Message Path Delays in Packet-Switching Communication Networks

IZHAK RUBIN

Abstract—A communication path (in isolation) in a packet-switching store-and-forward communication network, such as a computer- or satellite-communication network, is considered. Messages are assumed to arrive according to a Poisson stream, and message-lengths are considered to be random variables governed by an arbitrary distribution. Message lengths are divided into fixed-length packets which are sent independently over the N-channel communication path in a store-and-forward manner, and are reassembled at the destination terminal. Expressions for the distributions of the message waiting and delay times over the path are derived. Also, we obtain the limiting average message waiting times and required buffer sizes at the individual channels. The overall message waiting time is observed to depend only on the minimal channel capacity. The case of exponentially distributed message lengths serves as an illustrating example.

I. INTRODUCTION

A COMMUNICATION network is represented as a weighted graph. The branches of the graph represent the communication channels, while the vertices represent (source, repeater, or destination type) stations with storage (queueing) facilities. The branches are assigned capacity weights. Messages arrive at random at a source station and follow a specific route in the network towards their destination station. Message lengths are usually considered to be random variables. In a packet-switching communication network, the message is divided at the source station into fixed-length submessages, called packets. These packets are then sent independently through the network towards the destination station. At the latter stations, all the packets associated with a specific message are reassembled. The whole message is then transferred to the destination terminal (being the destination computer in case of a computer-communication network, or a specific terminal in a satellite-communication network). The network is also assumed to operate in a store-and-forward manner, so that at each station a queue of messages is generated and served according to a first-come-first-served discipline.

In this paper, we are considering an arbitrary N-channel path in a packet-switching store-and-forward communication network. Assuming Poisson arrivals at the source station of this path and random message lengths governed by an arbitrary distribution, we derive exact expressions for the distribution of the message delay over the communication path. We also obtain steady-state expressions for the average message waiting times and required buffer sizes at the individual channels.

Exact expressions for waiting and delay time distributions, as well as busy-period characteristics, along a communication path in a store-and-forward communication network, have recently been derived in [1] and [2]. These papers assume fixed message lengths. The results obtained in this paper thus follow from those obtained in [1] and [2]. Approximate time-delay results for store-and-forward communication networks, utilizing an “independence assumption” (which requires rechoosing the message length, at random, at any station) and assuming exponentially distributed message lengths, are reported in [3] and [4] and the references therein. Many of the recent time-delay reports have been associated with the Advanced Research Projects Agency (ARPA) computer-communication network (see [4]). The latter employs packets of 1000-bit length (and some others of shorter length, for situations in which the packet is not filled).

We note that our delay analysis involves a single path in isolation, so that the messages in this path do not interfere with and are not interfered with by any other messages in the network. Messages arriving at the path are spaced out in contiguous packets by the first channel. If the next channel has a higher capacity, the timing of packets is preserved but the duration of each is shortened. If the next channel has a lower capacity, the packets will be extended and if messages are close enough together, further delays will ensue. We show that the distribution of the overall message waiting time in the path is equal to that obtained by presenting all the messages to the channel with lowest capacity only.

Preliminary notions and definitions are introduced in Section II. Message time delays over a single-channel path are obtained in Section III. In Section IV, message waiting and delay times over an N-channel communication path are derived. As an illustrating example, we consider in Section V the case of exponentially distributed message lengths, and plot the resulting time delay versus the traffic intensity for several values of average packet to message lengths. The special case of packets of vanishing length, and arbitrary message-length distribution, is then observed.
II. PRELIMINARIES

A communication path, as shown in Fig. 1, is considered. The ith channel, whose capacity is $C_i$ [bits/s], is represented by the branch $(v_i,v_{i+1})$ between vertices $v_i$ and $v_{i+1}$, $i = 1,2,\ldots,N$. The vertices represent stations or buffer systems equipped with (infinite space) memory storage units or queueing facilities.

Messages arrive at input station $v_i$ at random times, following a Poisson stream with arrival intensity $\lambda$ [messages/s]. The message length is assumed to be random.

We let $Y_n$ [bits] denote the length of the nth arriving message at $v_i$, and assume $\{Y_n,n \geq 1\}$ to be an independent identically distributed (iid) sequence of (nonnegative valued) random variables, governed by the distribution function $F_Y(y)$, where

$$F_Y(y) = P\{Y_n \leq y\}. \quad (1)$$

The average message length is $\mu^{-1}$ [bits/message],

$$\mu^{-1} = E(Y_n) = \int_0^\infty ydF_Y(y). \quad (2)$$

At $v_i$, before being processed through the communication path, the message is divided into an integral number of packets of fixed length. We set the packet length to be $\alpha$ [bits/packet], and denote the number of packets representing the nth message by $M_n$. Thus, $\{M_n,n \geq 1\}$ is a sequence of iid random variables following the probability measure $g(m)$, where

$$g(m) \triangleq P\{M_n = m\} = F_Y(am) - F_Y(\alpha[m - 1]), \quad m \geq 1. \quad (3)$$

Clearly, $\sum_{m=1}^\infty g(m) = 1$ and the average number of packets per message is denoted as

$$\bar{M} = \sum_{m=1}^\infty mg(m) \quad \text{[packets/message]}. \quad (4)$$

For example, if the message length is exponentially distributed with mean $\mu^{-1}$,

$$F_Y(y) = (1 - \exp[-\mu y])u(y), \quad (5a)$$

where $u(y)$ denotes the unit step function; then (3) yields the probability $g_{\text{ex}}(m)$, where

$$g_{\text{ex}}(m) = \frac{\alpha^m}{\bar{M}}m!, \quad m \geq 1. \quad (5b)$$

and

$$q = \exp[-\mu \alpha], \quad p = (1-q), \quad (5c)$$

so that the number of packets per message follows a geometric distribution. Clearly, if the message length is not larger than a packet length, we have $g(1) = 1$ and $M_n = 1$ with probability 1 (and the situation studied in [1] subsequently follows).

The packets derived at $v_i$ from each message are then transferred through the communication path in a store-and-forward (packet-switching) manner. Thus, a packet arriving at $v_i$, $i = 1,2,\ldots,N$, is immediately transmitted through the ith channel if it is free, while, if the channel is busy, the packet joins the queue at $v_i$, and is served on a first-come first-served basis. While packets are processed independently at any station $v_i$, $i = 1,\ldots,N$, at the destination terminal $v_{N+1}$ all the packets belonging to one message are assembled and subsequently leave the network (being now transferred to the destination computer).

Each channel with its storage facilities can be considered as a queueing system. For that purpose, we consider a packet which is to be transmitted over the ith channel to be a customer which requires service from the ith server.

A message is thus considered to be a group of customers. Customers thus arrive at $v_i$ in groups, where the arrival process is Poisson with rate $\lambda$ and the group size is governed by distribution $g(m)$. The service time required by a customer (packet) from the server (channel) $i$ is equal to his transmission time over the channel and is given by

$$a_i = \alpha/C_i \quad \text{[s/packet]}. \quad (6)$$

The following notations will be utilized throughout the paper. We let $X_i^{(o)}$ denote the number of packets stored at $v_i$ or being transmitted through channel $i$ at time $t$. Thus, $\{X_i^{(o)}, t \geq 0\}$ is the queuing process associated with channel $i$. We assume $X_i^{(o)} = 0, i = 1,\ldots,N$. The (random) instants of arrival of packets at $v_i$ are denoted by $\{t_{n,i}^{(o)}, n = 1,2,\ldots,n_i = 1,2,\ldots,M_i\}$, where $t_{n,i}^{(o)}$ is the instant of arrival at $v_i$ of the $n$th packet associated with the nth message. The instants of message arrivals at $v_i$ are set to be $\{t_n^{(o)}, n = 1,2,\ldots\}$, where $t_n^{(o)} \triangleq t_{n,1}^{(o)}$. Similarly, we denote the instants of packet and message departure from channel $i$ by $\{r_{n,i}^{(o)}\}$ and $\{p_{n,i}^{(o)}\}$, respectively, where $r_{n,i}^{(o)} \triangleq r_{n,1}^{(o)}$, and $r_{n,i}^{(o)}$ is the departure time from channel $i$ of the jth packet associated with the nth message. Clearly, since we have batch arrivals at $v_i$, we have $t_{n,j}^{(o)} = t_{n,1}^{(o)} = t_n^{(o)}$, for each $j,n$. The packets associated with a specific message are then ordered according to their order of service. Thus, the $k$th packet of the nth message arriving at $v_i$ is the $k$th one to be transmitted over channel 1, among the message $M_i$ packets, and it departs into $v_i$ at time $r_{n,i}^{(o)}$. Clearly, for a communication path, $t_{n,j}^{(o)} = r_{n,j}^{(o)} = t_{n,j}^{(o)} = t_{n}^{(o)}, i = 1,2,\ldots,N - 1$.

The waiting time at $v_i$ of the jth packet associated with the nth arriving message is denoted as $W_{n,j}^{(o)}$. The $n$th
message waiting time at \( v_i \), \( \bar{W}_{n(i)} \) is set as
\[
\bar{W}_{n(i)} = \sum_{i=1}^{n} W_{n(i)},
\]
and is thus equal to the waiting time of the first associated packet. The packet delay time over the \( i \)th channel \( \gamma_{n,i} \) is given by the total of its waiting time and transmission delay time at channel \( i \). Thus, we have
\[
\gamma_{n,i} = W_{n,i} + \alpha_i.
\]
For the \( n \)th message, we define its delay time over the \( i \)th channel, \( i = 1,2,\cdots,N \), by
\[
\tilde{\gamma}_n = r_n - t_{n,i}(i),
\]
as the time difference between the departure of the last associated packet and the arrival of the message at \( v_i \). Similarly, the overall time-delay of the \( n \)th message through the \( N \)-channel communication path is given by
\[
\tilde{\gamma}_n = r_n - t_{n,i}(i),
\]
as the time difference between the instant the last associated packet leaves the \( N \)th channel and the instant of arrival of the \( n \)th message at \( v \).

In this paper, we will derive the steady-state distributions for the overall message time delays and the average message waiting times and delays along the individual channels.

**III. THE SINGLE-CHANNEL PATH**

Consider the case \( N = 1 \), so that the path includes only a single channel. In this case, if we are interested in the packets' waiting-times, we are considering a \( M/D/1 \) queueing system with group arrivals (see [5]-[7]). For this system, customers (packets) arrive according to a Poisson stream with intensity \( \lambda \), and each requires a fixed service of length \( a_i \). Waiting time results for the individual packets readily follow. However, we are interested here in the message time-delays. For that purpose, we need to obtain the distribution of the message waiting time \( W_n = W_{n,1} \). The latter random variable satisfies the relationship
\[
W_{n+1} = [W_n + M_1a_1 - T_{n+1}^{(i)}]_n,
\]
where \( [X]^+ = \max (0,X) \), and \( T_{n+1}^{(i)} = l_{n+1}^{(i)} - l_n^{(i)} \) denotes the \( n+1 \)st message interarrival time at channel \( i \). Relation (11) indicates that for deriving the message waiting times, we need consider a \( M/G/1 \) queueing system with unit Poisson arrivals of intensity \( \lambda \) and service times equal to \( M_1a_1 \).

**Theorem 1:** For the single-channel path, the message waiting-time sequence \( \{W_n, n \geq 1\} \) is governed by the same statistics as the corresponding sequence for a \( M/G/1 \) queueing system with Poisson arrivals of intensity \( \lambda \) and iid service times \( \{X_n, n \geq 1\} \), distributed according to
\[
P\{X_n = ma_1\} = g(m), \quad m = 1,2,3,\cdots.
\]
Hence, the limiting waiting-time distribution, when \( \rho_1 = \lambda E(M_1a_1) = \lambda a_1 < 1 \), where \( \bar{a}_1 = \sum M_1a_1 \), is given by
\[
W^{(i)}(t) = \lim_{n \to \infty} P(W^{(i)}(n) \leq t) = (1 - \rho_1) \sum_{n=0}^{\infty} \rho_n^{[a_1]} \int_0^t [1 - B(x)] \, dx, \quad t > 0,
\]
where
\[
B(t) = P\{X_n \leq t\} = \sum_{m=1}^{[t/a_1]} g(m),
\]
\( [x] \) is the largest integer not larger than \( x \), and \( [F(t)]^{a_1} \) denotes the \( n \)th convolution of \( F(t) \). The limiting message mean waiting time is given by
\[
\tilde{\gamma}_n = \tilde{\gamma}_n^{(i)} = W_n^{(i)} + M_1a_1,
\]
where
\[
C_{M^2} = \frac{\text{var}(M_1)}{[E(M_1)]^2}
\]
is the coefficient of variation associated with random variable \( M_1 \) and distribution \( g(m) \). For \( \rho_1 \geq 1 \), we have \( W^{(i)}(t) = 0 \), for each \( t > 0 \).

**Proof:** The theorem follows from (11) since for a \( M/G/1 \) system we have \( W_n = [W_n + X_n - T_{n+1}]_n \), where the service time \( X_n \) and the interarrival time \( T_{n+1} \) are independent random variables, the latter being exponentially distributed. Hence, (13) and (14) follow (see [5, pp. 255-256]). Equation (14) is known as the Pol- laezek-Khintchine formula.

The message delay time \( \tilde{\gamma}_n \) defined by (9) is now equal to the overall message delay time \( \tilde{\gamma}_n \). The latter are clearly given by
\[
\tilde{\gamma}_n = \tilde{\gamma}_n^{(i)} = W_n^{(i)} + M_1a_1.
\]
Hence, we conclude the following results, observing \( W_n^{(i)} \) and \( M_1a_1 \) to be statistically independent.

**Corollary 1:** For \( \rho_1 < 1 \), and a single channel path, the steady-state distribution of the message delay time is given by
\[
\gamma^{(i)}(t) = \lim_{n \to \infty} P(|\gamma_n| \leq t) = \sum_{n=1}^{\infty} W^{(i)}(t - ma_1)g(m),
\]
where \( W^{(i)}(t) \) is given by (13). The limiting mean delay time is equal to
\[
\mu^{(i)} = \lim_{n \to \infty} E(|\gamma_n|) = \frac{1}{2(1 - \rho_1)} (1 + C_{M^2}) + \bar{a}_1.
\]
Let \( X^{(i)} \) denote the number of messages at \( v_i \) at time \( t \), where a message is counted as long as any of its packets is in the system (waiting or being transmitted). Then, by Little's theorem we have for the limiting average message queue size \( X^{(i)} \),
\[
X^{(i)} = \lim_{t \to \infty} E(X^{(i)}(t)) = \lambda \mu^{(i)}.
\]
Consequently, the average message storage capacity required at \( v_i \), \( \bar{M}^{(i)} \) equals

\[
\bar{M}^{(i)} \triangleq \beta \lim_{t \to \infty} E(X_i^{(i)}) = \lambda \beta \gamma^{(i)} \text{ bits,} \quad (20)
\]

where \( \gamma^{(i)} \) is given by (18), and \( \beta \triangleq \bar{M} \alpha \).

We note that (14) reduces to the average waiting time considered in [1] when a fixed one packet message is considered (and \( C_M^2 = 0 \)). The related expression for an exponentially distributed message length results when \( C_M^2 = 1 \). The average waiting time increases linearly with \( C_M^2 \). When considering exponentially distributed message lengths as in (5a), \( g(m) \) is given by (5b), and subsequently \( C_M^2 \) is obtained to be

\[
C_M^2 = q = \exp (-\mu \alpha). \quad (21)
\]

Here, also

\[
\bar{a}_i = E(M_a) a_i = p^{-1} a_i = \left[ 1 - \exp (-\mu \alpha) \right]^{-1} a_i. \quad (22)
\]

Hence, the limiting average delay time \( \gamma^{(i)} \) of (18) is given for exponentially distributed message lengths by

\[
\gamma^{(i)} = \bar{a}_i + \frac{1}{2} \frac{\rho \bar{a}_i}{2(1-\rho_i)} \left[ 1 + \exp (-\mu \alpha) \right], \quad (23)
\]

where \( \rho_i = \lambda \bar{a}_i \) and \( \bar{a}_i \) is given by (22). Note that \( \gamma^{(i)} \) in (23) is a function of \( a_i = \alpha / C_i \) and of \( \mu \alpha \), and that \( (\mu \alpha)^{-1} \) indicates the average number of packets per message. We also observe that as we let \( \alpha \to 0 \), we obtain \( \bar{a}_i \to (\alpha C_i)^{-1} \) and (23) yields the expected delay formula for directly processed messages (so that no fixed packets need be generated).

IV. THE N-CHANNEL COMMUNICATION PATH

We consider now an \( N \)-channel path. Here we have for the message interarrival time,

\[
T_{n+1}^{(i)} = I_{n+1}^{(i)} - I_n^{(i)} = T_{n+1}^{(i-1)} - T_n^{(i-1)}
\]

\[
= \begin{cases} 
\frac{M_a a_{i-1}}{T_n^{(i-1)}} & \text{if } W_{n+1}^{(i-1)} > 0, \\
\frac{M_a a_{i-1} + I_{n+1}^{(i-1)}}{T_n^{(i-1)}} & \text{if } W_{n+1}^{(i-1)} = 0,
\end{cases} \quad (24)
\]

where \( I_{n+1}^{(i-1)} \) denotes the duration of the idle period prior to \( I_{n+1}^{(i-1)} \) at channel \( (i - 1) \). The message waiting time at channel \( i \) follows the relationship (derived by observing the following Theorem 2 and Corollary 2, see also [8, Lemma 1])

\[
W_{n+1}^{(i)} = \left[ W_n^{(i)} + M_a a_i - T_{n+1}^{(i)} \right]^+. \quad (25)
\]

Utilizing (24) and (25), the analysis follows exactly as in [1]. In particular, a sequence of ladder indices \( \{ k_i \} \) is defined so that \( k_1 = 1, k_2, \ldots, k_m \) is the index of the first channel \( i, k_{n+1} < i < N \), so that \( a_i > a_{i-1} \). The channels \( 1 = k_1, k_2, \ldots, k_m \) are then called ladder channels. The packet transmission times for the ladder channels are \( a_{k_2} < a_{k_3} < \cdots < a_{k_m} \). Proceeding then as in [1], we obtain the following result.

Theorem 2: In an \( N \)-channel path, if \( i \) is not a ladder index for \( \{ a_i \} \), then

\[
W_n^{(i)} = 0,
\]

for each \( n \geq 1; i \geq 2 \).

In particular, we note (as in [1, corollary 1]) that for any ladder channel \( k_i, 2 \leq k_i \leq m \), and any \( n \geq 1 \), we have \( W_n^{(k_i)} = 0 \) if \( W_n^{(i)} = 0 \). However, in the present case \( W_n^{(i)} > 0 \) for any \( j > 1 \), since every packet (except the first one) has to wait for the packet leader to be served first. Hence we have deduced the following conclusion.

Corollary 2: For any ladder channel \( k_i, 2 \leq k_i \leq m \), and any \( n \geq 1 \), we have \( W_n^{(k_i)} = 0 \) if \( W_n^{(i)} = 0 \). Also, \( W_n^{(k_i)} > 0 \) for any \( j > 1 \), so that any nonleading packet has a positive waiting time for any ladder channel.

The overall \( n \)th message waiting time over the first \( k \) channels is defined as

\[
S_n^{(i)} \triangleq \sum_{i=1}^{k} W_n^{(i)} , \quad n \geq 1, \quad k \geq 1. \quad (26)
\]

Incorporating (24) and (25) in [1, lemma 2], we deduce the following result (for a proof, see also [8, appendix 1]).

Lemma 1: The random variables \( S_n^{(i)} \) satisfy the relationship, \( k \geq 1, n \geq 1 \),

\[
S_{n+1}^{(i)} = [S_n^{(i)} + M_a \max (a_1, a_2, \cdots, a_n) - T_{n+1}^{(i)}]^+. \quad (27)
\]

It is of particular interest to note that \( S_n^{(i)} \) is statistically independent of \( M_a \) and \( T_{n+1}^{(i)} \), so that \( S_n^{(i)} \) is governed by the same distribution as the waiting time variable in a \( M/G/1 \) queueing system with service time \( M_a \max (a_1, a_2, \cdots, a_n) \). Also, as in [1], the overall waiting time is invariant to the order of the channels. Hence, ordering the channels so that the one with minimal capacity is the first one yields zero message waiting times at all the other channels (by Theorem 2) and thus implies the latter distribution as well. Utilizing the latter ordering, we also conclude that

\[
r_n^{(N)} - r_n^{(N)} = (M_n - 1) a_{\max}. \quad (28)
\]

where \( a_{\max} = \max (a_1, a_2, \cdots, a_N) \), so that the first and the last packets of the \( n \)th message depart from the \( N \)th channel in a time difference given by (28). The latter is the reassembling delay for the \( n \)th message. Subsequently, the overall \( n \)th message time delay through the path \( \bar{r}_n \) as defined by (10) is given as

\[
\bar{r}_n \triangleq r_n^{(N)} - r_n^{(N)} = r_n^{(N)} - r_n^{(N)} + r_n^{(N)} - r_n^{(N)} = \sum_{i=1}^{N} W_i^{(i)} + \sum_{i=1}^{N} a_i + (M_n - 1) a_{\max}
\]

\[
= S_n^{(N)} + \sum_{i=1}^{N} a_i + (M_n - 1) a_{\max}. \quad (29)
\]

The time delay results can thus be summarized as follows.
Theorem 3:

1) The overall waiting time for the n-th message at an N-channel path, \( S_n^{(N)} \), has the same distribution as the waiting time \( W_n \) in a \( M/G/1 \) queueing system with Poisson arrivals with rate \( \lambda \) and service times distributed as \( M_{\text{max}} \). If \( \rho_{\text{max}} \triangleq \frac{\lambda M_{\text{max}}}{\lambda M_{\text{max}} - 1} < 1 \), the limiting distribution of the overall waiting time exists and is given by

\[
W(t) = \lim_{n \to \infty} P(S_n^{(N)} \leq t) = (1 - \rho_{\text{max}}) \sum_{m=0}^{\infty} \rho_{\text{max}}^m (\tilde{a}_{\text{max}})^{-1} \cdot \int_0^t [1 - B(\tau)] d\tau, \quad t > 0, \tag{30}
\]

where

\[
B(t) = P[M_{\text{max}} \leq t] = \sum_{m=0}^{\infty} g(m).
\]

If \( \rho_{\text{max}} \geq 1 \), then \( \lim_{n \to \infty} P(S_n^{(N)} \leq x) = 0 \) for each \( x \).

2) The overall limiting average waiting time over the \( N \)-channel path is given by

\[
W \triangleq \sum_{i=1}^{N} W^{(i)} = \lim_{n \to \infty} E[S_n^{(N)}] = \frac{1}{2} \rho_{\text{max}} \tilde{a}_{\text{max}} \left( 1 + C_M^2 \right), \tag{31}
\]

where \( C_M^2 \) is the coefficient of variation associated with \( g(m) \) and is given by (15).

3) The overall (steady-state) average time delay over the path is equal to

\[
\gamma \triangleq \lim_{n \to \infty} E[\gamma_n] = \frac{1}{2} \rho_{\text{max}} \tilde{a}_{\text{max}} (1 + C_M^2) + \sum_{i=1}^{N} a_i = \tilde{a}_{\text{max}} - a_{\text{max}}. \tag{32}
\]

4) The overall waiting and delay times over an \( N \)-channel path with capacities \( (C_1, C_2, \ldots, C_N) \) are the same as those over an \( N \)-channel path with capacities \( (C_{i1}, C_{i2}, \ldots, C_{iN}) \), where the latter sequence is an arbitrary ordering of \( (C_1, C_2, \ldots, C_N) \). The overall waiting time depends only on the minimal capacity \( min(C_1, C_2, \ldots, C_N) \).

Average message waiting times and queue sizes at the individual channels follow by observing that

\[
\tilde{W}_n^{(k)} = S_n^{(k)} - S_n^{(k-1)}. \tag{33}
\]

Noting that \( \tilde{W}_n^{(k)} \) is the waiting time for the first packet of the \( n \)-th message, and that the average message inter-arrival times are \( E[T_n^{(k)}] = \lambda^{-1} \), for each \( k \), the limiting average queue size at channel \( k \) is deduced, by Little's theorem, to be given as

\[
\tilde{X}^{(k)} \triangleq \lim_{n \to \infty} E[\tilde{X}_n^{(k)}] = \lambda \tilde{a}_k + \lim_{n \to \infty} E[\tilde{W}_n^{(k)}]. \tag{34}
\]

Consequently, we obtain the following result.

Theorem 4: For a ladder channel \( k_i, 1 \leq k_i \leq m \), if \( \rho_{k_i} = \lambda \tilde{a}_{k_i} < 1 \), the limiting average waiting time at the channel is equal to

\[
\tilde{W}^{(k_i)} = \lim_{n \to \infty} E[\tilde{W}_n^{(k_i)}] = (1 + C_M^2) \left[ \frac{1}{2(1 - \rho_{k_i})} a_{k_i} - \frac{1}{2(1 - \rho_{k_{i-1}})} a_{k_{i-1}} \right], \tag{35}
\]

where \( a_0 \triangleq 0 \). For \( \rho_{k_i} \geq 1 \), \( \tilde{W}^{(k_i)} = \infty \). The buffer memory size for channel \( k_i \) is equal to

\[
\tilde{M}^{(k_i)} = \beta \lim_{n \to \infty} E(\tilde{X}_n^{(k_i)}) = \lambda \beta [\tilde{W}^{(k_i)} + \tilde{a}_{k_i}] \text{ bits.} \tag{36}
\]

For a nonladder channel \( k \), \( \tilde{W}^{(k)} = 0 \) and the memory size required is \( \tilde{M}^{(k)} = \lambda \tilde{a}_k \) bits.

The techniques developed in [36] are utilized in [37] to derive further information regarding the evolution of the queueing stochastic processes at the individual channels (and obtain channel waiting-time distributions and busyperiod characteristics).

V. EXAMPLE—EXPONENTIALLY DISTRIBUTED MESSAGE LENGTHS, AND THE ZERO-PACKET CASE

Consider the arriving message lengths to be exponentially distributed with average length \( \mu^{-1} \). As indicated by (5), \( g(m) \) is then a geometric distribution. Using (21), (22), and (32), we then find the average time delay over the \( N \)-channel to be given by

\[
\gamma = \frac{1}{2(1 - \rho_{\text{max}})} \tilde{a}_{\text{max}} [1 + \exp(-\mu \alpha)] + \sum_{i=1}^{N} a_i = \tilde{a}_{\text{max}} - a_{\text{max}}. \tag{37}
\]

where

\[
a_{\text{max}} = \frac{a}{\min(C_1, C_2, \ldots, C_N)} \quad \text{and} \quad \tilde{a}_{\text{max}} = \tilde{M}_{\text{max}} = \frac{1}{1 - \exp(-\mu \alpha)} \frac{a}{C_{\text{min}}}, \tag{38}
\]

\[
\rho_{\text{max}} \triangleq \frac{\lambda M_{\text{max}}}{\lambda M_{\text{max}} - 1}.
\]

If we now let \( \alpha \to 0 \), we obtain \( a_{\text{max}} \to 0 \) and \( \tilde{a}_{\text{max}} \to 1/\mu C_{\text{min}} \triangleq a_{\text{max}}^{(e)} \).

\[
\lim_{a \to 0} a^{(e)}_{\text{max}} = \frac{\rho_{\text{max}}^{(e)}}{1 - \rho_{\text{max}}^{(e)}} a_{\text{max}}^{(e)} + a_{\text{max}}^{(e)} = \frac{a_{\text{max}}^{(e)}}{1 - \rho_{\text{max}}^{(e)}}, \tag{40}
\]

where \( \rho_{\text{max}}^{(e)} = \lambda a_{\text{max}}^{(e)} \). Equation (40) is the delay formula associated with a single \( M/M/1 \) queueing system with mean service time equal to \( 1/\mu C_{\text{min}} \). Thus, as packet sizes are reduced, the overall time delay converges to that associated with a single channel of capacity \( C_{\text{min}} \), which processes exponentially distributed message lengths. Similarly, for the case of a general message-length distribution (1), we obtain
where \( d_{\text{max}} = 1/\mu C_{\text{min}} \) and \( C_{\text{y}}^2 = \text{var}(Y_n) / [E(Y_n)] \).

Hence, we conclude that as \( \alpha \to 0 \), the overall message delay over the \( N \)-channel path reduces to that over a single \( M/G/1 \) channel with service time \( X_n = Y_n/C_{\text{min}} \), \( C_{\text{min}} = \min(C_1, C_2, \ldots, C_N) \).

We note that \( d_{\text{max}} \) (given by (39) for the exponential case) is minimized at \( \alpha = 0 \). Thus, as expected, the transmission time over the channel is reduced as the packet size is reduced. However, the coefficient of variation \( C_{\text{y}}^2 \) behaves differently. For the exponential case, \( C_{\text{y}}^2 = \exp(-\alpha \beta) \) which increases as \( \alpha \) is decreased. Thus, in the waiting-time expression a factor which causes an increasing delay as \( \alpha \) decreases is incorporated. This is due to the resulting increase in the variance of \( M_n \) as \( \alpha \) is decreased.

The variation of the message overall delay time is shown in Fig. 2 as a function of the traffic intensity \( \rho^* = \lambda/\mu \).

We assume exponentially distributed message lengths with mean \( \mu^{-1} \) and equal channel capacities \( C_i = 1 \), \( i = 1, 2, \ldots, N \). As an average delay index independent of the path length \( N \), we choose the delay \( \gamma_r^* \), where \( \gamma_r^* = \gamma_r - (N - 1)\alpha \alpha_i = \alpha/C_i = \alpha \), where \( \gamma_r \) is given by (37). The variation of \( \gamma_r^* \) versus \( \rho^* \) is observed for five values of packet length \( \alpha \); i.e., average packet length per message length \( (\mu \alpha) \) values of 0, 0.1, 0.5, 1, and 2, assuming \( \mu = 0.1 \). To obtain the overall message delay one needs to compute \( \gamma_r = \gamma_r^* + (N - 1)\alpha \). The figure indicates that for low traffic intensities one can choose longer packet lengths without causing too high a resulting delay, as compared to the \( \alpha = 0 \) case. For higher intensities, the largest packet length one wishes to choose is dictated by the allowable path delay, as indicated in the figure. We further note the basic sources of message delay. The delay curve for \( \alpha = 0 \) presents the basic message delay on the path. The curves for \( \alpha > 0 \) incorporate the additional delays caused by the incompletely filled packets and the increased length of the message packets.

VI. CONCLUSIONS

We have derived formulas for the distribution of the overall message waiting and delay times, over an \( N \)-channel path in a packet-switching communication network. We have obtained, as well, expressions for the steady-state average message waiting times and required buffer sizes at the individual channels. We have assumed messages with random message lengths to be divided into fixed packets, which are then sent independently through the communication path and are reassembled when departing at the \( N \)th channel. The results indicate that the overall message waiting time over the path follows the waiting-time distribution derived for an \( M/G/1 \) queueing system with service times distributed as \( \alpha M_n/C_{\text{min}} \) where \( C_{\text{min}} \) is the minimal channel capacity over the path, \( \alpha \) is the packet length, and \( M_n \) denotes the number of packets associated with the \( n \)th arriving message. The overall reassembling time, for the \( n \)th message, is observed to be equal to \( \alpha (M_n - 1)/C_{\text{min}} \). The overall average message delay time over the communication path is given by (32).

Considering exponentially distributed message lengths as an illustrating example, the message delay time is given by (37) and plotted in Fig. 2 as a function of the traffic intensity for several values of packet length. We note that our results can readily be applied to arbitrary input processes (see Lemma 1), using \( G/G/1 \) results. A detailed derivation of the message and packet waiting-time, delay-time, idle-period, and busy-period distributions at the channels along the path is presented in [8]. Time-delay problems for more involved topological and flow situations are currently under investigation. In particular, we note that the problem of message time-delay calculation when the communication route is composed of \( k \) vertex disjoint paths is equivalent to that of an equivalent queueing system with \( k \) channels (servers) each with capacity equal to the minimal capacity of the related path. Also, using the results of this paper, an analysis technique is being developed for evaluating message delays in a packet-switching network with interfering paths. The latter message delays will then be minimized by incorporating appropriate adaptive packet routing procedures. The latter manifest the advantage in using packet-switching.

ACKNOWLEDGMENT

The author wishes to thank Prof. L. Kleinrock of the Computer Science Department, University of California, Los Angeles, for a useful discussion and for suggesting the application of the results of [1] and [2] to a packet-switching situation.

REFERENCES


Nonlinear Analysis of Correlative Tracking Systems
Using Renewal Process Theory

HEINRICH MEYR

Abstract—A new method is presented which describes the behavior of an \((N + 1)\)-order tracking system in which the nonlinearity is either periodic [phase-locked loop (PLL) type] or nonperiodic [delay-locked loop (DLL) type]. The cycle slipping of such systems is modeled by means of renewal Markov processes. A fundamental relation between the probability density function (pdf) of the single process and the renewal process is derived which holds in the stationary state. Based on this relation it is shown that the stationary pdf, the mean time between two cycle slips, and the average number of cycles to the right (left) can be obtained by solving a single Fokker-Planck equation of the renewal process.

The method is applied to the special case of a PLL and compared with the so-called periodic-extension (PE) approach. It is shown that the pdf obtained via the renewal-process approach can be reduced to agree with the PE solution for the first-order loop in the steady state only. The reasoning and its implications are discussed. In fact, it is shown that the approach based upon renewal-process theory yields more information about the system's behavior than does the PE solution.

INTRODUCTION

Correlative tracking systems, such as phase-locked loops (PLL's), Costas loops, and delay-locked loops (DLL's) have attracted much interest among researchers in the field of telecommunication and synchro-nization theory and are widely used in practice. A recent bibliog-raph [1] cites more than 800 papers and at present others are in the stirring. Even though much has been published regarding their behavior and perfor-mance [2], several important problems have not been addressed nor even formulated. New applications arise, for example, in modern mass transportation systems which require precise velocity and distance measurement systems. One proposed solution to the problem \([3]-[6]\) uses a DLL system to estimate the time difference between two versions of the same stochastic signal. It can be shown that this time difference is inversely proportional to the velocity. In the analysis of such a system \([6]\), the intrinsic noise \([7], [8]\) has been neglected for most communications applications; however, in mass transportation systems, it is of central importance and determines the limit of usefulness of such systems.

This paper consists of three sections. In Section I, background material and motivation for the problems to be treated are given. In Section II, a new method for the analysis of correlative tracking systems is presented. This method is based on the theory of renewal Markov processes and can be applied to systems with periodic nonlinearities as well as to systems with nonperiodic nonlinearities. It is shown that the stationary probability density function (pdf), the mean time between two cycles \(E(\tau_L)\) of the phase process, and the average number of