the FEC gain for FEC(10,11) and FEC(20,22) on a 10 Mbps link. Comparing the FEC gain for the exponential and the deterministic PSD we see that it is higher in the case of the deterministic PSD, and the difference between the results of the distributions increases as the average load decreases. This result is supported by Figure 7 showing the average loss run length, the lower average loss run length of the deterministic PSD at a given average load contributes to the higher efficiency of FEC. The results showing the probability of losses in a block of packets are in accordance as well, the probability of a large number of lost packets is lower in the case of the deterministic PSD. Comparing the results for the FEC gain to those for the Kullback-Leibler distance we have to note that while the FEC gain measures the difference of the distributions shown in 9 and 10 for high values of j , the Kullback-Leibler distance puts much weight on the values of j where $P(j,n)$ is high, i.e. typically for low values of j . This explains why the difference between the distributions is high for low loads in terms of FEC gain, while it is high for high loads in terms of Kullback-Leibler distance. As a consequence these two measures measure different aspects of the difference between the results with the different PSDs. Comparing Figures 13 and 14 we conclude that FEC(10,11) and FEC(20,22) behave similarly, and thus in the following we will only show figures for FEC(20,22) for brevity.

The figures presented show that results for the video trace are similar to those for the MMPP model in terms of FEC gain, which confirms the results presented in the

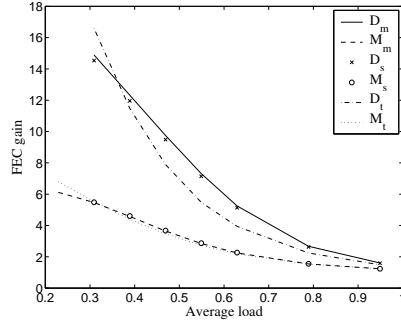


Figure 13: FEC gain vs. average load for FEC(10,11) on a 10 Mbps link.

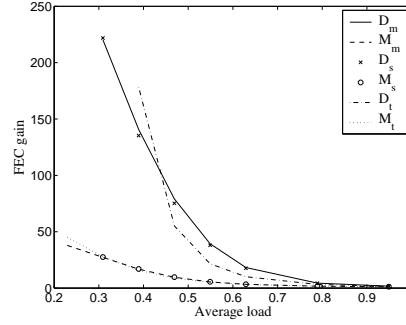


Figure 14: FEC gain vs. average load for FEC(20,22) on a 10 Mbps link.

literature on the efficiency of Markovian modeling of LRD traffic. Due to the similarity of the results and to maintain clarity of the figures, in the following subsections we will only show results obtained with the MMPP model. In all of the figures the simulations showed a perfect match with the results of the analytical model, and thus in the following we will not show simulation results.

The figures show that there is a significant difference between the results with the two packet size distributions: the uncorrected loss probability and the average burst size are lower while the gain achievable by using FEC is higher in the case of the deterministic PSD. The difference however is partly due to the different stationary loss probabilities at a given average load. We eliminate the effects of the stationary loss probability in the following subsection.

4.2 Equal average packet loss probabilities

In this section we show the results as a function of the stationary loss probability. The rationale behind this is that applications can measure the stationary loss probability in the network but not the average load. Thus from an application's point of view it is interesting to know whether the PSD in the network influences the gain achievable by using FEC. In order to compare the packet loss process at a certain stationary loss probability, we take the results with the deterministic PSD at a given average load and decrease the background traffic of the mathematical model with exponential PSD to match the average packet loss probability.

Figure 15 shows the probability of loosing j packets in a block of 22 packets. The average load of the scenarios with deterministic PSD is $\rho = 0.8$, while the average load of the scenarios with exponential PSD is set to match the stationary loss probability of the corresponding scenarios with deterministic PSD and the same link speed by changing the background traffic intensity. Comparing the figure to Fig. 10 we see that difference between the distributions with the different PSDs decreased significantly, but the probability of a large number of lost packets in a block is still lower in the case of the deterministic PSD despite of the equal average packet loss probabilities. Figure 16 shows the Kullback-Leibler distance between the results obtained with the two distributions as a function of the stationary loss probability. The distance between the results with the two PSDs decreased significantly (three orders of magnitude) compared to Fig. 12. Figure 17 shows the average loss run length on a 10 Mbps link. Figure

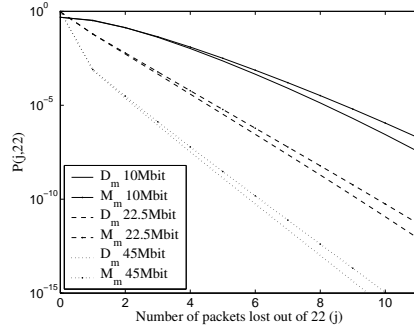


Figure 15: Probability of losing j packets in a block of 22 packets. The stationary loss probability of the scenarios with the same link speed is equal.

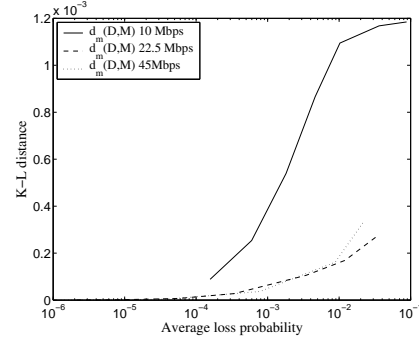


Figure 16: Kullback-Leibler distance vs. stationary loss probability for $P(j,22)$.

18 shows the FEC gain on a 10 Mbps link. The results for the loss run length support the results for the FEC gain, at a given stationary loss probability the deterministic PSD results in shorter loss bursts, which makes FEC more efficient. Comparing these figures to Figures 7 and 14 respectively also shows that the difference between the results obtained with the different distributions is lower at a particular average loss probability than at a particular average load level. Thus the observed difference in Subsection 4.1 was partly due to the different stationary loss probabilities at a particular average load level.

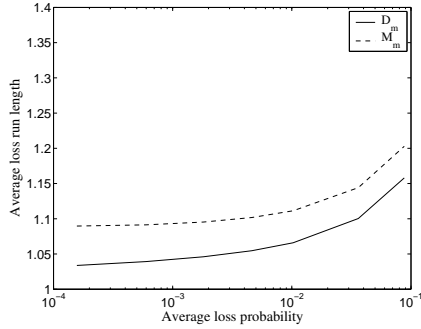


Figure 17: Average loss run length vs. stationary loss probability on a 10 Mbps link.

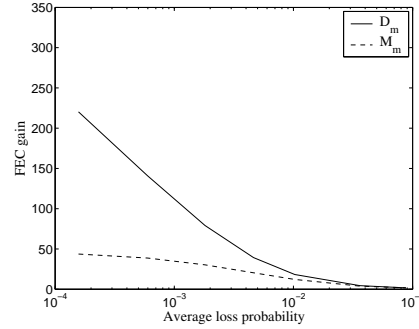


Figure 18: FEC gain vs. stationary loss probability on a 10 Mbps link for FEC(20,22).

The remaining difference between the packet loss processes and the related FEC performance (a factor of three in the considered scenario) can be due to the difference in the level of statistical multiplexing (the background traffic intensity was decreased in the case of the exponential PSD to match the stationary loss probability of the scenario with the deterministic PSD and as a result the packet loss process became more correlated) and to the difference between the packet size distributions.

Before we start to investigate in what extent the differences are due to the differ-

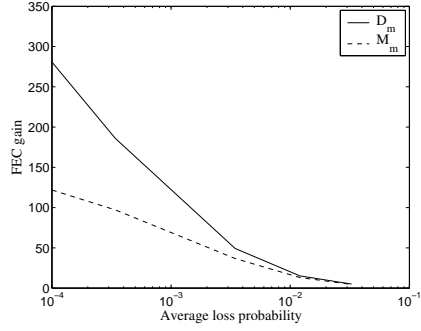


Figure 19: FEC gain vs. stationary loss probability on a 22.5 Mbps link for FEC(20,22).

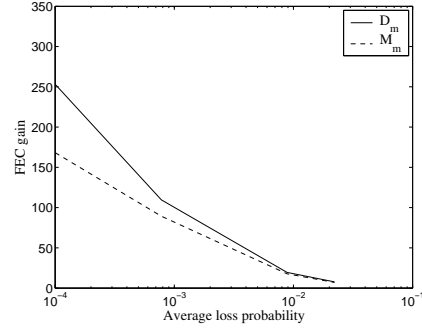


Figure 20: FEC gain vs. stationary loss probability on a 45 Mbps link for FEC(20,22).

ences in the levels of statistical multiplexing and how much to the difference in the PSD, let us compare results with different link speeds. Figures 19 and 20 show the FEC gain on a 22.5 Mbps and a 45 Mbps link respectively. Comparing the figures we can see that the difference between results with the two PSDs in terms of FEC gain decreases as the link speed increases (from 10 Mbps to 45 Mbps). This observation supports the hypothesis that the difference between results with the different distributions is in part due to the different levels of statistical multiplexing. The reason for the phenomenon is that the higher the link speed the less the background traffic has to be changed to keep the stationary loss probability constant, and thus the change in the level of statistical multiplexing decreases as the link speed increases. As an example let us consider the 10 Mbps and 45 Mbps links at a stationary loss probability of 10^{-3} . The background traffic with the deterministic PSD is 3.7 Mbps on the 10 Mbps link, and 38.9 Mbps on the 45 Mbps link. The background traffic of the scenario with exponential PSD has to be set to 1.96 Mbps and 33.5 Mbps on the 10 Mbps and the 45 Mbps links respectively to have an average loss probability of 10^{-3} . Thus the background traffic has to be decreased by 47 percent in the case of the 10 Mbps link while by only 14 percent in the case of the 45 Mbps link.

4.3 Isolating the effects of the packet size distribution

In the following subsection we separate the effects of the level of statistical multiplexing and the PSD. We do it by decreasing the arrival intensity of both the background traffic and the tagged stream in the mathematical model with exponential PSD in order to match the stationary loss probability given by the model with deterministic PSD. Thus we keep both the stationary loss probability and the level of statistical multiplexing constant (doing so is equivalent to matching the stationary loss probability through increasing the link speed).

Figure 21 shows the probability of losing j packets in a block of 22 packets. The average load of the scenarios with deterministic PSD is $\rho = 0.8$, while the average load of the scenarios with exponential PSD is set to match the stationary loss probability of the corresponding scenarios with deterministic PSD and the same link speed. Figure 22 shows the Kullback-Leibler distance for $P(j, 22)$ as a function of the stationary loss probability for the considered three link capacities. Comparing these figures to Figures

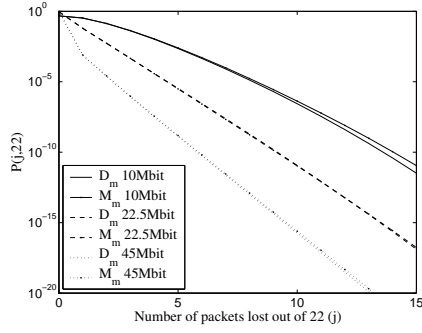


Figure 21: Probability of losing j packets in a block of 22 packets. The stationary loss probability of the scenarios with the same link speed is equal. (same level of statistical multiplexing).

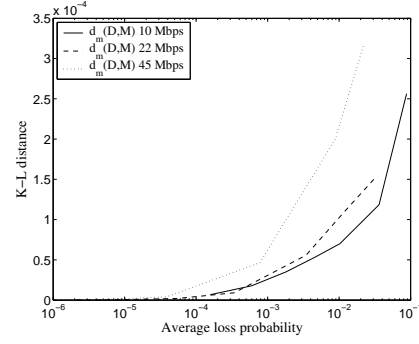


Figure 22: Kullback-Leibler distance vs. stationary loss probability for $P(j,22)$ (same level of statistical multiplexing).

15 and 16 respectively we can see a further decrease in the distance of the results with the different PSDs. Figure 23 shows the average loss run length as a function of the stationary loss probability. Comparing this to Fig. 17 shows a decrease between the results with the two distributions due to the decrease of the average loss run length for the exponential PSD especially at low stationary loss probabilities. The same effect can be seen comparing Fig. 24 to Fig. 18, which shows the FEC gain on a 10 Mbps link as a function of the stationary loss probability for FEC(20,22). Comparing the two figures we see that the FEC gain for the exponential PSD got closer to that for the deterministic one. Thus the difference in the FEC gain considering a particular stationary loss probability is mainly due to the different levels of statistical multiplexing and in a lower extent to the different PSDs.

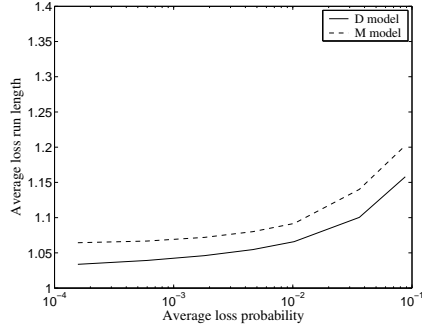


Figure 23: Average loss run length vs. stationary loss probability on a 10 Mbps link (same level of statistical multiplexing).

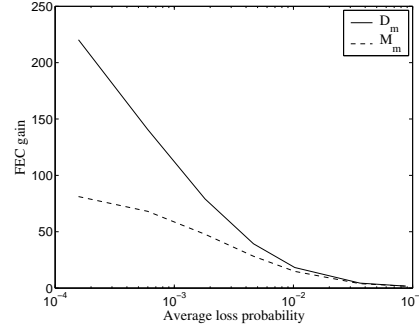


Figure 24: FEC gain vs. stationary loss probability on a 10 Mbps link for FEC(20,22) (same level of statistical multiplexing).

5 Conclusion

In this paper we presented analytical models to evaluate the loss process of a single multiplexer fed by a bursty source and Poisson background traffic for two different packet size distributions and used extensive simulations to validate them. The models can be extended to include bursty background traffic by using an MMPP instead of the Poisson process, and random losses as described in [Dube et al., 2003]. We used the models to compare the probability of losses in a block, the related FEC performance and the average loss run length in the case of exponential and deterministic packet size distributions. Based on the results we conclude that the deterministic packet size distribution not only yields a lower stationary loss probability compared to the exponential packet size distribution but also a more independent packet loss process resulting in a lower average loss run length and more efficient FEC. We showed that the difference between the results at a particular average load level is mainly due to the different stationary loss probabilities while at a particular stationary loss probability it is mainly due to the different levels of statistical multiplexing. We showed that considering a particular stationary loss probability the difference between results obtained with the two distributions decreases as the link speed increases. Thus the packet size distribution influences the efficiency of FEC on a backbone link less than it does in access networks. At the same time in access networks even single multimedia streams can have a significant impact on the characteristics of the aggregate traffic and thus influence the performance of FEC by minimizing the variance of their packet size distribution. Furthermore, we conclude that applications experiencing a certain stationary loss probability in the network can have rough estimates on the performance of FEC independent of the packet size distribution in the network. The results presented here add a new aspect to the evaluation of the potential of FEC and error concealment to correct losses, and serve as a basis for future research on the efficiency of end-to-end error control techniques.

6 Appendices

A Workload distribution

The Laplace transform of the virtual waiting time distribution for the MMPP/G/1/K queue is given in [Blondia, 1989]. Following the arguments presented there one can derive the Laplace transform of the workload distribution

$$\begin{aligned}
 V(s) = & \frac{1}{[\mu - \bar{\pi}_0(\hat{Q} - \hat{\Lambda})^{-1}\bar{e}]} \{ \bar{\pi}_0[-s(I - \hat{\Lambda} + \hat{Q})^{-1}(\hat{Q} - \hat{\Lambda})^{-1} \\
 & + \sum_{k=1}^{N-1} T[\hat{\Lambda}S]^{k-1}(sI + \hat{Q})S[G^*(s)]^k - [G^*(s)]^{N-1} \sum_{k=0}^{N-1} \bar{\pi}_k T[\hat{\Lambda}S]^{N-k-1}] \\
 & + [G^*(s)]^{N-1} \sum_{k=1}^{N-1} \bar{\pi}_k \sum_{j=N-k}^{\infty} \left[\sum_{n=0}^j A_k T[\hat{\Lambda}S]^{n-k} - G^*(s)T[\hat{\Lambda}S]^n \right] \},
 \end{aligned}$$

where $S = (\hat{\Lambda} - sI - \hat{Q})^{-1}$, $T = (sI - \hat{\Lambda} + \hat{Q})^{-1}$, $G^*(s)$ is the Laplace transform of the service time distribution. A_k is an $L \times L$ matrix whose (l, m) th element denotes the conditional probability of the MMPP reaching phase m and having k arrivals during a service time, starting from phase l . Instead of calculating the inverse Laplace transform

of the above expression we use an approximation based on the steady state distribution of the remaining exponential stages in an $MMPP/E_r/1/K$ queue, where E_r denotes an r stage Erlang distribution. r is chosen to be cN , where $c \geq 1$ is an arbitrary whole number and N is defined in Section 3.3. Given $\pi(k, l)$, $0 \leq k \leq cNK$, $1 \leq l \leq L$, the steady state distribution of the remaining exponential stages in the $MMPP/E_r/1/K$ queue we calculate the queue length distribution as seen by an arriving packet as

$$\Pi(k, l) = \frac{\pi(k, l)\lambda_l}{\sum_{l=1}^L \lambda_l \sum_{i=0}^{cNK} \pi(i, l)}. \quad (22)$$

Given the queue length distribution as seen by an arriving packet, $\Pi(k, l)$, $0 \leq k \leq cNK$, $1 \leq l \leq L$, the workload distribution $V(i, l)$, $0 \leq i \leq NK$, $1 \leq l \leq L$ is approximated by

$$V(i, l) = \begin{cases} \Pi(0, l) & i = 0 \\ \sum_{k=(i-1)c+1}^{ic} \Pi(k, l) & 0 < i \leq NK. \end{cases} \quad (23)$$

B Interarrival-time distribution

The probability $Q_i^{l,m}(k)$ denotes the joint conditional probability that between two arrivals from the joint arrival stream there are k exponential service completions out of i and the state of the MMPP at the moment of the next arrival is m given that at the time of the last arrival the MMPP was in state l as defined in Subsection 3.1. To calculate $Q_i^{l,m}(k)$ we will first calculate $f_{l,m}(t)$ the joint probability density function of the interarrival-time distribution of the joint arrival process and that the next arrival will occur in state m of the MMPP given that the last arrival was in state l of the MMPP as it was defined in Subsection 3.2.

Using the notations introduced in Section 3 the Laplace transform $f_{l,m*}(s)$ of $f_{l,m}(t)$ is given by [Fischer and Meier-Hellstern, 1992]

$$f_{l,m*}(s) = \mathcal{L} \left\{ e^{(\hat{Q}-\hat{\Lambda})x} \hat{\Lambda} \right\} = (sI - \hat{Q} + \hat{\Lambda})^{-1} \hat{\Lambda} = \frac{Adj(sI - \hat{Q} + \hat{\Lambda}) \hat{\Lambda}}{det(sI - \hat{Q} + \hat{\Lambda})}, \quad (24)$$

where $Adj(\cdot)$ is the adjoint of (\cdot) . The inverse Laplace transform of (24) can be calculated by finding the roots of $det(sI - \hat{Q} + \hat{\Lambda})$ and performing partial fraction decomposition. If $det(sI - \hat{Q} + \hat{\Lambda})$ has no repeated roots (i.e. $\lambda_i \neq \lambda_j$ if $i \neq j$), then $f_{l,m*}(s)$ can be written as

$$f_{l,m*}(s) = \sum_{j=1}^L \frac{B_j^{l,m}}{s + \beta_j}, \quad (25)$$

where β_j are the roots of $t(s) = det[sI - \hat{Q} + \hat{\Lambda}]$, i.e. $t(s) = \prod_{j=1}^L (s + \beta_j)$, and $B_j^{l,m}$ are the coefficients resulting from the partial fraction decomposition. The roots of $t(s)$ can be found analytically as long as $L \leq 4$ and numerically for $L > 4$. Performing the inverse Laplace transform we get

$$f_{l,m}(t) = \sum_{j=1}^L B_j^{l,m} e^{-\beta_j t}. \quad (26)$$

Based on (26) the infinite integrals in equations (15), (16), (17) and (18) can be calculated as

$$\int_x^\infty f_{l,m}(t) dt = \sum_{j=1}^L \frac{B_j^{l,m}}{\beta_j} e^{-\beta_j x}. \quad (27)$$

Now we turn to the calculation of $Q_i^{l,m}(k)$. We denote with $P^{l,m}(k)$ the joint probability of having k Poisson arrivals with intensity μ between two arrivals from the MMPP and the next arrival from the MMPP coming in state m of the MMPP given that the last arrival came in state l . The z-transform $P^{l,m}(z)$ of $P^{l,m}(k)$ is given by [Kleinrock, 1975]

$$P^{l,m}(z) = \sum_{k=0}^{\infty} \left(\int_0^{\infty} \frac{(\mu t)^k}{k!} e^{-\mu t} f_{l,m}(t) dt \right) z^k = f_{l,m*}(\mu - \mu z). \quad (28)$$

Thus we can express $P^{l,m}(k)$ from (28) by performing the inverse z-transform after the substitution. Using the notation $\alpha_j = 1 + \beta_j/\mu$ and $A_j^{l,m} = B_j^{l,m}/(\mu\alpha_j)$ we get

$$P^{l,m}(k) = \sum_{j=1}^L A_j^{l,m} \frac{1}{\alpha_j^k}. \quad (29)$$

Given the probability $P^{l,m}(k)$ one can express $Q_i^{l,m}(k)$ as

$$\begin{aligned} Q_i^{l,m}(k) &= P^{l,m}(k) & \text{if } k < i \\ Q_i^{l,m}(k) &= \sum_{j=i}^{\infty} P^{l,m}(j) & \text{if } k = i, \end{aligned} \quad (30)$$

so that we have

$$Q_i^{l,m}(k) = \begin{cases} \sum_{j=1}^L A_j^{l,m} \left(\frac{1}{\alpha_j} \right)^k & 0 \leq k < i \\ \sum_{j=1}^L \frac{A_j^{l,m}}{1 - 1/\alpha_j} \left(\frac{1}{\alpha_j} \right)^i & k = i. \end{cases} \quad (31)$$

References

- [Abramov, 2004] Abramov, V. (2004). Asymptotic behavior of the number of lost messages. *SIAM Journal on Applied Mathematics*, 64:746–761.
- [Abramowitz and Stegun, 1972] Abramowitz, M. and Stegun, I. A. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.
- [Ait-Hellal et al., 1999] Ait-Hellal, O., Altman, E., Jean-Marie, A., and Kurkova, I. A. (1999). On loss probabilities in presence of redundant packets and several traffic sources. *Performance Evaluation*, 36-37(1-4):485–518.
- [Altman and Jean-Marie, 1998] Altman, E. and Jean-Marie, A. (1998). Loss probabilities for messages with redundant packets feeding a finite buffer. *IEEE J. Select. Areas Commun.*, 16(5):779–787.
- [Andersen and Nielsen, 1998] Andersen, A. T. and Nielsen, B. F. (1998). A markovian approach for modeling packet traffic with long-range dependence. *IEEE J. Select. Areas Commun.*, 16(5):719–732.
- [Beran et al., 1995] Beran, J., Sherman, R., Taqqu, M., and Willinger, W. (1995). Long-range dependence in variable-bit-rate video traffic. *IEEE Trans. Commun.*, 43(2/3/4):1566–1579.
- [Biersack, 1992] Biersack, E. (1992). Performance evaluation of forward error correction in ATM networks. In *Proc. of ACM SIGCOMM*, pages 248–257.
- [Blondia, 1989] Blondia, C. (1989). The N/G/1 finite capacity queue. *Commun. Statist. - Stochastic Models*, 5(2):273–294.
- [Cao et al., 2002] Cao, J., Cleveland, W. S., Lin, D., and Sun, D. X. (2002). Internet traffic tends toward poisson and independent as the load increases. In *Nonlinear Estimation and Classification*. Springer.

- [Cidon et al., 1993] Cidon, I., Khamisy, A., and Sidi, M. (1993). Analysis of packet loss processes in high speed networks. *IEEE Trans. Inform. Theory*, 39(1):98–108.
- [Dán and Fodor, 2003] Dán, G. and Fodor, V. (2003). Quality differentiation with source shaping and forward error correction. In *Proc. of MIPS'03*, pages 222–233.
- [Dube et al., 2003] Dube, P., Ait Hellal, O., and Altman, E. (2003). On loss probabilities in presence of redundant packets with random drop. *Performance Evaluation*, 53(3-4):147–167.
- [Dube and Altman, 2002] Dube, P. and Altman, E. (2002). Utility analysis of simple FEC schemes for VoIP. In *Proc. of IFIP Networking 2002*, pages 226–239.
- [Fischer and Meier-Hellstern, 1992] Fischer, W. and Meier-Hellstern, K. (1992). The markov-modulated poisson process MMPP cookbook. *Performance Evaluation*, 18(2):149–171.
- [Frossard, 2001] Frossard, P. (2001). FEC performances in multimedia streaming. *IEEE Commun. Lett.*, 5(3):122–124.
- [Gilbert, 1960] Gilbert, E. (1960). Capacity of a burst-noise channel. *Bell Syst. Tech. J.*, 69:1253–1265.
- [Gross and Harris, 1998] Gross, D. and Harris, C. M. (1998). *Fundamentals of Queueing Theory*. Wiley, New York.
- [Grossglauser and Bolot, 1999] Grossglauser, M. and Bolot, J.-C. (1999). On the relevance of long-range dependence in network traffic. *IEEE/ACM Trans. Networking*, 7(5):629–640.
- [Gurewitz et al., 2000] Gurewitz, O., Sidi, M., and Cidon, I. (2000). The ballot theorem strikes again: Packet loss process distribution. *IEEE Trans. Inform. Theory*, 46(7):2595–2599.
- [Heffes and Lucantoni, 1986] Heffes, H. and Lucantoni, D. M. (1986). A markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE J. Select. Areas Commun.*, 4(6):856–868.
- [Hoffman and Fernando, 1998] Hoffman, D. and Fernando, G. (1998). RTP payload format for MPEG1/MPEG2 video. RFC 2250.
- [Kawahara et al., 1994] Kawahara, K., Kumazoe, K., Takine, T., and Oie, Y. (1994). Forward error correction in ATM networks: An analysis of cell loss distribution in a block. In *Proc. of IEEE INFOCOM*, pages 1150–1159.
- [Kleinrock, 1975] Kleinrock, L. (1975). *Queueing Systems*, volume I. Wiley, New York.
- [Klemm et al., 1993] Klemm, A., Lindemann, C., and Lohmann, M. (1993). Modeling IP traffic using the batch Markovian arrival process. *Performance Evaluation*, 54(2):149–173.
- [Kullback, 1959] Kullback, S. (1959). *Information Theory and Statistics*. Wiley, New York.
- [Liang et al., 2003] Liang, Y., Apostolopoulos, J., and Girod, B. (2003). Analysis of packet loss for compressed video: does burst-length matter? In *Proc. of IEEE ICASSP*, pages 684–687.
- [Mohamed et al., 2004] Mohamed, S., Rubino, G., and Varela, M. (2004). Performance evaluation of real-time speech through a packet network: a random neural networks-based approach. *Performance Evaluation*, 57:141–162.
- [Neuts, 1979] Neuts, M. (1979). A versatile markovian point process. *J. Appl. Prob.*, 16:764–779.
- [Robert and Le Boudec, 1997] Robert, S. and Le Boudec, J. (1997). New models for pseudo self-similar traffic. *Performance Evaluation*, 30(1-2):57–68.
- [Ryu and Elwalid, 1996] Ryu, B. and Elwalid, A. (1996). The importance of long-range dependence of VBR video traffic in ATM traffic engineering: Myths and realities. In *Proc. of ACM SIGCOMM*, pages 3–14.
- [Salvador et al., 2003] Salvador, P., Valadas, R., and Pacheco, A. (2003). Multiscale fitting procedure using Markov-modulated Poisson processes. *Telecommunication Systems*, 23(1-2):123–148.

- [Schulzrinne et al., 1992] Schulzrinne, H., Kurose, J., and Towsley, D. (1992). Loss correlation for queues with bursty input streams. In *Proc. of IEEE ICC*, pages 219–224.
- [Yoshihara et al., 2001] Yoshihara, T., Kasahara, S., and Takahashi, Y. (2001). Practical time-scale fitting of self-similar traffic with markov-modulated poisson process. *Telecommunication Systems*, 17(1-2):185–211.