Throughput Analysis of Stop-and-wait Retransmission Schemes for k-reliable Multicast

Bongtae Kim, Arne A. Nilsson
Department of Electrical and Computer Engineering &
Center for Communications and Signal Processing
North Carolina State University
Raleigh, NC 27695

Harry G. Perros
Department of Computer Science &
Center for Communications and Signal Processing
North Carolina State University
Raleigh, NC 27695

April 18, 1994

Abstract

This paper considers a class of stop-and-wait multicast retransmission schemes in which a message is considered to be successfully transmitted when a predefined number of acknowledgements (less than the number of receivers) is received. These retransmission schemes can be used for applications which need high throughput with an acceptable level of reliability. Four different stop-and-wait retransmission schemes are presented and analytical expressions of throughput for each scheme are derived. Throughput comparisons are provided through numerical examples.

1 Introduction

Various retransmission strategies proposed in the literature fall into one of two main categories: stop-and-wait schemes and continuous schemes. The basic idea of a stop-and-wait scheme is to ensure that each message has been received correctly before transmitting the next message. In continuous schemes, up to a predefined number of messages can be sent by the transmitter without receiving an acknowledgement. Among the continuous schemes, a further distinction is made between schemes where errored messages and all subsequent messages are retransmitted (such as go-back-N), and schemes where only errored messages are retransmitted (such as selective-repeat). All variants of these error control schemes have been evaluated for point-to-multipoint communication environments.

Calo and Easton [2] proposed a broadcast protocol for large file transfer in a satellite network. The protocol is basically a stop-and-wait multidestination scheme. In their protocol, the file is divided into blocks of uniform size and each block, in turn, is subdivided into N frames. The transmitter sends all N frames and receivers respond with acknowledgement frames. The first cycle of the block transfer ends when sufficient time has elapsed for all acknowledgment frames to reach the transmitter. The second cycle begins as the sender starts to retransmit those frames not yet acknowledged by all receivers. When all frames have been acknowledged by all receivers, the last cycle ends. Analytical expressions for the throughput were obtained for two special cases: 1) the up-link is error free; 2) the acknowledgements are error free. For the general case, upper and lower bounds are derived.

Mase *et al.* [4] described a go-back-N scheme, termed "end-to-end error control scheme", suitable for point-to-multipoint satellite communication by using a control message to indicate the occurrence of retransmission to all of the receivers. They also presented a "tandem error control scheme", where the up-link and the down-link use separate go-back-N error control schemes. The idea is to decrease the round trip delay so that the system throughput can be increased. The performance analysis is done by simulation except for the case of a single receiver.

Gopal and Jaffe [3] presented three different go-back-N protocols suitable for point-to-multipoint communication. In their work, the transmitter maintains *ack-outstanding lists* of transmitted messages which contain the identity of all receivers from whom acknowledgements are expected for those messages. The ack-outstanding lists are updated as new acknowledgements are received and when the ack-outstanding list is empty, the transmitter transmits the next message. Otherwise it retransmits the message and all messages which follow it in the transmission sequence. The three protocols, which are memoryless, limited memory and full memory, differ in the way that they maintain and update the ack-outstanding lists. Analytic expressions for the throughput of these three schemes are given. For the scheme that can achieve the highest throughput, the full memory scheme, an embedded Markov chain is solved in order to obtain the throughput. An exact solution is obtained only for the case of two receivers, and an approximation solution is presented to solve the case of more than two receivers.

Wang and Silvester [5] proposed a scheme, in which the sender transmits multiple copies of the same message to the receivers in order to maximize throughput. Using dynamic programming, they found the optimal number of copies in terms of round-trip propagation delay, the error probability and the number of receivers that have not yet received the message. In [6] the same authors analyzed the delay performance of this scheme.

Ammar and Wu [1] proposed a scheme in which the set of destinations is split into disjoint groups. The transmitter carries a separate conversation with each group. The conversations are time multiplexed over a single channel. They derived expressions for the maximum throughput achievable with their protocols. They also addressed issues on the optimal grouping of destinations in order to maximize throughput. Their results indicate that the destination set splitting can improve the throughput of point-to-multipoint error control protocols, particularly if the receivers' capabilities are not identical.

These reliable schemes guarantee both error free and ordered delivery of messages to all participating receivers. However, their throughput can be substantially reduced when one (or more) of the participating receivers is slower than the others. This can be due to network congestion or limited buffer and/or processing at the receiver. Furthermore, it is possible that some applications may not need acknowledgements all the time from all receivers. Rather, they may need a high throughput with an acceptable level of reliability.

In a view of these consideration, in this paper, we propose a class of stop-and-wait multicast retransmission schemes in which a message is considered to be correctly transmitted, when a predefined number of acknowledgements are received. This predefined number of acknowledgements (hereafter referred to as the *ACK threshold*) is less than the number of receivers. Thus, if there are R receivers and the ACK threshold is k, $k < R$, the transmitter will assume that a message has been correctly transmitted at the moment when it receives the k-th acknowledgement. When the ACK threshold is set to the number of receivers, the error control scheme becomes fully reliable.

In the following section, we briefly describe our proposed retransmission schemes. In section 3, we develop analytical expressions of throughput for four different schemes applying order statistics. Numerical examples are given in section 4. We conclude our work in section 5.

2 The Stop-and-wait Retransmission Schemes

We consider a system with one transmitter and R receivers. All data are transmitted in the form of messages or clearly delimited blocks of information. A message may have an arbitrary but bounded length. All messages have a sequence number (SN) which uniquely identifies the messages. The sequence number field in a message is large enough to permit the detection of duplicate message at the receiver. The round-trip delay is bounded. A message also contains an error checking code which enables each receiver to detect transmission errors.

As in the go-back-N protocols by Gopal and Jaffe [3], and by Wang and Silvester [5], we define the following schemes: (a) memoryless stop-and-wait scheme with single copy transmission (ML-S), (b) full-memory stop-and-wait scheme with single copy transmission (FM-S), (c) memoryless stop-and-wait scheme with m-copy transmission (ML-M), and (d) full-memory stop-and-wait scheme with m-copy transmission (FM-M). These schemes are explained below.

Operation of A Receiver

When a receiver receives a message, it first checks for errors. A message received in error is discarded. If an error-free message has a correct sequence number (the next or higher sequence number of the previously accepted message), the message is accepted and acknowledged with the sequence number of the received message and the identity of the receiver. If an error-free message has the same sequence number as the previously accepted message, the message is acknowledged and then discarded.

Operation of A Transmitter

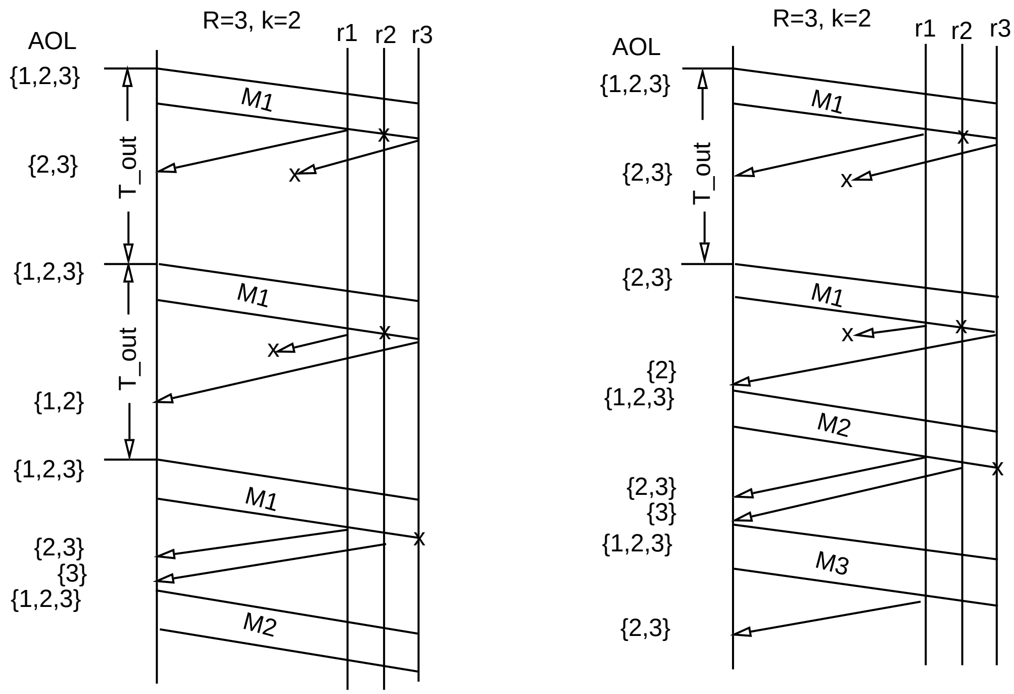
For each scheme, the transmitter operates differently. We first describe the functions that are common to all schemes. The transmitter starts a timer immediately upon transmission of a message. If the message is transmitted for the first time, the transmitter initializes an *acknowledgement outstanding list* (AOL) which contains the identity of all receivers from which an acknowledgement for that message is expected.

The AOL is updated as error free acknowledgements are received. If a message is retransmitted, or m copies of the same message have been transmitted, the transmitter may receive more than one ACK for a message from the same receiver. Upon receipt of the first error free ACK from a receiver, that receiver is removed from the AOL. All subsequent acknowledgements for the same message from the same receiver are discarded. When k (where k is the ACK threshold) receivers are removed from the AOL, the transmitter assumes that the message has been correctly transmitted and starts to transmit the next message. If the timer expires before k receivers are removed from the AOL, the transmitter retransmits the message. In the memoryless schemes, the transmitter reinitializes the AOL whenever it retransmits a message and acknowledgements received after the timer expires are ignored. In the full-memory schemes, however, the AOL is not reinitialized when a message is retransmitted. Rather, the AOL contains the identities of those receivers which have not as yet acknowledged the message. In the single copy transmission schemes, the transmitter transmits or retransmits a single copy of a message. In m-copy transmission schemes, the transmitter transmits (or retransmits) m copies of the same message.

Figure 1 illustrates the memoryless and full-memory schemes with single copy transmission for 3 receivers (R=3) and ACK threshold k=2. When message M1 is transmitted, both memoryless and full-memory schemes initialize the AOL to {1,2,3}, i.e., it contains the identity of all receivers participating in the multicast. When the acknowledgement from receiver 1 (r1) is received, the receiver is removed from the AOL and the new AOL is {2,3}. Since only one acknowledgement, which is less than the ACK threshold $k = 2$, is received within the timeout period (T_out), the transmitter retransmits the same message M1. The memoryless scheme reinitializes the AOL to {1,2,3}, whereas in the full-memory scheme the AOL {2,3} is not reinitialized and it contains the receivers which have not as yet acknowledged. Figure 2 shows these two schemes with m-copy transmission for R=3, k=2 and m=2, where m is the number of copies for the same message. In this figure, both copies of M1 are acknowledged by receiver 1 (r1), but only the first acknowledgement updates the AOL to {2,3}. Notice that if the first acknowledgement is lost in transit, the second acknowledgement may update the AOL. In both figures, 'x' indicates that a message (or an acknowledgement) is in error or it is lost.

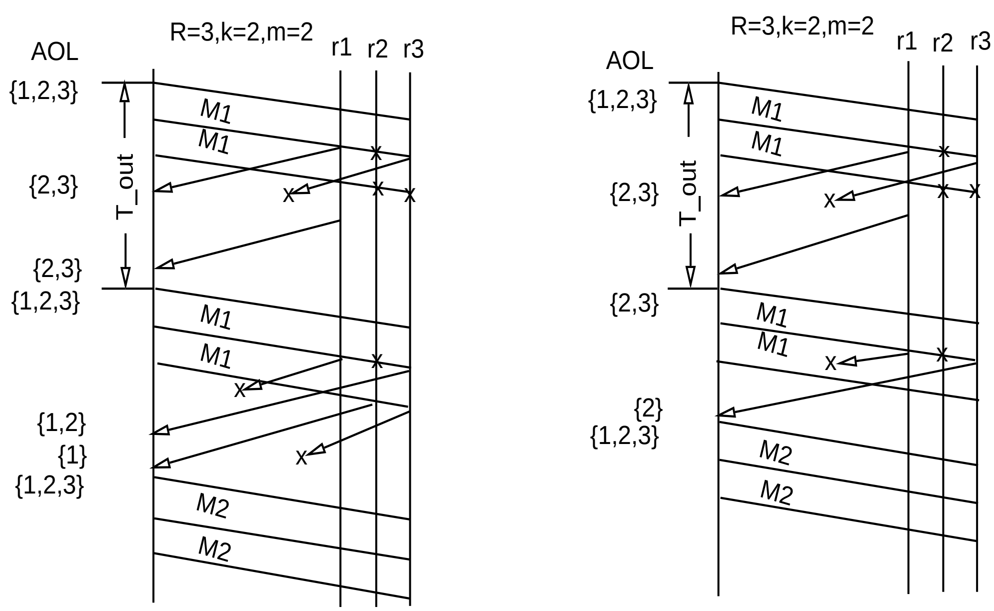
3 Throughput Analysis

The round-trip delay for a particular receiver is defined as the time interval elapsing from the moment that the transmission of a message is started to the moment an acknowledgement from the receiver is received. The round-trip delay, which includes delays such as message transmission time, round-trip propagation delay, processing delays, and queuing delays, is assumed to follow an arbitrary distribution. The round-trip delay for different receivers is independent and identically distributed. We define the throughput η_k as the ratio of the expected message transmission time to the expected time required to successfully transmit a message (i.e., the expected time to receive the k-th acknowledgement). Below, we first develop a basic model which is common to all four schemes. Subsequently, we extend the analysis to model each of the four schemes.



(a) Memoryless Scheme (b) Full-memory Scheme

Figure 1: Single copy transmission schemes



(a) Memoryless Scheme (b) Full-memory Scheme

Figure 2: M-copy transmission schemes

3.1 The Basic Model

We start with the case in which the transmitter always receives all acknowledgements from all receivers. Acknowledgements may be delayed, but they will eventually arrive at the transmitter error free (i.e., the probability of receiving an acknowledgement error free is p=1). Let T_i be the round-trip delay for a message from the i-th receiver. We assume that all T_i $(1 \leq i \leq R)$, are independent and identically distributed with a common distribution function $F(t)$. Let

$$\begin{gathered}
\tau_1 \text{ be the smallest of } T_1,\, T_2,\, \cdots T_R, \\
\tau_2 \text{ be the next larger of } T_1,\, T_2,\, \cdots T_R, \\
\vdots \\
\tau_k \text{ be the k-th largest of } T_1,\, T_2,\, \cdots T_R, \\
\vdots
\end{gathered}$$

τ_k is called the k-th order statistic for the class $\{\ T_i : 1 \leq i \leq R\ \}$. We want to determine the distribution function $F_k(t) = P\{\tau_k \leq t\}$, that is the probability that k or more of the T_i's have a value no greater than t. We may view this problem as a Bernoulli sequence of R trials. There is a success in the i-th trial if and only if $T_i \leq t$. Since $P\{T_i \leq t\} = F(t)$, we have

$$\begin{aligned}
F_k(t) &= P\{\tau_k \leq t\} \\
&= P\{\text{k or more of the } T_i \text{ lie in (0,t]}\} \\
&= \sum_{j=k}^{R} \frac{R!}{j!(R-j)!} F(t)^j [1-F(t)]^{R-j} \qquad (1)
\end{aligned}$$

The probability density function for τ_k can be obtained by taking the first derivative of $F_k(t)$ (see Appendix A). We have

$$f_k(t) = \frac{R!}{(k-1)!(R-k)!} F(t)^{k-1} f(t) [1-F(t)]^{R-k}, \qquad (2)$$

where $f(t)$ is the probability density function of $F(t)$. Now we consider the case in which acknowledgements may not be received. Let p $(0 < p < 1)$ be the probability of receiving an acknowledgement error free. The probability that the round-trip delay is no greater than t can be expressed as

$$P\{T_i \leq t\} = \begin{cases} F(t) & \text{with } p \\ 0 & \text{with } 1-p \end{cases}$$

and the probability that exactly n acknowledgements are received is

$$P\{n \text{ ACKs are received}\} = \frac{R!}{n!(R-n)!} p^n (1-p)^{R-n}.$$

Given that n acknowledgements have been received, $n \geq k$, the round-trip delay distribution for the k-th acknowledgement is

$$F_{k|n}(t) = \sum_{j=k}^{n} \frac{n!}{j!(n-j)!} F(t)^j [1-F(t)]^{n-j}. \qquad (3)$$

The probability of receiving k or more acknowledgements by time t is derived as

$$\begin{aligned}
F_k(t) &= \sum_{n=k}^{R} F_{k|n}(t) \frac{R!}{n!(R-n)!} p^n (1-p)^{R-n} \\
&= \sum_{n=k}^{R} \sum_{j=k}^{n} \frac{n!}{j!(n-j)!} F(t)^j [1-F(t)]^{n-j} \frac{R!}{n!(R-n)!} p^n (1-p)^{R-n} \\
&= \sum_{j=k}^{R} \frac{R!}{j!(R-j)!} [pF(t)]^j [1-pF(t)]^{R-j}. \qquad (4)
\end{aligned}$$

The detailed derivation of equation (4) can be found in Appendix A. The probability density function can be obtained by taking the derivative of $F_k(t)$ (see Appendix B). We have

$$f_k(t) = \frac{R!}{(k-1)!(R-k)!} [pF(t)]^{k-1} p f(t) [1-pF(t)]^{R-k}. \qquad (5)$$

Note that if we set p=1 in equation (5), we obtain equation (2).

3.2 The memoryless scheme with single copy transmission

Since the acknowledgement outstanding list (AOL) is reinitialized upon retransmission of a message, the expected time E_k to receive the k-th acknowledgement within the timeout period T_O is

$$\begin{aligned}
E_k &= E[\tau_k | \tau_k \le T_O] \\
&= \sum_{n=k}^{R} \left[\int_0^{T_O} t f_{k|n}(t|\tau_k \le T_O) dt \right] \frac{R!}{n!(R-n)!} p^n (1-p)^{R-n} \\
&= \frac{k}{F_k(T_O)} \binom{R}{k} \int_0^{T_O} t [pF(t)]^{k-1} p f(t) [1-pF(t)]^{R-k} dt. \qquad (6)
\end{aligned}$$

The detailed derivation of equation (6) can be found in Appendix C. The probability of receiving k or more acknowledgements within a timeout period is $F_k(T_O)$. Let $P_S = F_k(T_O)$ and let γ be the total time required to transmit a message successfully. The successful transmission of a message follows a geometric distribution with parameter P_S. Thus, we have

$$\begin{aligned}
\gamma &= \sum_{i=1}^{\infty} (1-P_S)^{i-1} P_S [(i-1)T_O + E_k] \\
&= \frac{1-P_S}{P_S} T_O + E_k. \qquad (7)
\end{aligned}$$

If the average message transmission time is T_F, then the throughput η_k is

$$\eta_k = \frac{T_F}{\gamma} = \frac{P_S T_F}{P_S E_k + (1-P_S) T_O}. \qquad (8)$$

3.3 The full-memory scheme with single copy transmission

In the full memory scheme, a receiver is removed from the AOL upon the receipt of the first acknowledgement. We derive the distribution function for the first acknowledgement (the minimum round-trip delay distribution) received after a transmission of a message or subsequent retransmissions. As before, we assume that the round-trip delay follows an arbitrary probability density function $f(t) \geq 0$ if $t \geq 0$ and $f(t) = 0$ if $t < 0$. We define the following probability density functions:

$$
\begin{aligned}
g_1(t) &= \begin{cases} pf(t) & t \geq 0 \\ 0 & t < 0 \end{cases} \\
g_2(t) &= \begin{cases} pf(t - T_O) & t \geq T_O \\ 0 & t < T_O \end{cases} \\
&\vdots \\
g_i(t) &= \begin{cases} pf(t - (i-1)T_O) & t \geq (i-1)T_O \\ 0 & t < (i-1)T_O \end{cases} \\
&\vdots
\end{aligned}
$$

where $g_i(t)$ is the probability density function of the round-trip delay for the (i-1)st retransmission shifted by (i-1)T_O and p is the probability of receiving an acknowledgement. The probability density function for the first acknowledgement can be obtained as follows. Before the first timeout period expires, the transmitter expects only one acknowledgement from each receiver. Hence, the probability density function of the time to receive the first acknowledgement from a particular receiver is $g_1(t)$ which is the original round-trip delay distribution function. If an acknowledgement is not received until the timer expires, the transmitter retransmits the same message. In the time interval $T_O \leq t < 2T_O$, two acknowledgements are outstanding and the probability density function for the first acknowledgement is

$$g_1(t) \int_t^\infty g_2(x)dx + g_2(t) \int_t^\infty g_1(x)dx.$$

Putting these together, we have the following probability density function for the first acknowledgement from a particular receiver:

$$
f^*(t) = \begin{cases} g_1(t) & 0 \leq t < T_O \\ g_1(t)\bar{G}_2(t) + g_2(t)\bar{G}_1(t) & T_O \leq t < 2T_O \\ \vdots & \\ \sum_{i=1}^{n} g_i(t) \prod_{j=1, j \neq i}^{n} \bar{G}_j(t) & (n-1)T_O \leq t < nT_O \\ \vdots & \end{cases} \tag{9}
$$

where $\bar{G}_i(t) = \int_t^\infty g_i(x)dx = 1 - \int_0^t g_i(x)dx$. The probability distribution (cumulative) function $F^*(t)$ can be obtained by integrating $f^*(t)$ up to time t. The probability density function $f_k^*(t)$ for the k-th AOL update can be obtained by replacing $f^*(t)$ and $F^*(t)$ with

$f(t)$ and $F(t)$ in equation (2). The expected time to receive the k-th acknowledgement is

$$E_k^* = \int_0^\infty t f_k^*(t) dt \tag{10}$$

and the expression for the throughput η_k^* is

$$\eta_k^* = \frac{T_F}{E_k^*}. \tag{11}$$

3.4 The memoryless scheme with m-copy transmission

In this scheme, m copies of a message are transmitted sequentially within a timeout period. As before, we assume that the round-trip delay follows an arbitrary probability density function $f(t)$, where $f(t) \geq 0$ if $t \geq 0$ and $f(t) = 0$ if $t < 0$. We define the following probability density functions:

$$\begin{aligned}
h_1(t) &= \begin{cases} pf(t) & t \geq 0 \\ 0 & t < 0 \end{cases} \\
h_2(t) &= \begin{cases} pf(t - T_F) & t \geq T_F \\ 0 & t < T_F \end{cases} \\
&\vdots \\
h_m(t) &= \begin{cases} pf(t - (i-1)T_F) & t \geq (i-1)T_F \\ 0 & t < (i-1)T_F \end{cases}
\end{aligned}$$

where $h_i(t)$ is the probability density function of the round-trip delay for the i-th copy transmission of a message and p is the probability of receiving an acknowledgement. Although m copies of a message are transmitted within a timeout period, the transmitter removes receivers in the AOL upon receipt of the first acknowledgements of the message from each receiver. The probability density function for the first acknowledgement from a particular receiver is

$$f^m(t) = \begin{cases} h_1(t) & 0 \leq t < T_F \\ h_1(t)\bar{H}_2(t) + h_2(t)\bar{H}_1(t) & T_F \leq t < 2T_F \\ \vdots & \\ \sum_{i=1}^{m} h_i(t) \prod_{j=1, j\neq i}^{m} \bar{H}_j(t) & (m-1)T_F \leq t \end{cases} \tag{12}$$

where $\bar{H}_i(t) = \int_t^\infty h_i(x)dx = 1 - \int_0^t h_i(x)dx$. The probability distribution (cumulative) function $F^m(t)$ can be obtained by integrating equation (12) up to time t. We can obtain $F_k^m(t)$ and $f_k^m(t)$ by replacing $F^m(t)$ and $f^m(t)$ with $F(t)$ and $f(t)$ in equations (1) and (2), respectively. In turn, we can obtain E_k^m as in equation (6):

$$E_k^m = \frac{k}{F_k^m(T_O)} \binom{R}{k} \int_0^{T_O} t[F^m(t)]^{k-1} f^m(t)[1 - F^m(t)]^{R-k} dt. \tag{13}$$

Letting $F_k^m(T_O) = P_S^m$, the throughput η_k^m of this scheme is

$$\eta_k^m = \frac{P_S^m T_F}{P_S^m E_k^m + (1 - P_S^m)T_O}. \tag{14}$$

3.5 The full-memory scheme with m-copy transmission

This scheme combines the m-copy transmission and the full-memory scheme. Using the probability density function derived in section 3.4, we define the following probability density functions:

$$
\begin{aligned}
q_1(t) &= \begin{cases} f^m(t) & t \geq 0 \\ 0 & t < 0 \end{cases} \\
q_2(t) &= \begin{cases} f^m(t - T_O) & t \geq T_O \\ 0 & t < T_O \end{cases} \\
&\vdots \\
q_i(t) &= \begin{cases} f^m(t - (i-1)T_O) & t \geq (i-1)T_O \\ 0 & t < (i-1)T_O \end{cases} \\
&\vdots
\end{aligned}
$$

where $q_i(t)$ is the probability density function of the round-trip delay for the (i-1)-st retransmission shifted by (i-1)T_O and p is the probability of receiving an acknowledgement. The probability density function for the first acknowledgement is:

$$
f^\star(t) = \begin{cases} q_1(t) & 0 \leq t < T_O \\ q_1(t)\bar{Q}_2(t) + q_2(t)\bar{Q}_1(t) & T_O \leq t < 2T_O \\ \vdots & \\ \displaystyle\sum_{i=1}^{n} q_i(t) \prod_{j=1, j\neq i}^{n} \bar{Q}_j(t) & (n-1)T_O \leq t < nT_O \\ \vdots & \end{cases} \tag{15}
$$

where $\bar{Q}_i(t) = \int_t^\infty q_i(x)dx = 1 - \int_0^t q_i(x)dx$. The probability distribution (cumulative) function $F^\star(t)$ can be obtained by integrating the above probability density function (15) up to time t. As in section 3.3, replacing $f^\star(t)$ and $F^\star(t)$ with $f(t)$ and $F(t)$ in equations (2), we can obtain the probability density function for the k-th acknowledgement $f_k^\star(t)$. The expected time to receive the k-th acknowledgement is

$$
E_k^\star = \int_0^\infty t f_k^\star(t)dt, \tag{16}
$$

and the expression for the throughput $\eta_k^\star$ is

$$
\eta_k^\star = \frac{T_F}{E_k^\star}. \tag{17}
$$

4 Numerical Examples

In this section, we give plots of the throughput of the retransmission schemes for different cases. We assume the average message transmission time T_F=1 and that the total number of receivers R=50. Three arbitrary probability distribution functions, an exponential, a 2-phase Erlang and a 2-phase Coxian, are chosen to represent the round trip delay. Parameters for these three distributions are: μ=0.4 for the 2-stage Erlang, μ_1=0.5, μ_2=0.01 and a=0.03 for the 2-phase Coxian, and μ=0.2 for the exponential distribution. All three distribution have the same mean of 5. The squared coefficient of variation C^2 is 0.5 and 23.8 for the 2-phase Erlang and the 2-phase Coxian respectively. These three distributions were shifted in time in order to account for fixed time components such as round-trip propagation delay, the minimum processing delay, and the minimum time to transmit a message. The fixed (shifted) time s is assumed to be 1.5 for all three distributions. Accordingly, the three shifted distributions have a mean of 6.5 and the squared coefficient of variation $C^2_{Exp} = 0.6$, $C^2_{Erl} = 0.3$ and $C^2_{Cox} = 14$.

Figures 3 to 5 give the throughput of each of the four schemes assuming that the round-trip delay follows the above shifted exponential distribution, and that the probability of receiving an acknowledgement from a particular receiver p=0.99. This corresponds to the bit error rate $BER \simeq 10^{-5}$ for 1,000 bit message and $BER \simeq 10^{-6}$ for 10,000 bit message if acknowledgements are always received error free.

Figure 3 gives the throughput of each of the four schemes in terms of the timeout period T_O. The ACK threshold k is set to 45 (i.e., 90% of the total acknowledgements). For both ML-M and FM-M schemes, three copies of the same message (m=3) are transmitted or retransmitted continuously within a timeout period. When the timeout period is 1 (same as the message transmission time T_F), both FM-S and FM-M schemes achieves the same maximum throughput of 0.1757. FM-M maintains a maximum throughput until the timeout period is $3T_F$. In FM-M, if $T_O \leq mT_F$, the same message is continuously transmitted until the k-th acknowledgement is received. When the timeout period is greater than mT_F, there exists channel idle time between retransmissions, thus the throughput of the full memory schemes (FM-M and FM-S) decreases. The throughput of the memoryless schemes (ML-M and ML-S), however, increases because the probability of a successful transmission (i.e., the probability of receiving k or more acknowledgements) in a timeout period increases. When the timeout period is greater than 9, the m-copy schemes (ML-M and FM-M) have the same throughput, and when the timeout period is greater than 20, the single copy schemes (FM-S and ML-S) have the same throughput. Further increase of the timeout period over 9 and 20 respectively has very little impact on the throughput due to tradeoffs between the probability of successful transmission and the channel idle time. We note that the m-copy schemes always perform better over single copy schemes.

Figure 4 shows the throughput of all four schemes in terms of the ACK threshold k. The timeout period T_O is set to 13.5. This specific timeout period is the 90th percentile of the probability distribution function of the time required to receive an acknowledgement from a particular receiver (i.e., 0.9=p[1-exp(-$\mu(T_O - s)$], where p=0.99, s=1.5 and μ=0.2). As before, both ML-M and FM-M schemes transmit 3 copies (m=3) of the same message at each timeout period. When the ACK threshold is less than 6, all four schemes have the same throughput. Here, the probability of receiving the k-th acknowledgement from the first transmission of a message or the transmission of the first copy of a message is high. Notice that the maximum achievable throughput with this round-trip delay distribution is $1/s$=0.667. As the ACK threshold increases, the m-copy schemes (ML-M and FM-M) have better throughput. As the ACK threshold approaches to R (total number of receivers), the throughput of ML-S is sharply dropped. This is because the probability of receiving the k-th acknowledgement within a timeout period is very low.

Figure 5 shows the throughput of the m-copy schemes in terms of the number of copies m. The timeout period T_O is set to 13.5, the ACK threshold k=45, and the probability of receiving an acknowledgement p=0.99. As the number of copies increases, the throughput is improved. For this example, if more than 4 copies are transmitted, no further performance gain is observed.

In Figure 6, we plot the throughput of ML-S for the three shifted round-trip probability density functions described at the beginning of this section. For each distribution, the timeout period was fixed to the 90th percentile as in the case of Figure 4. The mass in the 2-phase Coxian distribution is mostly concentrated toward zero, it thus has the highest throughput. As the ACK threshold increases, however, the tail probabilities dominate the throughput.

Figures 7 and 8 give the throughput of ML-S and ML-M in terms of the timeout period (T_O) for the three shifted round-trip delay distributions. When the probability of receiving an acknowledgement p is reduced from 0.99 to 0.9, the m-copy scheme still maintains higher throughput than the single copy scheme. Figure 9 shows the throughput of the ML-M scheme in terms of the number of copies (m) for the three shifted round-trip delay distributions.

5 Conclusions

We obtained analytic expressions for the throughput of several stop-and-wait retransmission schemes for k-reliable multicast. The analytical expressions were derived in terms of the timeout period, the acknowledgement threshold and the probability of receiving an acknowledgement.

The full-memory schemes perform better over memoryless schemes for both high ACK threshold and short timeout period. The m-copy retransmission schemes give higher throughput performance. The number of copies per transmission that optimizes throughput, may vary depending on the round-trip delay distribution.

As the ACK threshold increases, the throughput efficiency decreases. The correct balance between efficiency and reliability must be evaluated for each application to be supported by multicast services to determine the specific variety of multicasting to be used. The proposed k-reliable schemes can be used for applications which need high throughput with an acceptable level of reliability.

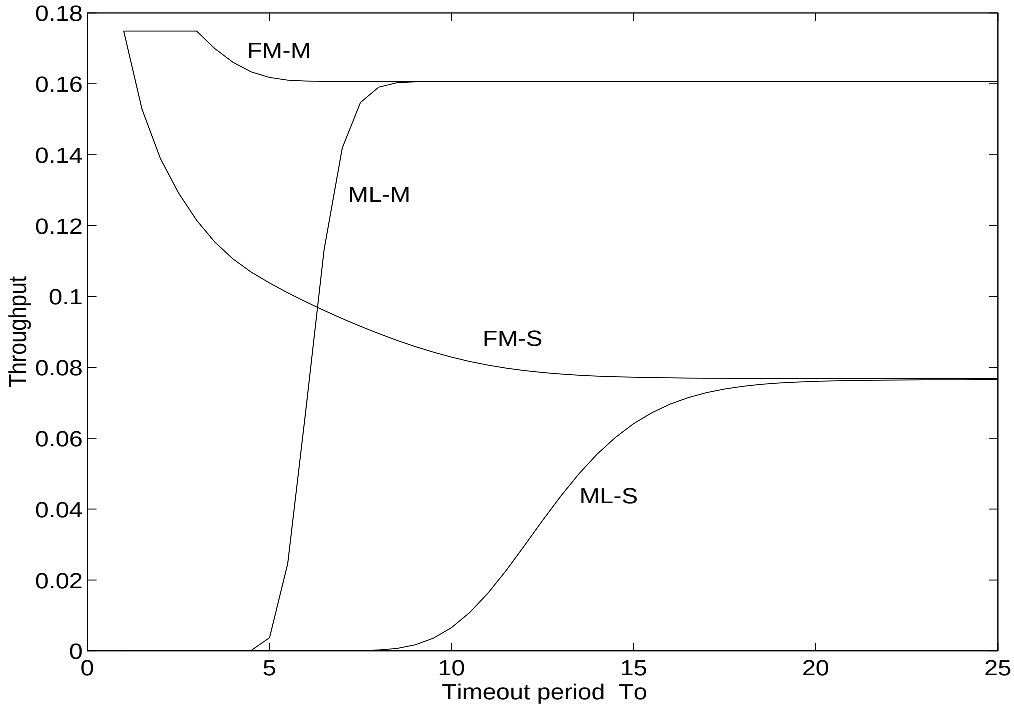


Figure 3: Throughput vs timeout period (T_O). R=50, k=45, p=0.99, m=3. Round-trip delay : shifted exponential distribution with μ=0.2, s=1.5.

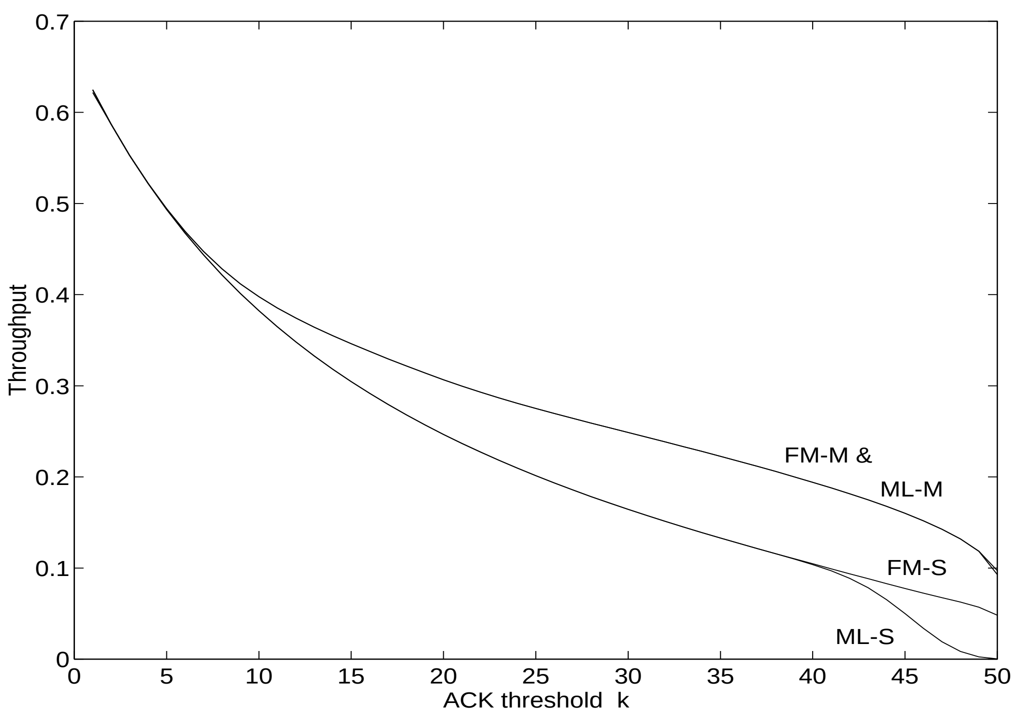


Figure 4: Throughput vs ACK threshold (k). R=50, p=0.99, T_O=13.5, m=3. Round-trip delay : shifted exponential distribution with μ=0.2, s=1.5.

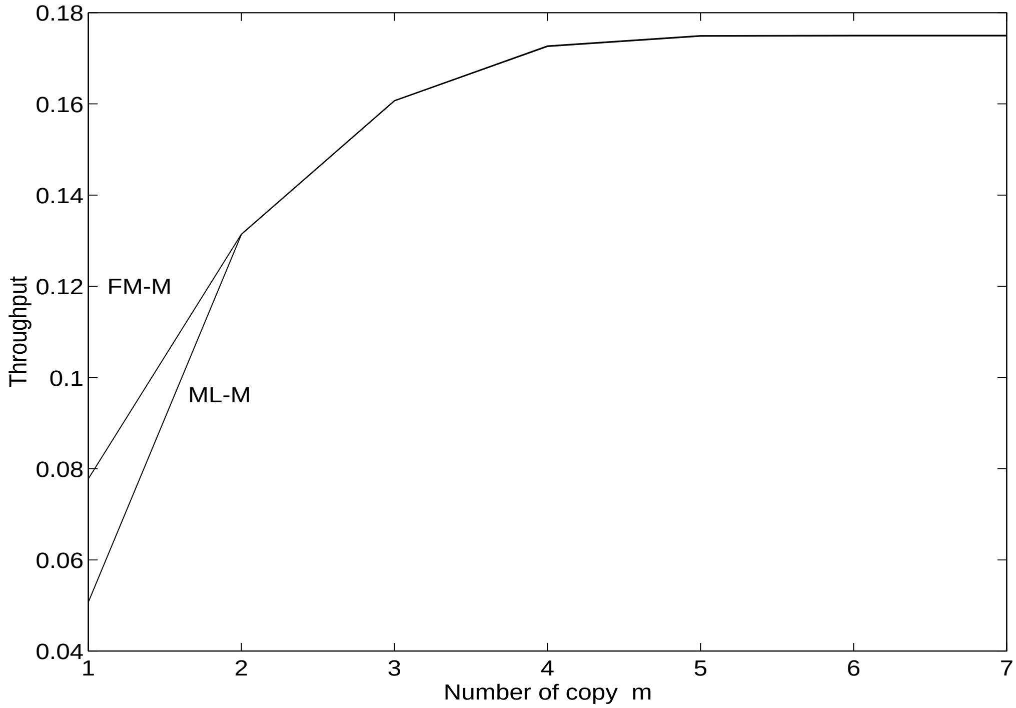


Figure 5: Throughput vs number of copies (m). R=50, k=45, p=0.99, T_O=13.5. Round-trip delay : shifted exponential distribution with μ=0.2, s=1.5.

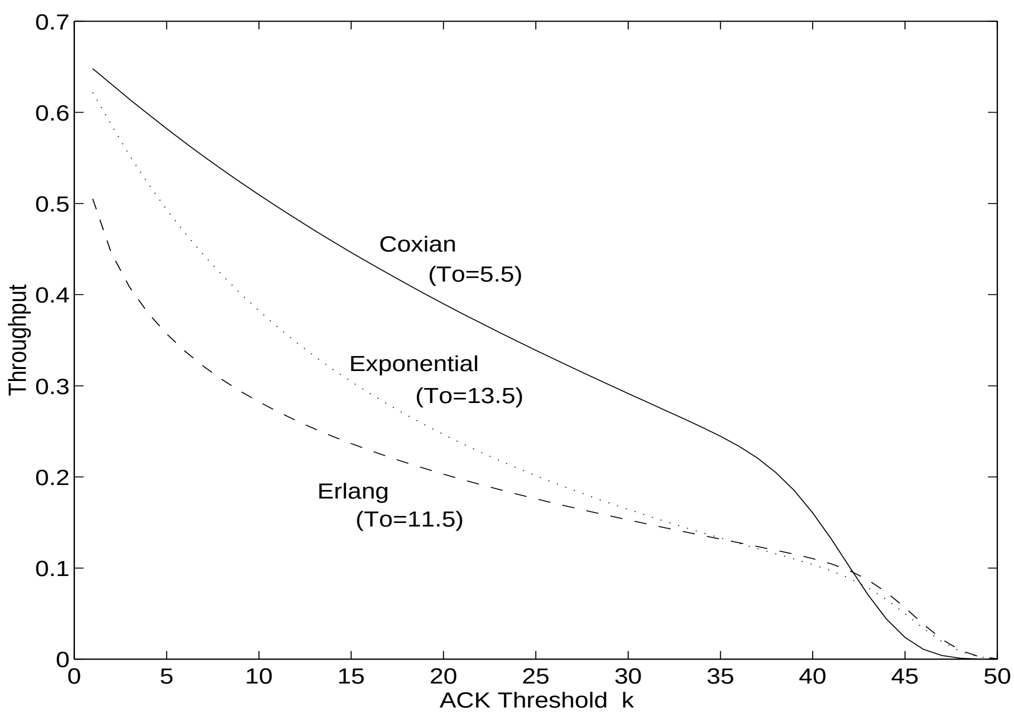


Figure 6: Throughput vs ACK threshold (k) for different round-trip delay distributions. R=50, p=0.99, s=1.5.

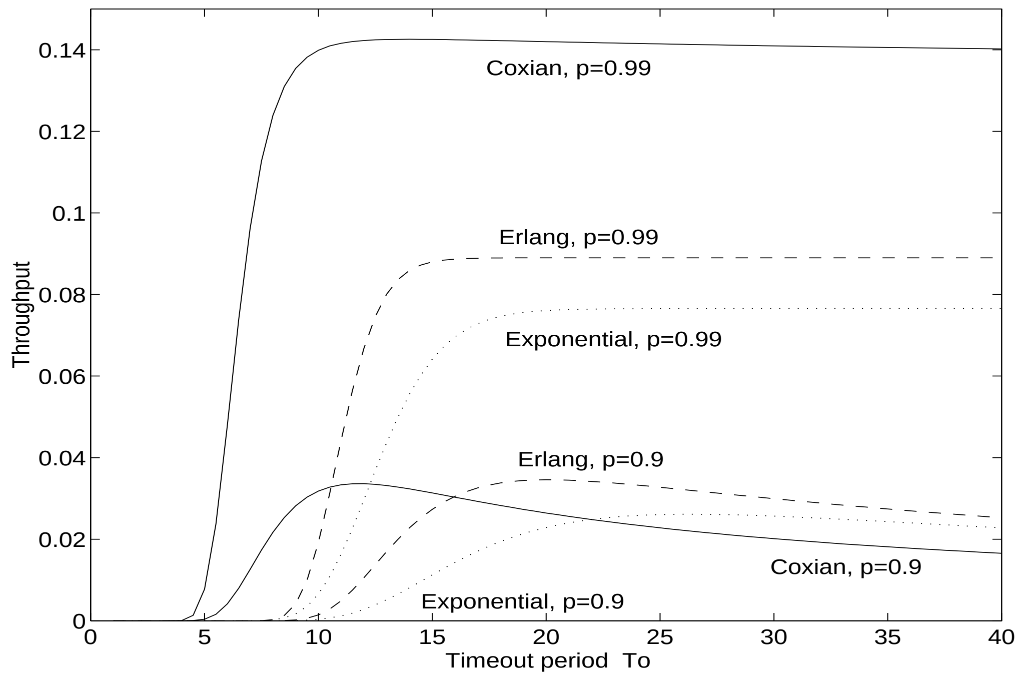


Figure 7: Throughput of ML-S vs timeout period (T_O) for different round-trip delay distributions. R=50, k=45, s=1.5.

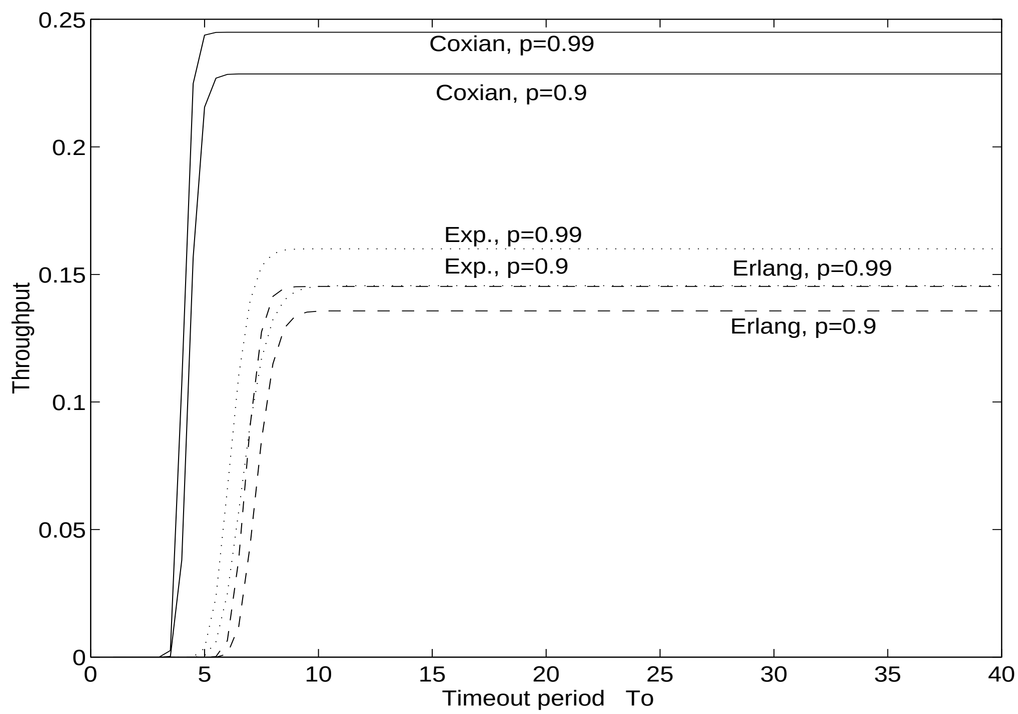


Figure 8: Throughput of ML-M vs timeout period (T_O) for different round-trip delay distributions. R=50, k=45, m=3, s=1.5.

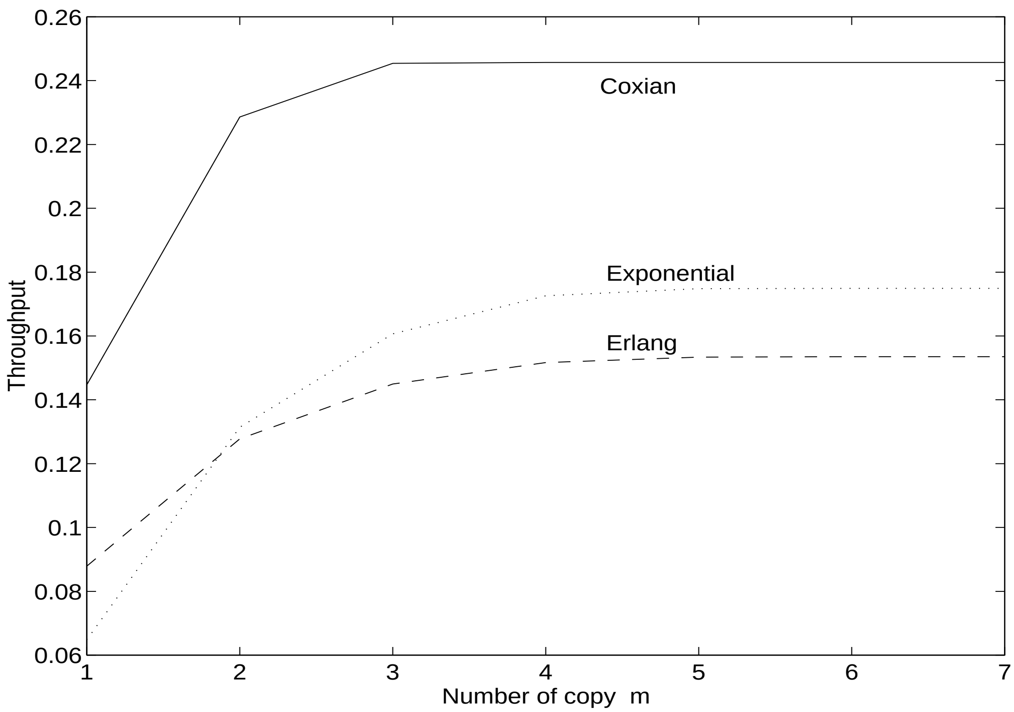


Figure 9: Throughput of ML-M vs the number of copies (m) for different round-trip delay distributions. R=50, k=45, p=0.99, T_O=15.

Appendix A

The probability density function for τ_k can be obtained by taking the derivative of $F_k(t)$:

$$
\begin{aligned}
f_k(t) &= \frac{d}{dt}\sum_{j=k}^{R}\binom{R}{j}F(t)^j[1-F(t)]^{R-j} \\
&= \sum_{j=k}^{R}\frac{R!}{(R-j)!(j-1)!}F(t)^{(j-1)}f(t)[1-F(t)]^{(R-j)} \\
&+ \sum_{j=k}^{R}\frac{R!}{(R-j-1)!j!}F(t)^{j}f(t)[1-F(t)]^{(R-j-1)} \\
&= \frac{R!}{(R-k)!(k-1)!}F(t)^{(k-1)}f(t)[1-F(t)]^{(R-k)} \\
&- \frac{R!}{(R-k-1)!k!}F(t)^{k}f(t)[1-F(t)]^{(R-k-1)} \\
&+ \frac{R!}{(R-k-1)!k!}F(t)^{k}f(t)[1-F(t)]^{(R-k-1)} \\
&- \frac{R!}{(R-k-2)!(k+1)!}F(t)^{(k+1)}f(t)[1-F(t)]^{(R-k-2)} \\
&+ \frac{R!}{(R-k-2)!(k+1)!}F(t)^{(k+1)}f(t)[1-F(t)]^{(R-k-2)}
\end{aligned}
$$

$$
\begin{aligned}
&- \quad \frac{R!}{(R-k-3)!(k+2)!} F(t)^{(k+2)} f(t)[1-F(t)]^{(R-k-3)} \\
&\quad \vdots \\
&+ \quad \frac{R!}{(R-1)!} F(t)^{(R-1)} f(t)[1-F(t)]^{0} \\
&= \quad \frac{R!}{(R-k)!(k-1)!} F(t)^{k-1} f(t)[1-F(t)]^{R-k}. \qquad (18)
\end{aligned}
$$

The probability density function can also be obtained by applying the multinomial law. For a given distribution, R observations are drawn at random from the distribution and then ordered by size so that τ_k represents the k-th largest value. Then, the probability that τ_k will lie between t and $t+dt$ can be obtained by considering the general multinomial law. We want to obtain the probability that, when picking R variates at random from the given distribution, the variates fall $(k-1)$ times in the interval $(0,t)$, once in $(t, t+dt)$ and $(R-k)$ times in $(t+dt, \infty)$. Substituting the corresponding probabilities in the multinomial, we have

$$
\begin{aligned}
f_k(t)dt &= \frac{R!}{(k-1)!1!(R-k)!} [\int_0^t f(t)dt]^{k-1} f(t)dt [\int_t^\infty f(t)dt]^{R-k} \\
&= \frac{R!}{(k-1)!(R-k)!} F(t)^{k-1} f(t)dt [1-F(t)]^{R-k}.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
f_k(t) &= \frac{R!}{(k-1)!(R-k)!} F(t)^{k-1} f(t)[1-F(t)]^{R-k} \\
&= k \binom{R}{k} F(t)^{k-1} f(t)[1-F(t)]^{R-k}. \qquad (19)
\end{aligned}
$$

The probability distribution function $F_k(t)$ can be obtained by integrating the probability density function $f_k(t)$:

$$
\begin{aligned}
F_k(t) &= \int_0^t \frac{R!}{(k-1)!(R-k)!} F(\tau)^{k-1} f(\tau)(1-F(\tau))^{R-k} d\tau \\
&= \left[\frac{R!}{k!(R-k)!} F(\tau)^{k}(1-F(\tau))^{R-k} \right]_0^t \\
&+ \int_0^t \frac{R!}{(R-k-1)!} F(\tau)^{k} f(\tau)(1-F(\tau))^{R-k-1} d\tau \\
&= \left[\frac{R!}{k!(R-k)!} F(\tau)^{k}(1-F(\tau))^{R-k} \right]_0^t \\
&+ \left[\frac{R!}{(k+1)!(R-k-1)!} F(\tau)^{k+1}(1-F(\tau))^{R-k-1} \right]_0^t \\
&+ \int_0^t \frac{R!}{(k+2)!(R-k-2)!} F(\tau)^{k+1} f(\tau)(1-F(\tau))^{R-k-2} d\tau
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{R!}{k!(R-k)!}F(t)^k(1-F(t))^{R-k} \\
&+ \frac{R!}{(k+1)!(R-k-1)!}F(t)^{k+1}(1-F(t))^{R-k-1} \\
&+ \frac{R!}{(k+2)!(R-k-2)!}F(t)^{k+2}(1-F(t))^{R-k-2} \\
&\quad \vdots \\
&+ \frac{R!}{R!0!}F(t)^R(1-F(t))^0 \\
&= \sum_{j=k}^{R} \frac{R!}{(R-j)!j!}F(t)^j(1-F(t))^{R-j}. \qquad (20)
\end{aligned}
$$

Appendix B

The probability of receiving at least k acknowledgements for a message is derived as

$$
\begin{aligned}
F_k(t) &= \sum_{n=k}^{R} F_{k|n}(t)\frac{R!}{n!(R-n)!}p^n(1-p)^{R-n} \\
&= \sum_{n=k}^{R}\sum_{j=k}^{n} \frac{n!}{j!(n-j)!}F(t)^j[1-F(t)]^{n-j}\frac{R!}{n!(R-n)!}p^n(1-p)^{R-n} \\
&= \sum_{j=k}^{R}\sum_{n=j}^{R} \frac{n!}{j!(n-j)!}\frac{R!}{n!(R-n)!}F(t)^j[1-F(t)]^{n-j}p^n(1-p)^{R-n} \\
&= \sum_{j=k}^{R} \frac{R!}{j!(R-j)!}[pF(t)]^j \sum_{n=j}^{R} \frac{(R-j)!}{(n-j)!(R-n)!}[p-pF(t)]^{n-j}(1-p)^{R-j-(n-j)} \\
&= \sum_{j=k}^{R} \frac{R!}{j!(R-j)!}[pF(t)]^j \sum_{l=0}^{R-j} \frac{(R-j)!}{l!(R-j-l)!}[p-pF(t)]^{l}(1-p)^{R-j-l} \\
&= \sum_{j=k}^{R} \frac{R!}{j!(R-j)!}[pF(t)]^j[p-pF(t)+1-p]^{R-j} \\
&= \sum_{j=k}^{R} \frac{R!}{j!(R-j)!}[pF(t)]^j[1-pF(t)]^{R-j}. \qquad (21)
\end{aligned}
$$

The probability density function for the k-th acknowledgement can be obtained by taking the derivative of equation (21):

$$
\begin{aligned}
f_k(t) &= \frac{d}{dt}\sum_{j=k}^{R}\binom{R}{j}[pF(t)]^j[1-pF(t)]^{R-j} \\
&= \sum_{j=k}^{R} \frac{R!}{(R-j)!(j-1)!}[pF(t)]^{(j-1)}pf(t)[1-pF(t)]^{(R-j)} \\
&+ \sum_{j=k}^{R} \frac{R!}{(R-j-1)!j!}[pF(t)]^{j}pf(t)[1-pF(t)]^{(R-j-1)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{R!}{(R-k)!(k-1)!}[pF(t)]^{(k-1)}pf(t)[1-pF(t)]^{(R-k)} \\
&- \frac{R!}{(R-k-1)!k!}[pF(t)]^{k}pf(t)[1-pF(t)]^{(R-k-1)} \\
&+ \frac{R!}{(R-k-1)!k!}[pF(t)]^{k}pf(t)[1-pF(t)]^{(R-k-1)} \\
&- \frac{R!}{(R-k-2)!(k+1)!}[pF(t)]^{(k+1)}pf(t)[1-pF(t)]^{(R-k-2)} \\
&+ \frac{R!}{(R-k-2)!(k+1)!}[pF(t)]^{(k+1)}pf(t)[1-pF(t)]^{(R-k-2)} \\
&- \frac{R!}{(R-k-3)!(k+2)!}[pF(t)]^{(k+2)}pf(t)[1-pF(t)]^{(R-k-3)} \\
&\quad \vdots \\
&+ \frac{R!}{(R-1)!}[pF(t)]^{(R-1)}pf(t)[1-pF(t)]^{0} \\
&= \frac{R!}{(R-k)!(k-1)!}[pF(t)]^{k-1}pf(t)[1-pF(t)]^{R-k}. \qquad (22)
\end{aligned}
$$

Appendix C

The expected time to receive the k-th ACK is obtained as follows:

$$
\begin{aligned}
E_k &= E[\tau_k|\tau_k \leq T_O] \\
&= \sum_{n=k}^{R}\left[\int_0^{T_O} t f_{k|n}(t|\tau_k \leq T_O)dt\right]\frac{R!}{n!(R-n)!}p^n(1-p)^{R-n} \\
&= \sum_{n=k}^{R}\left[\frac{1}{F_k(T_O)}\int_0^{T_O} t\frac{n!}{(k-1)!(n-k)!}F(t)^{k-1}f(t)[1-F(t)]^{n-k}dt\right]\times \\
&\qquad \left[\frac{R!}{n!(R-n)!}p^n(1-p)^{R-n}\right] \\
&= \left[\frac{1}{F_k(T_O)}\frac{R!}{(R-k)!(k-1)!}\int_0^{T_O} tF(t)^{k-1}f(t)p^k\right]\times \\
&\qquad \left[\sum_{n=k}^{R}\frac{(R-k)!}{(n-k)!(R-n)!}[p-pF(t)]^{n-k}(1-p)^{R-k-(n-k)}dt\right] \\
&= \left[\frac{1}{F_k(T_O)}\frac{R!}{(R-k)!(k-1)!}\int_0^{T_O} tF(t)^{k-1}f(t)p^k\right]\times \\
&\qquad \left[\sum_{l=0}^{R-k}\frac{(R-k)!}{l!(R-k-l)!}[p-pF(t)]^{l}(1-p)^{R-k-l}dt\right] \\
&= \frac{k}{F_k(T_O)}\binom{R}{k}\int_0^{T_O} t[pF(t)]^{k-1}pf(t)[1-pF(t)]^{R-k}dt. \qquad (23)
\end{aligned}
$$

References

[1] M. H. Ammar and L. R. Wu, "Improving Throughput of Point-to-Multipoint ARQ Protocols Through Destination Set Splitting," Proceedings of INFOCOM '92, pp.262-271, IEEE Computer Society, 1992.

[2] S.B. Calo and M.C. Easton, "A Broadcast Protocol for File Transfers to Multiple Sites," IEEE Trans. Commun., Vol. COM-29, pp.670-678, Nov. 1981

[3] I.S. Gopal and J.M. Jaffe, "Point-to-multipoint Communication over Broadcast Links," IEEE Trans. Commun., Vol. COM-32, pp.480-486, September 1984.

[4] K. Mase, T. Takenaka, H. Yamamoto, and M. Shinohara, "Go-back-N ARQ Schemes for Point-to-multipoint Satellite Communications," IEEE Trans. Commun., Vol. COM-31, pp.583-589, April 1983.

[5] J.L. Wang and J.A. Silvester, "Optimal Adaptive Multireceiver ARQ protocols," IEEE Trans. on Commun., Vol. COM-41, pp. 1816-1829, 1993.

[6] J.L. Wang and J.A. Silvester, "Delay Minimization of the Adaptive Go-Back-N ARQ protocols for Point-to-multipoint Communication," Proceedings of INFOCOM '89, pp.584-593, IEEE Computer Society, 1989.

[7] G. P. Wadsworth and J.G. Bryan, "Applications of Probability and Random Variables," 2nd edition, pp.169-174, 1974, McGraw-Hill.