Decreasing Loss Probabilities by Redundancy and Interleaving: A Queueing Analysis

Alain Jean-Marie, Parijat Dube, Damien Artiges, Eitan Altman

Decreasing Loss Probabilities by Redundancy and Interleaving: a queueing analysis

Alain Jean-Marie[a], Parijat Dube[b], Damien Artiges[c] and Eitan Altman[c]

[a]LIRMM, Univ. de Montpellier II/CNRS, 161 Rue Ada, 34392 Montpellier, France

[b]IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA

[c]INRIA, B.P. 93, 06902 Sophia-Antipolis Cedex, France

Abstract

We study forward error correction (FEC) that reduces loss probabilities of messages, based on adding redundant packets and interleaving, as proposed in [9]. Without interleaving, losses occur due to a locality phenomenon: If a packet is lost then the probability of another loss closely after that may be significantly larger than the probability of a loss much later; Thus losses tend to cluster. We study FEC with interleaving using different approaches: Ballot theorems, recursions and a purely algebraic approach, and complement it with numerical investigation.

1. INTRODUCTION

FEC and interleaving are frequently used at a bit-by-bit basis at the link level to recover from channel errors. Yet, they may be useful at a packet level in case where congestion causes packet losses. Simple packet based FEC is already implemented in audio applications [4] and have been proposed also for video [5]. Although FEC can improve the quality of voice [2], the overhead in adding redundancy may decrease the goodput of information (i.e. the throughput of packets well received) [2].

We assume here that losses can be modeled through buffer overflow at a bottleneck link. A sequence of some number n of consecutive packets form a message. Not only is the loss probability of a single packet of interest but also the loss probability of the whole message. In order to reduce the probability of losses of messages, it has been proposed [9] to add redundancy; if j redundant packets are added to a message, then loss only occurs if more than j out of n consecutive packets are lost. We can thus consider new supermessages consisting of n consecutive packets, such that if $k \le j$ packets in a supermessage are lost, they can be reconstructed from the remaining packets. Explicit expressions for the loss probabilities were obtained in [1,3] based on recursive schemes proposed in [6]. The analysis showed that the redundancy mechanism is efficient only for low and for high loads [3], since the gain of adding redundancy was compensated by additional losses caused by the increased transmission rate (due to the extra redundancy).

One reason why the additional redundancy was not so effective is related to the following locality phenomenon. If a packet is lost at time t, then the probability to have another loss closely after t may be significantly larger than the probability of a loss at a time quite larger than t. Hence, if a packet in a message is lost, then with a probability that is high in some sense, other packets will be lost in the same message, and then the additional redundancy fails to avoid the loss of the message. This observation motivated Shacham and McKenney

[9] to introduce the following interleaving scheme. We first add j redundant packets to each message, as before. Again, a message together with its redundant packets is called a supermessage. A group of d consecutive supermessages is called a block. The order of transmission of packets within a block is performed in a round-robin way. Thus, the packets of message k, $k = 1, ..., d$, are transmitted at times $k, k + d, k + 2d, ..., k + (n - 1)d$. Again, loss in a supermessage only occurs if more than j packets in it are lost. Since packets corresponding to the same message are now transmitted in intervals of d packets, the locality effect is expected to be reduced as d increases.

In a recent related work [8], the authors consider an interlaced scheme (like ours combined FEC and interleaving). However they look only at losses of messages (which they call cells) as a whole and not at packets (which they call symbols) constituting the messages. Thus they effectively look at the probability of loss of j messages in a group of n messages. Thus there analysis is similar to the works in [1,3,6] which consider no interleaving and which look at the loss of j packets in a group of n packets.

The structure of the paper is the following. We first analyze the locality phenomenon of losses in Sec. 2. We then obtain expressions for message loss probability in Sec. 3 using three different approaches. The third recursive approach is extended in Sec. 4 to the case of multiple sessions. Numerical investigation appears in Sec. 5. and we conclude in Sec. 6.

2. Locality and Dependence Phenomena

We consider an M/M/1 queue with a buffer of size K. Service is exponential with parameter μ and interarrival times are exponential with parameter λ. A loss occurs whenever an arriving packet finds the queue full. Define $\rho = \lambda/\mu$. Cidon et al. observed in [6] the *dependence phenomenon:* the probability of losing at least one packet out of n consecutive packets is smaller than the probability of losing at least one packet out of n consecutive packets if losses are assumed to be independent (the loss probability of a single packet under the independence loss assumption being equal to the stationary probability of losing a packet in an M/M/1/K queue).

We shall quantify this phenomenon and introduce another related one, called the *locality phenomenon.* Let $P(j, n)$ be the probability of losing j packets out of n consecutive packets. Let $P(> j, n)$ denote the probability of losing more than j packets out of n consecutive packets. Expressions for $P(j, n)$ and $P(> j, n)$ as well as their probability generating functions are given in [3], see also [6] for a recursive scheme. Let $P_C(> j, n)$ ("C" for "correlated") denote the probability of losing more than j packets out of n consecutive packets, given that at least one packet is lost. We have: $P_C(> j, n) = P(> j, n)/P(> 0, n)$. We compare this quantity to the one that would be obtained if losses are assumed to be independent. Let p be the loss probability of an arbitrary packet. It is given by $p = \rho^K R_K$, where $R_K = \left(\sum_{i=0}^{K} \rho^i \right)^{-1}$ (see next section). Let $P_I(> j, n)$ ("I" for "independent") be the probability of more than j losses out of n consecutive packets, given that at least one packet is lost, if losses are assumed to be independent. Then

$$P_I(> j, n) \;\; = \;\; \frac{\sum_{g=j+1}^{n} \binom{n}{g} p^g (1 - p)^{n-g}}{\sum_{g=1}^{n} \binom{n}{g} p^g (1 - p)^{n-g}} = \frac{\sum_{g=j+1}^{n} \binom{n}{g} p^g (1 - p)^{n-g}}{1 - p^n}$$

Define $\alpha(j, n) = P_C(> j, n)/P_I(> j, n)$. When the ratio $\alpha(j, n)$ is larger than one, then the probability $P(> j, n)$ is smaller under the assumption of independent packet losses, which means that the system suffers from the locality of losses, and when $\alpha(j, n) < 1$, the opposite is true, which means that the system benefits from the dependence phenomenon. We now introduce the *locality phenomenon:* for $\rho \leq 1$, $\alpha(j, n) > 1$ for $j = 1, n = 2$. It also holds for any $1 \leq j \leq n - 1$ for sufficiently small ρ.

We show that $\alpha(1, 2) > 1$ for $\rho \leq 1$. Let A_k be the event of having k packets at a time just before packet number r arrives, L_1 be the event of losing the rth packet, and L_2 be

the event of losing the $r + 1$st packet. Observe that $P(A_{K-1}) = p/\rho$. Then

$$P(> 0, 2) = \sum_{k=0}^{K} P(L_1 \cup L_2 | A_k) P(A_k) = p + P(A_{K-1}) P(L_2 | A_{K-1}) = p + \frac{p}{\rho} \frac{\lambda}{\lambda + \mu} = p \frac{2 + \rho}{1 + \rho}$$

$$P(> 1, 2) = P(L_1) P(L_2 | L_1) = p \frac{\lambda}{\lambda + \mu} = \frac{p\rho}{1 + \rho} \ . \tag{1}$$

Hence, $P_C(> 1, 2) = \rho/(2 + \rho)$. On the other hand, $P_I(> 1, 2) = p^2/(1 - p^2)$. A simple calculation shows that for any $0 < \rho \leq 1$, $\alpha(1, 2) > 1$. For $n > 2$, $j \geq 1$, we may expect situations where $\alpha(j, n)$ is either larger or lower than one. For $\alpha(j, n)$ to be smaller than one, the dependence phenomenon has to outweigh the locality phenomenon. The dependence phenomenon can be explained as follows: in the same way that packet losses tend to cluster, "non-losses" also occur in groups. In particular, if a packet arrives when the queue length is small, then the next n consecutive packets all benefit together from this, and few packets will be lost among them. If losses are independent, then the packets will not make advantage of that anymore. For high ρ, the effect of that locality becomes less important with respect to the decrease in the loss probabilities due to the dependence phenomenon.

3. Loss Probabilities: Model and Analysis

We analyse the redundancy scheme with round-robin transmission. Firstly j redundant packets are added to each message. A message together with its redundant packets is called a supermessage. A group of d consecutive supermessages is called a block. The order of transmission of packets within a block is performed in a round-robin way. Thus, the n packets of message k, $k = 1, \ldots, d$ are transmitted at times $k, k+d, k+2d, \ldots, k+(n-1)d$. Again, a loss in a supermessage only occurs if more than j packets in it are lost. Since packets corresponding to the same message are now transmitted in intervals of d packets, the locality effect is reduced as d increases. We call d the *distance* parameter. For the analysis we still consider an M/M/1/K queue. The *aggregate* packet process is therefore a Poisson process with a rate called λ. Packet service times are assumed exponentially distributed with parameter μ. A loss occurs whenever an arriving packet finds the buffer full. Define $\rho = \lambda/\mu$ and assume that the queue is stationary. The loss probability of interest when j redundant packets are added to a message of length $n - j$ is the probability of more than j losses in a supermessage of length n which we denote by $P(> j, n, d)$ as with j redundant packets a message is recovered correctly at the receiver if the number of packets lost (in the channel) is less than or equal to j. Observe that $P(> j, n, d) = 1 - \sum_{h=0}^{j} P(h, n, d)$, where for any h, $P(h, n, d)$ is defined as the probability of h losses in a supermessage of length n (data+redundancy) when packets of a message are interleaved with $d - 1$ packets of other messages. Of course $P(h, n, d) = 0$, for $h > n$. Let $P_i(j, n, d)$ be the probability of j losses in a supermessage of length n given that there are i packets in the buffer just before the arrival of the first packet of the supermessage and $\Pi(i)$ be the stationary probability of having i packets in the buffer. By the *Poisson Arrivals See Time Averages (PASTA)* property, this is also the probability that an arriving packet sees i packets in the queue. Then we have:

$$P(j, n, d) = \sum_{i=0}^{K} \Pi(i) P_i(j, n, d) \ , \quad \text{where } \Pi(i) = \frac{1 - \rho}{1 - \rho^{K+1}} \rho^i \ . \tag{2}$$

3.1. First Approach

Our first approach for evaluating $P_i(j, n, d)$, inspired by [7], uses Ballot theorems. We look at the probability of losses in a tagged supermessage. For the analysis below we

define a *group* as including all the packets of a block arriving at the buffer starting from the first packet of the supermessage till the last packet of the same supermessage. Thus with our scheme there shall be $(n-1)d+1$ packets in a group. We now look at losses in a group [1]. Like in [7], we define three types events, mutually exclusive in each family, as follows:

(i) $\mathcal{V}_i(k_1)$: event that the first packet of the group to be lost is k_1 given that upon the arrival of the first packet of the group there are i packets in the buffer.

(ii) $\mathcal{S}(k_l, k_{l+1})$: event that packet k_{l+1} is lost given that packet k_l was lost.

(iii) $\mathcal{U}(k_j)$: event that packet k_j is the last to be lost.

Let $v_i(k_1)$, $s(k_l, k_{l+1})$ and $u(k_j)$ be the probabilities of events $\mathcal{V}_i(k_1), \mathcal{S}(k_l, k_{l+1})$ and $\mathcal{U}(k_j)$. We next consider the group as a whole. Let us define $P_i^g(z, y)$ as the probability of z losses in a group of y packets given that the queue length just before the arrival of the first packet of the group is i. Let k_g, $1 \leq g \leq z$ be the positions of the z losses in the group. We have following Lemma from [7], a consequence of the strong Markov property of the process.

Lemma 1 ([7]) *The expression for $P_i^g(z, (n-1)d+1)$ can be written as:*

$$\sum_{k_1=1}^{[(n-1)d+1-z+1]} \sum_{k_2=k_1+1}^{[(n-1)d-z+2]} \cdots \sum_{k_z=k_{z-1}+1}^{(n-1)d} v_i(k_1)s(k_1, k_2) \ldots s(k_{z-1}, k_z)u(k_z) \ .$$

Observe that the probabilities $v_i(k_1), s(k_l, k_{l+1})$ and $u(k_j)$ can be evaluated using Ballot theorems (see [7]). Let f_m, $1 \leq m \leq j$ be the positions of the lost packets of the supermessage where f_1 takes value in the set $\mathcal{F}_1 = \{1, (1+d), \ldots, (1+(n-(j-1)-1)d)\}$, f_2 takes value in $\mathcal{F}_2 = \{(f_1+d), \ldots, (1+(n-j-1)d)\}$ and in general f_m takes value in $\mathcal{F}_m = \{(f_{m-1}+d), \ldots, 1+(n-(j-m)-1)d\}$. Let $\mathcal{F}$ be the set $\{f_1, f_2, \ldots, f_j\}$. Let $\mathcal{K}$ be the set $\{k_1, \ldots, k_z\}$.

Proposition 1 ([7]) $P_i(j, n, d)$ *is given by:*

$$\sum_{f_1 \in \mathcal{F}_1} \cdots \sum_{f_j \in \mathcal{F}_j} \sum_{z=j}^{(n-1)d+1} P_i^g(z, (n-1)d+1) \mathbf{1}(\mathcal{F} \subseteq \mathcal{K}) \ , \tag{3}$$

where $\mathbf{1}(A)$ is the indicator function for event A, and $P_i^g(.)$ is given by Lemma 1.

Thus knowing $P_i(j, n, d)$ we can calculate $P(j, n, d)$ from (2).

3.2. Second Approach

We here present a more direct approach for obtaining $P_i(j, n, d)$ which does not (explicitly) involve the results of [7] based on Ballot theorems. We define following types of events associated exclusively with the packets belonging to the supermessage [2]:

(i) $\mathcal{V}_i^m(k_1)$: event that the first packet (of the supermessage) to be lost is k_1 given that upon the arrival of the first packet of the supermessage there are i packets in the buffer.

(ii) $\mathcal{S}^m(k_l, k_{l+1})$: event that packet k_{l+1} is lost given that packet k_l was lost.

(iii) $\mathcal{U}^m(k_j)$: event that packet k_j is the last to be lost.

Observe that if packet k_i belongs to the tagged supermessage then its position (if we number the packets in a group from 1 to $(n-1)d+1$) in the group is of the form $(f_i-1)d+1$, where $f_i = 1, \ldots, n$. Let $v_i^m(k_1)$, $s^m(k_l, k_{l+1})$ and $u^m(k_j)$ be the probabilities of events $\mathcal{V}_i^m(k_1), \mathcal{S}^m(k_l, k_{l+1})$ and $\mathcal{U}^m(k_j)$. We have the following:

---

[1] In order to compute the loss probability of a super message we can assume without loss of generality that it is the first one in a block. The loss probability of the supermessage will not depend on this assumption

[2] We shall be using the superscript m in the notations just to distinguish them from notations in the first approach.

Lemma 2 $P_i(j, n, d)$ *is given by:*

$$\sum_{f_1=1}^{n} \sum_{f_2=1}^{n} \ldots \sum_{f_j=1}^{n} v_i^m(k_1) s^m(k_1, k_2) \ldots s^m(k_{j-1}, k_j) u(k_j) \tag{4}$$

or in matrix form: $P_i(j, n, d) \;=\; \boldsymbol{V_i^m} \, (\boldsymbol{S^m})^{j-1} \, (\boldsymbol{U^m})^T \tag{5}$

where $\boldsymbol{V_i^m}$ and $\boldsymbol{U^m}$ are n-dimensional row vectors with elements $v_i^m(.)$ and $u^m(.)$, respectively, and $\boldsymbol{S^m}$ is an $n \times n$ matrix with elements $s^m(.,.)$.

Proof: From the definition we have :

$$P_i(j, n, d) \;=\; \sum_{f_1=1}^{n-j+1} \sum_{f_2=f_1+1}^{n-j+2} \ldots \sum_{f_j=f_{j-1}+1}^{n} v_i^m(k_1) s^m(k_1, k_2) \ldots s^m(k_{j-1}, k_j) u^m(k_j) \tag{6}$$

where $k_g = (f_g - 1)d + 1$, $g = 1, 2, \ldots, j$. Also by definition $s^m(k_l, k_{l+1}) = 0$ when $f_l \geq f_{l+1}$ for $1 \leq l \leq j - 1$. Thus (4) and hence (5) follows from (6). We next proceed to compute the elements of $\boldsymbol{V_i^m}, \boldsymbol{U^m}$ and $\boldsymbol{S^m}$.

(i) *Expression for $v_i^m(k_1)$:* If the first packet of the supermessage arrives at a full buffer, it is lost and no other packet can be the first to be lost. Let us number the packets of the group starting with the first packet of the (tagged) supermessage. Thus the first packet of the supermessage has position 1, the second packet has position $d + 1$ and in general the mth packet will have position $(m - 1)d + 1$. Let $k_1 = (f_1 - 1)d + 1$, for some f_1. If k_1 is the first packet to be lost then the queue length just before the arrivals of $1, \ldots, f_1 - 1$ packets of the supermessage should not be K. We first calculate the d step probability of going from a state i (the number of packets just before the arrival of the first packet of the supermessage) to state j (the number of packets just before the arrival of the second packet of the supermessage or $d+1$th packet of the group). Let P be the transition matrix associated with the queue length (seen by an arrival) process of the underlying $M/M/1/K$ queue. Then the probability that the number of packets just before the arrival of the first packet of the supermessage is i (with $i \neq K$) and the number of packets just before the arrival of the second packet of the supermessage is j (with $j \neq K$) is simply the element ij of the d-step transition matrix P^d of the embedded Markov chain (of the queue length process) at arrival epochs in a $M/M/1/K$ queue. The transition probabilities p_{ij} are well known (see e.g. [6]): with $\alpha = 1/(1 + \rho)$,

$$i < K: \quad p_{ij} \;=\; \begin{cases} \alpha^{i+1} & \text{if } j = 0 \\ \rho\, \alpha^{i+2-j} & \text{if } 0 < j \leq i+1 \\ 0 & \text{o.w.}, \end{cases} \qquad p_{Kj} = \begin{cases} \alpha^{K} & \text{if } j = 0 \\ \rho\, \alpha^{K+1-j} & \text{if } 0 < j \leq K \\ 0 & \text{o.w.} \end{cases}$$

Thus $v_i^m(k_1)$ can be written as:

$$v_K^m(k_1) \;=\; \begin{cases} 1 & \text{if } k_1 = 1 (f_1 = 1) \\ 0 & \text{otherwise (o.w.)} \end{cases}$$

and $v_i^m(k_1) = 0$ for $k_1 \leq K - i$, for $k_1 = (f_1 - 1)d + 1$, $f_1 > 2$,

$$v_i^m(k_1) \;=\; \sum_{j_1=0}^{K-1} (P^d)_{i j_1} \sum_{j_2=0}^{K-1} (P^d)_{j_1 j_2} \ldots \sum_{j_{f_1-2}=0}^{K-1} (P^d)_{j_{f_1-3} j_{f_1-2}} (P^d)_{j_{f_1-2} K}$$

and for $k_1 = (f_1 - 1)d + 1$, $f_1 = 2$, $v_i^m(k_1) = (P^d)_{iK}$. In order to simplify this expression, we next define a matrix $\bar{P}$ as the matrix obtained from P^d matrix by replacing the Kth row and Kth column in P^d matrix with 0s. Then we can write $v_i^m(k_1)$ as:

$$
v_i^m(k_1) \;=\; \begin{cases} 0 & \text{if } k_1 \le K - i \\[2mm] \displaystyle\sum_{j_{f_1-2}=0}^{K-1} (\bar{P})_{ij_{f_1-2}}^{f_1-2} (P^d)_{j_{f_1-2}K} \;=\; \left((\bar{P})^{f_1-2}\, P^d\right)_{iK} & \text{o.w.} \end{cases}
$$

(ii) *Expression for* $s^m(k_l, k_{l+1})$: If packet k_l is lost then the queue length just before the arrival of the k_lth packet is K, similarly if packet k_{l+1} is lost the queue length just before the arrival of this packet should also be K. Let $k_l = (f_l - 1)d + 1$ and $k_{l+1} = (f_{l+1} - 1)d + 1$. Thus there are $(f_{l+1} - f_l - 1)$ packets of the supermessage between packets numbered k_l and k_{l+1} arriving at the buffer that are not lost. Thus we can write the expression for $s^m(k_l, k_{l+1})$ again in terms of the transition matrix P^d as:

$$
\sum_{j_{f_l}=0}^{K-1} (P^d)_{Kj_{f_l}} \sum_{j_{f_l+1}=0}^{K-1} (P^d)_{j_{f_l}j_{f_l+1}} \cdots \sum_{j_{f_{l+1}-2}=0}^{K-1} (P^d)_{j_{f_{l+1}-3}j_{f_{l+1}-2}} (P^d)_{j_{f_{l+1}-2}K}
$$

Again we can express $s^m(k_l, k_{l+1})$ in terms of $\bar{P}$ as:

$$
s^m(k_l, k_{l+1}) \;=\; \sum_{j_{f_l}=0}^{K-1} (P^d)_{Kj_{f_l}} \sum_{j_{f_{l+1}-2}=0}^{K-1} (\bar{P})_{j_{f_l}j_{f_{l+1}-2}}^{f_{l+1}-f_l-2} (P^d)_{j_{f_{l+1}-2}K} = \left(P^d\, (\bar{P})^{f_{l+1}-f_l-2}\, P^d\right)_{KK}
$$

(iii) *Expression for* $u^m(k_j)$: If $k_j = (n-1)d + 1$, that is k_j is the last packet of the supermessage (and hence of the group) then no packets can be lost after it, i.e., $u^m((n-1)d+1) = 1$. Also if k_j is lost then the queue length just before the arrival of k_j should be K. If $k_j = (f_j - 1)d + 1$, then there are still $n - f_j$ more packets of the supermessage that will come to the buffer and if k_j is the last packet to be lost then none of these $n - f_j$ packets should be lost. Thus we can write the expression for $u^m(k_j)$ as:

$$
\begin{cases} 1 & k_j = (n-1)d + 1 \\[2mm] \displaystyle\sum_{j_{f_j}=0}^{K-1} (P^d)_{Kj_{f_j}} \sum_{j_{f_j+1}=0}^{K-1} (P^d)_{j_{f_j}j_{f_j+1}} \cdots \sum_{j_{n-1}=0}^{K-1} (P^d)_{j_{n-2}j_{n-1}} & k_j < (n-1)d + 1. \end{cases}
$$

And finally $u^m(k_j)$ can be written as:

$$
\begin{cases} 1 & k_j = (n-1)d + 1 \\[2mm] \displaystyle\sum_{j_{f_j}=0}^{K-1} (P^d)_{Kj_{f_j}} \sum_{j_{n-1}=0}^{K} (\bar{P})_{j_{f_j}j_{n-1}}^{n-f_j-1} \;=\; \sum_{j_{n-1}=0}^{K} \left(P^d\, (\bar{P})^{n-f_j-1}\right)_{Kj_{n-1}} & k_j < (n-1)d + 1 \end{cases}
$$

3.3. Third Approach: Using Recursions

It is also possible to write recursions to calculate $P_i(j, n, d)$. Those recursions generalize the results in [6], which was a special case of our model with $d = 1$. Let us again denote the element (i, k) of P^d by $(P^d)_{i,k}$. Then we have the following relations:

$$
i < K: \; P_i(j, 1, d) \;=\; \begin{cases} 1 & j = 0 \\ 0 & j \ge 1 \end{cases} \qquad P_K(j, 1, d) = \begin{cases} 1 & j = 1 \\ 0 & j = 0, \end{cases} \quad j \ge 2 \; . \tag{7}
$$

For $n \geq 2$ we have:

$$P_i(j,n,d) \;=\; \begin{cases} \displaystyle\sum_{k=0}^{K} (P^d)_{i,k}\, P_k(j,n-1) & 0 \leq i \leq K-1 \\[2em] \displaystyle\sum_{k=0}^{K} (P^d)_{Kk}\, P_k(j-1,n-1) & i = K. \end{cases} \tag{8}$$

4. Multiple Sessions

We next deal with case of multiple sessions and propose a recursive procedure for computing the losses. We note, however, that in presence of multiplexing between a large number of sessions, consecutive packets of a given session are already separated by a large number of other packets, which results in an effect similar to the one expected by interleaving. Assume that packets arrive to the system from S independent sources each implementing interleaving and redundancy. The total arrival process to the system from source s, $s = 1, 2, \ldots, S$ is assumed to be Poisson with rate λ_s. The overall arrival process to the system is also Poisson with rate $\lambda = \sum_{s=1}^{S} \lambda_s$. Let d_s and n_s be the distance parameter and the supermessage length for source s. Denote by $P_i^s(j,n_s,d_s)$ (resp. $P_i^{\bar{s}}(j,n_s,d_s)$), the probabilities of j losses in a supermessage of n_s packets originating from source s, given that there are i packets in the system just before the arrival of the first packet in the supermessage (resp. just before the arrival of a packet from any other source (denoted by $\bar{s}$)). The probability of an arrival in the overall Poisson arrival stream being from source s is equal to λ_s/λ. Define $p = \lambda_s/\lambda$ and $\bar{p} = 1 - p$. We next provide recursive expressions for computing $P_i^s(j,n,d_s)$ for $s \in \{1, 2, \ldots, S\}$:

$$P_i^s(j,1,d_s) \;=\; \begin{cases} 1 & j = 0, \quad i = 0, 1, \ldots, K-1 \\ 0 & j \geq 1 \end{cases}, \qquad P_K(j,1,d_s) = \begin{cases} 1 & j = 1 \\ 0 & j = 0, \quad j \geq 2. \end{cases}$$

We shall condition on the number of packets arriving between two consecutive packets of the same supermessage. Let k be the queue length seen by an arriving packet of the tagged supermessage. For $n_s \geq 2$ we have the recursive equations for $0 \leq i \leq K-1$: For $n_s \geq 2$ we have the recursive equations for $0 \leq i \leq K-1$:

$$P_i^s(j,n_s,d_s) = \sum_{k=0}^{K} \sum_{z=(d_s-1)}^{\infty} \left[\binom{z}{d_s-1} p^{d_s-1} \bar{p}^{z-(d_s-1)} (P^z)_{i,k} \right] \left[p P_k^s(j,n_s-1,d_s) + \bar{p} P_k^{\bar{s}}(j,n_s-1,d_s) \right]$$

and for $i = K$:

$$\begin{aligned} P_K^s(j,n_s,d_s) \;=\; & \sum_{k=0}^{K} \sum_{z=(d_s-1)}^{\infty} \left[\binom{z}{d_s-1} p^{d_s-1} \bar{p}^{z-(d_s-1)} (P^z)_{Kk} \right] \\ & \times \left[p P_{K-k}^s(j-1,n_s-1,d_s) + \bar{p} P_{K-k}^{\bar{s}}(j-1,n_s-1,d_s) \right] \end{aligned} \tag{9}$$

where $P_i^{\bar{s}}(j,n_s,d_s)$ for $n_s \geq 1$ is given by:

$$\begin{cases} \sum_{k=0}^{K} p_{ik} \left[p P_k^s(j,n_s,d_s) + \bar{p} P_k^{\bar{s}}(j,n_s,d_s) \right] & 0 \leq i \leq K-1 \\ P_{K-1}^{\bar{s}}(j,n_s,d_s) & i = K \end{cases}$$

Having obtained $P_i^s(j,n_s,d_s)$ one can obtain $P^s(j,n_s,d_s) = \sum_{i=0}^{K} \Pi(i) P_i^s(j,n_s,d_s)$ where $\Pi(i)$ is given by (2) with $\rho = \lambda/\mu = \sum_{s=1}^{S} \lambda_s/\mu$ and hence $P^s(>j,n_s,d_s)$.
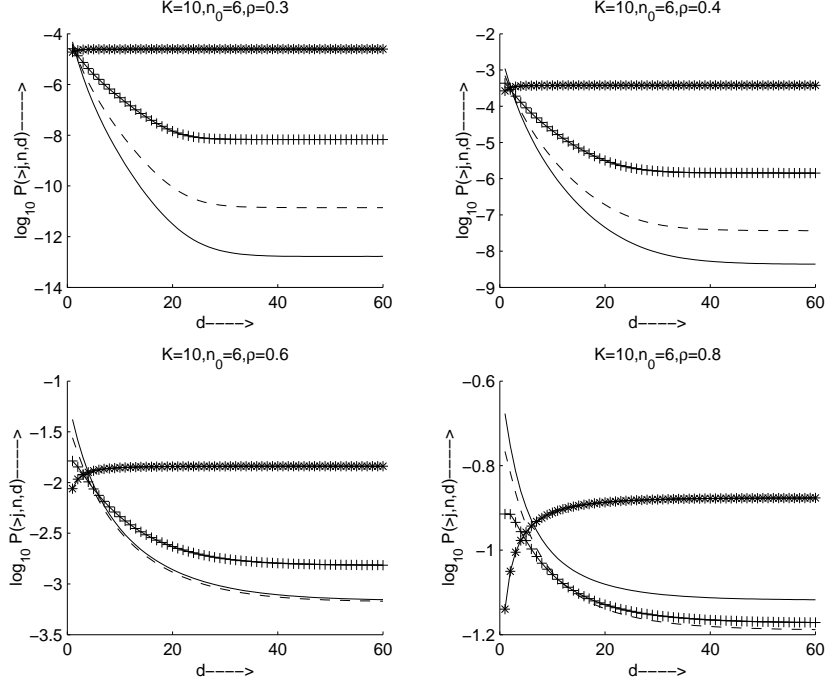
Figure 1. $\log_{10} P(> j, n, d)$ for $\rho = 0.3, 0.4, 0.6, 0.8$ as a function of d. We represent the loss probabilities for $j = 0$ by $*-$, for $j = 1$ by $+-$, for $j = 2$ by $--$ and for $j = 3$ by $-$.

5. Numerical results

We present below numerical results for messages of length $n_0 = 6$, and queue of capacity $K = 10$. We plot the loss probabilities of messages as a function of the distance parameter d and the number of redundant packets j in each message of n_0 packets using the third approach. Thus the length of the supermessage is $n = n_0 + j$. Recall that the loss probability we are looking for when j is the amount of redundancy is $P(> j, n, d)$. Each figure is given with a different load ρ. In order that loss probabilities corresponding to different redundancy numbers j be comparable, we use the following convention on ρ. Let $\rho(j)$ denote the load function (λ/μ) of the queue that corresponds to a redundancy of j. The given ρ in each figure corresponds to the case of $j = 0$, i.e., $\rho = \rho(0)$. Different curves in the same figure that correspond to different redundancies j, have in common the same rate at which messages arrive. Therefore, the actual rate $\rho(j)$ corresponding to redundancy j in each figure is given by $\rho(j) = \rho(n_0 + j)/n_0 = \rho n/n_0$. Note that adding redundancy indeed increases the rate at which packets arrive, if the rate at which messages arrive is unchanged, which has as an effect to increase loss probabilities.

Light and medium load: $\rho \leq 0.8$. We begin by analyzing the case $\rho < 1$. [3] argues that for $d = 1$, redundancy results in deteriorating the loss probabilities for $\rho > 0.31$ with the assumed values for n_0 and K. For ρ larger than 2 redundancy became again profitable. We were thus especially interested in the range $0.3 < \rho < 1$. Fig. 1 considers $\rho = 0.3, 0.4, 0.6, 0.8$. We observe that for $j \geq 1$, the loss probabilities considerably decrease in d for a given load and after some large d it remains nearly constant. The gain is more prominent at low loads and decreases with increasing loads. E.g., for $\rho = 0.4$, $d = 40$, the gain (as compared to the loss probability without interleaving, at $d = 1$) is of a factor 10^3, 10^4, 10^5 for $j = 1, 2, 3$ resp. and for $d = 60$ the gain factor is pretty much the same as for $d = 40$. For $\rho = 0.8$, the gain at $d = 40$ is reduced by a factor of 10
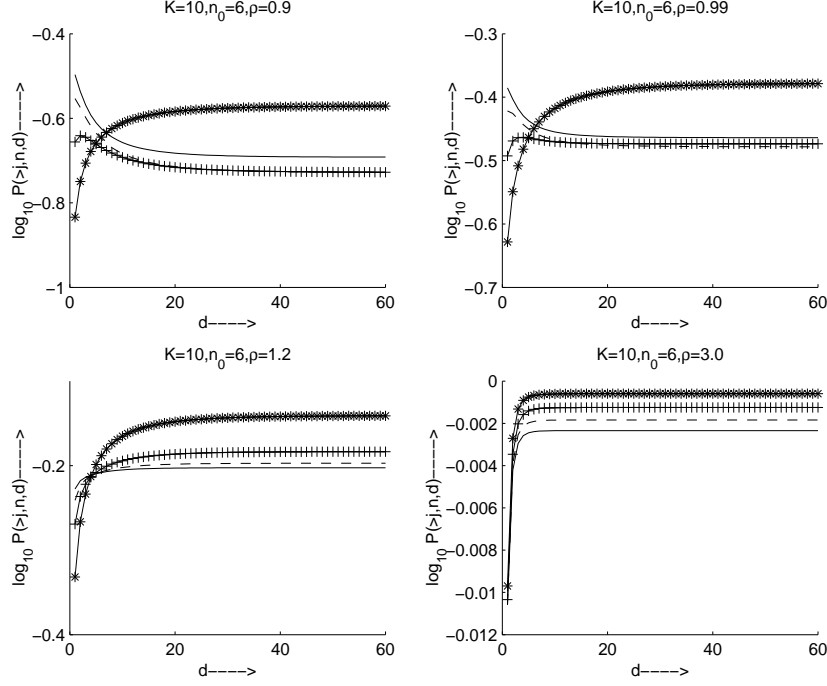
Figure 2. $\log_{10} P(> j, n, d)$ for $\rho = 0.9, 0.99, 1.2, 3.0$ as a function of d. We represent the loss probabilities for $j = 0$ by $*-$, for $j = 1$ by $+-$, for $j = 2$ by $--$ and for $j = 3$ by $-$.

for all $j \geq 1$. We further observe that for $\rho = 0.3, 0.4$, the higher the redundancy, the smaller are the loss probabilities, at least for $j \leq 3$. *An interesting phenomenon is the non-monotonicity of losses with j and d, calling for the need of optimal design parameters.* We note that for each d, there is an optimal redundancy such that for smaller redundancy or higher redundancy the loss probability is higher. For higher redundancy, the effect of increasing the rate of packet arrivals (so as to keep the message arrival rate unchanged after adding redundancy) due to the additional redundancy, becomes dominant, i.e. more important than the gain due to the possibility of reconstructing lost packets by the additional redundancy. Conversely, for lower redundancy, the effect of having less redundancy becomes more important and causes increase in the loss probabilities. For low d, the optimal redundancy is 0, and it increases with d; it remains fixed for $d \geq 5$. For $\rho = 0.3, 0.4$ the optimal redundancy is 3 is for $d = 1$ and for $\rho = 0.6$, it is 2 which has almost the same performance as 3. For $\rho = 0.8$ the gain in adding redundancy and increasing d still exists, but it becomes quite small. For large d, the optimal redundancy is $j = 2$ and it only slightly outperforms redundancy 1; $j = 3$ is strictly worse than $j = 1, 2$.

Heavy and High load: $\rho \geq 0.9$ For ρ larger than 0.9 we always get the best performance for $d = 1$. This is illustrated in Fig. 2. Yet, as was observed in [3], for very high load, redundancy decreases the loss probabilities, as can be seen from this same figure.

Note that in practice, a well designed flow control CAC (Call Admission Control) will not allow operation under heavy and high load conditions. This is the case in some ATM applications, where loss probabilities are required to be of the order of 10^{-9}.

The fact that for heavy and high load $d = 1$ is optimal can be explained by the Dependence phenomenon: when the packet loss probability becomes quite high, if a packet is not lost, then the probability that neighboring packets are not lost is higher than that of a typical packet. Therefore it is desirable not to disperse messages, i.e. to use $d = 1$.

K	n_0	$\rho = 0.4$		$\rho = 0.9$		$\rho = 3.0$	
		j^*	d^*	j^*	d^*	j^*	d^*
10	5	3	20	0	1	2	1
	10	3	20	3	20	3	1
20	5	2	20	0	1	2	1
	10	3	20	0	1	3	1
30	5	1	20	0	1	2	1
	10	3	15	0	1	3	1

Table 1

Optimal j^* and d^* for message length n_0, buffer size K and ρ.

Optimal Redundancy and Interleaving Based on the numerical study of our model we observe that the optimal (here the optimality criterion is minimising the message loss probability) redundancy j^* and the optimal distance parameter for interleaving d^* depends on the load for a given message length and buffer size. Thus based on our analysis numerical calculations can be done for different sets of parameters n_0, K, ρ and tables/charts for d^* and j^* can be obtained. These may be useful in providing "engineering" guidelines for designing (near) optimal FEC schemes. In the Table 1 we give the values of j^* and d^* for buffer sizes $10, 15, 20$, message sizes $5, 10$ and three loads $0.4, 0.9, 3.0$. Here we have restricted the value of d^* to be in $[1, 20]$.

6. Conclusion

In this paper through a simple $M/M/1/K$ model we analysed the (congestion) loss probabilities of packets. We quantified two phenomena, the *dependence* and the *locality*. We then proposed three different approaches for obtaining expressions for loss probabilities $P(j, n, d)$ in the presence of interleaving and redundancy. In passing, we establish links between combinatorial methods based on the Ballot theorem, and algebraic methods based on transition matrices.

REFERENCES

1. O. Ait Hellal, E. Altman, A. Jean-Marie and I.A. Kurkova, "On Loss Probabilities in Presence of Redundant Packets and Several Traffic Sources", *Performance Evaluation*, pp. 486-518, Vol 36-37, 1999.
2. E. Altman, C. Barakat and V. M. Ramos-Ramos, "Queuing analysis of simple FEC schemes for Voice over IP", *Comm. Networks*, **39**(2), 185-206, June, 2002.
3. E. Altman and A. Jean-Marie, "The Loss Process of Messages in an M/M/1/K Queue", *In Proc. of Infocom '94*, Toronto, Canada, pp. 1191-1198, June 1994.
4. J.-C. Bolot, S. Fosse-Parisis and D. Towsley, "Adaptive FEC-Based Error Control for Interactive Audio in the Internet", *In Proc. Infocom '99*, New York, NY, March 1999.
5. J-C. Bolot and T. Turletti, "Adaptive Error Control for Packet Video in the Internet", *Proc. IEEE International Conference on Image Processing ICIP'96*, Lausanne, 1996.
6. I. Cidon, A. Khamisy and M. Sidi, "Analysis of Packet Loss Processes in High-Speed Networks", *IEEE Trans. on Inform. Theory*, Vol. 39, No. 1, pp. 98-108, 1993.
7. O. Gurewitz, M. Sidi and I. Cidon, "The Ballot Theorem Strikes Again: Packet Loss Process Distribution", *IEEE Trans. on Infor. Theory*, Vol. 46, No. 7, Nov. 2000.
8. R. Kurceren and J. W. Modestino, "Use of FEC Coding for Real-Time Transport on ATM Networks," manuscript from private communications.
9. N. Shacham and P. McKenney, "Packet Recovery in High-Speed Networks Using Coding and Buffer Management," *In Proc. of Infocom '90*, SF0, CA, pp. 124-131, May 1990.