is also in error. It is evident that, with this representation, not all transitions are permissible. For example, the transition 00 → 10 is not permissible because it would require that the present transmission be both error free (00) and erroneous (10); consequently, the associated transition probability is zero. As before, we use M to denote the now 4×4 error transition matrix; the ijth element of M, which represents the probability of transition from state i to stage j, will be denoted by $a(i, j)$.

By adopting the same definition of a cycle as before, we have two types of cycles corresponding to initial states 01 and 11, respectively. If X_1 corresponds to the first block of a cycle, then in the present situation, we also need to consider a previous transmission X_0. First suppose, $X_0 = 0$ and $X_1 = 1$ (i.e., we are starting from the initial state 01), then the number of accepted blocks K_0 in the cycle, by considering all relevant possibilities, has distribution

$$\Pr[K_0=1]=a(01, 00|m)a(00, 01)+a(01,10|m)a(10, 01)$$

$$\Pr[K_0=k>1]=a(01, 00|m)a(00, 00)^{k-1}a(00, 01) + a(01, 10)|m)a(10, 00)a(00, 00)^{k-2}a(00, 01)$$

where $a(i, j|m)$ corresponds to transition from state i to state j after m steps; it is shown in [3] that $a(i, j|m)$ simply equals the ijth element of the matrix M^m. From the above distribution, the conditional average $E(K_0)$ may be obtained. Likewise, the average number of accepted blocks $E(K_1)$ conditional on the initial state 11 may be similarly obtained. If $v = (p_{00}, p_{01}, p_{10}, p_{11})$ represents the steady-state probability vector of M, then it may be obtained by solving the system of linear equations [3]: $vM = v$. The unconditional mean number of accepted blocks per cycle is thus $E(K) = p_{01}E(K_0) + p_{11}E(K_1)$. Hence, the efficiency for this situation can be obtained from (2.6) as before.

Repeating the same procedure would enable the evaluation of error patterns which may depend on an arbitrary number of past transmissions. In general, if dependency is required on the past n transmissions as well as the present one, then the number of entries in the corresponding error transition matrix is $2^{2(n+1)}$, which grows exponentially in n. This matrix, however, is a sparse one whose demand on memory will not be excessive because for an n-step dependent chain $\Pr[b_1b_2 \cdots b_{n+1} \rightarrow d_1d_2 \cdots d_{n+1}]$ is only nonzero if $b_2 = d_1$, $b_3 = d_2$, $\cdots$, $b_{n+1} = d_n$. Thus, each row of the matrix can have at most two nonzero elements, and so the proportion of nonzero elements in the $2^{n+1} \times 2^{n+1}$ matrix is at most 2^{-n}. Numerical evaluation of throughput efficiency accordingly appears to be quite feasible.

V. Summary and Conclusions

The efficiency of the go-back-N ARQ scheme under Markov error patterns has been analyzed and a new formula is presented which subsumes the conventional efficiency formula as a special case. A measure of transmission clustering—the clustering coefficient—is introduced and is found to play a major role in governing operating efficiency. We find that a Markov system with a high clustering coefficient is generally more efficient than an equivalent random error system; but one with a low clustering coefficient is less efficient than the random error system. By suitably extending the Markov representation, it is shown that the present method may be carried through to the analysis of quite general error patterns, and simple numerical procedures for efficiency computation is presented.

Acknowledgment

The authors are grateful to the three anonymous referees for their helpful comments and suggestions on an earlier version of the paper.

References

[1] "Data communications—high level data link control procedures—Elements of procedures," independent numbering, ISO Rep. 4336, 1976.

[2] J. P. A. Adoul, "Error intervals and cluster density in channel modeling," *IEEE Trans. Inform. Theory,* vol. IT-20, pp. 125-129, Jan. 1974.

[3] D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes.* London, England: Chapman and Hall, 1965.

[4] B. D. Fritchman, "A binary channel characterization using partitioned Markov chains," *IEEE Trans. Inform. Theory,* vol. IT-13, pp. 221-227, Apr. 1967.

[5] W. A. Fuller, *Introduction to Statistical Time Series.* New York: Wiley, 1976.

[6] L. N. Kanal and A. R. K. Sastry, "Models for channels with memory and their applications to error control," *Proc. IEEE,* vol. 66, pp. 724-744, July 1978, and pp. 125-129, Jan. 1974.

[7] S. Lin and P. S. Yu, "An effective error control sheme for satellite communications," *IEEE Trans. Commun.,* vol. COM-28, pp. 395-401, Mar. 1980.

[8] A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions.* Englewood Cliffs, NJ: Prentice-Hall, 1962.

Design and Analysis of a Contention-Based Lookahead Reservation Protocol on a Multichannel Local Area Network

P. C. WONG AND T. S. YUM

Abstract—**The contention-based lookahead reservation (CLAR) protocol is proposed to be used on a multichannel local area network. The protocol can provide fast-circuit-switching services which are particularly advantageous for networks supporting integrated services [1]. The delay and throughput performance for message transmission are obtained, and they agree closely with that obtained by simulation. The delay performance of CLAR is similar to the M-CSMA [2], [3] protocol for an M-channel network, but only CLAR can give a stable maximum throughput of $(M - 1)/M$ independent of the cable length. Moreover, CLAR requires only two sets of transceivers, while M-CSMA requires M. The lookahead reservation technique can provide 9 percent throughput increase for fixed size messages and 19 percent for geometrically distributed messages.**

I. Introduction

Many different types of channel access methods are possible for local area networks. Roughly speaking, they can be classified into 1) token-passing protocols [3], [4], 2) contention protocols [5], [6], and 3) reservation protocols [7]-[9]. Among them, the CSMA-type protocols have received a lot of attention for their simplicity and other nice features. However, the CSMA-type networks do have some drawbacks. First of all, if the channel traffic surges temporarily, packet collisions on the channel will increase, resulting in a decreased channel throughput. Second, for a network with small packet size and

Paper approved by the Editor for Local Area Networks of the IEEE Communications Society. Manuscript received September 9, 1986; revised May 15, 1987. This paper was presented at IEEE INFOCOM '87, San Francisco, CA, April 1987.

P. C. Wong is with the Department of Computing, Hong Kong Polytechnic, Hong Kong and the Department of Electronics, Chinese University of Hong Kong, Shatin NT, Hong Kong.

T. S. Yum is with the Department of Electronics, Chinese University of Hong Kong, Shatin NT, Hong Kong.

IEEE Log Number 8718651.

0090-6778/88/0200-0234$01.00 © 1988 IEEE

long cable length, the network performance will be poor [10].

Marson and Roffinella [2] proposed a multichannel CSMA-type computer network. They found that by dividing the channel into M subchannels, the throughput is increased. However, the protocol used on the network would require M transceivers for each station, and each station has to monitor the status of all subchannels for finding the appropriate channel to transmit packets. This means that a complicated transceiver is needed for each station. Performance degradation due to the increase of cable length (using the CSMA protocols) still exists, which limits the size of the network. The resulting network, therefore, is neither well suited for low-cost microcomputer communications nor for large-scale and long-distance packet transmission.

To combine the advantages of the CSMA protocol and multichannel topology, we propose, in Section II, the contention-based lookahead reservation (CLAR) protocol for use in a multichannel network. We then analyze this protocol in Section III and discuss its performance in Section IV.

II. Network Access Protocol

Fig. 1 shows an M-channel network. Here one channel is dedicated for signaling, which includes functions such as channel reservation, acknowledgments, channel release, system setup, etc. We choose the slotted ALOHA protocol for use on the signaling channel. This protocol has a very simple hardware implementation and its performance is well understood. But more importantly, its throughput is independent of cable length and slot size. The remaining M-1 channels are for data transmission through reservation on the signaling channel. Each station requires two transceivers. One transceiver is for the signaling channel, and the other one can be switched to any of the M-1 data channels.

Each station keeps a table recording the current information on the data channels. A typical table is shown in Table I. The second column of the table indicates the channel status. B or "busy" indicates that the channel is not available for reservation. R or "ready" indicates that the transmission on that channel is about to finish and is now open for contention by all stations. A station succeeded in reserving a "ready" channel therefore has to wait for the end of the previous transmission before it can start its own. I or "idle" indicates that the channel is allowed for immediate transmission. The third and fourth columns show the transmitting and the receiving station numbers, respectively.

A station with a message to transmit will first look up the channel status table to select an "idle" channel. If none is available, a "ready" channel is selected. It then contends on the signaling channel to transmit a channel reservation signal unit or SU (Fig. 2). If a station succeeds in transmitting a reservation packet, the immediately following slot is reserved for the destination station to reply with a "reservation acknowledgment (RACK)" signal unit. All stations will refrain from transmission in that slot.

After the RACK SU is received from the destination station, the source station will wait, if necessary, for the reserved ready channel to go idle and then transmit its data packets. After detecting the reservation acknowledgment SU, all stations change the status of the reserved channel to "busy." When the source station has L slot of bits remaining to be sent, it broadcasts a channel release SU through the signaling channel. All stations then change the status of that data channel to "ready," which means that the channel is ready for reservation again. Note that this "lookahead" reservation mechanism is deliberately introduced so that the transmission of the present message on the data channel and the scheduling of the next message on the same channel are done simultaneously. The next message can be transmitted as soon as the present message transmission is finished. The lookahead interval L should be chosen to be much larger than the sum of the average channel-release and channel-reservation delays.

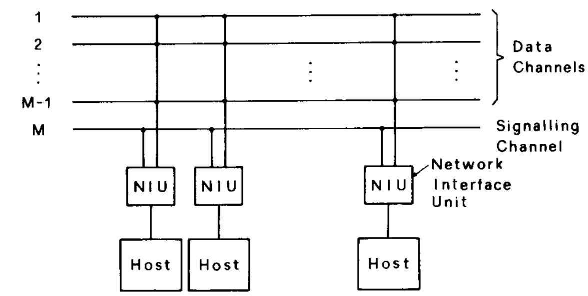

Fig. 1. An M-channel local area network.

TABLE I
CHANNEL STATUS TABLE

Channel Number	Channel Status: B/R/I	Transmit Station	Receive Station
1	B	15	6
2	I	–	–
.	.	.	.
8	R	13	8
9	B	18	27

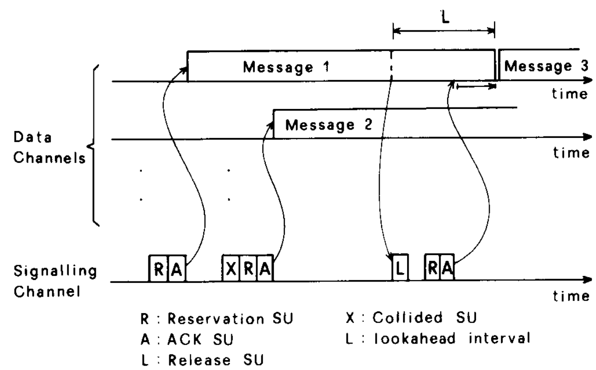

Fig. 2. Reservation and release of data channels through the signaling channel.

III. Analysis

We present in the following a simplified model which can give a lower bound on the network throughput and an upper bound on the average message delay. Let there be a total of N stations in the network. A station is READY when it has finished its previous transmission. In READY states, messages are generated according to independent Bernoulli trials. This means that at every slot, a READY station generates a message with probability q. The READY state duration therefore is geometrically distributed. A station enters the QUEUEING state when a new message is generated. If an idle or ready channel is available, it will immediately contend for that channel. After a successful contention, a QUEUED station becomes a RESERVED station and enters a reserved queue. As soon as a channel becomes idle, the corresponding reserved station can start transmission.

We assume that the message sizes are geometrically distributed with mean $\bar{B} = 1/v$ where v is the probability that a station would complete its message transmission at each time slot. This choice is for mathematical convenience since a geometric distribution possesses the memoryless property needed for Markovian analysis. We further assume that the lookahead interval is sufficiently long so that channel-release contention is always completed before the end of the interval. Since a time slot is always reserved for RACK after each successful reservation, it need not be modeled as a separate transaction.

The multichannel network is modeled as a closed queueing network shown in Fig. 3. A total of N customers are circulating in the network. We divide the network model into three sections: the arrival section, the contention section, and the channel section.

1) The channel section contains $M - 1$ parallel servers representing the $M - 1$ data channels and a "waiting room" of size $M - 1$ which can accommodate up to $M - 1$ customers (stations) that have channels reserved. Let $\tilde{m}$ be the total number of customers in the channel section.

2) The contention section is a single-server queue where the $\tilde{n}$ QUEUED customers are waiting for channel contention. Channel contention begins whenever there is room for more customers in the channel section, i.e., $\tilde{m} < 2M - 2$. When $\tilde{m} = 2M - 2$, the contention server stops working.

3) The arrival section has $N - \tilde{n} - \tilde{m}$ READY customers. At every time slot, each READY customer leaves the arrival section with probability q.

Let (n, m) be the state vector of the above closed queue network. At each time slot, the following events could occur.

1) Generation of i new messages: state transition from (n, m) to $(n + i, m)$.

2) Service completion of j messages: state transition from (n, m) to $(n, m - j)$.

3) Successful channel contention: state transition from (n, m) to $(n - 1, m + 1)$.

Let $r_i(n, m)$ = Prob [i stations generate messages in state (n, m)]. Let the time slot be small enough so that the probability of four or more message arrivals in a slot is negligible. We therefore have

$$r_i(n, m)=\begin{cases}(1-q)^{N-n-m} & i=0\\ (N-n-m)q(1-q)^{N-n-m-1} & i=1\\ \binom{N-n-m}{2} q^2(1-q)^{N-n-m-2} & i=2\\ 1-r_0-r_1-r_2 & i=3.\end{cases}$$

Similarly, let $q_j(m)$ = Prob [j messages have finished transmission in a slot in state (n, m)]. This probability is independent of n. Again, with some slots small enough, the probability of four or more departures from the channel section in a time slot is assumed negligible. We have

$$q_j(m)=\begin{cases}(1-v)^m & j=0\\ mv(1-v)^{m-1} & j=1\\ \binom{m}{2} v^2(1-v)^{m-2} & j=2\\ 1-q_0-q_1-q_2 & j=3.\end{cases}$$

The probability of successful contention in state (n, m) is

$$C(n, m)=\begin{cases}u & \text{for } m<2M-2 \text{ and } n>0\\ 0 & \text{otherwise}\end{cases}$$

where u is an average value obtained from simulation results for a heavily loaded signaling channel. Then the transition probabilities are given by

$$\text{Prob }[(n, m)\rightarrow(n+i-1, m-j+1)] = r_i q_j C(n, m), \qquad i, j\in\{0, 1, 2, 3\}$$

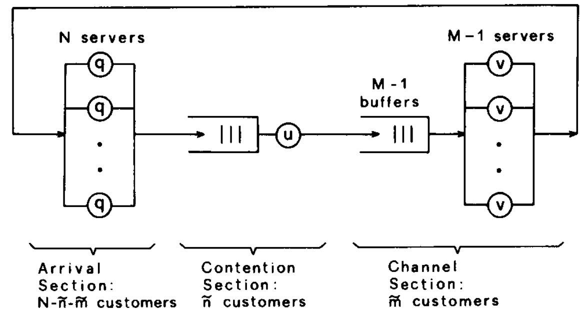

Fig. 3. Queueing model for the CLAR protocol in an M-channel LAN.

for a successful contention; and

$$\text{Prob }[(n, m)\rightarrow(n+i, m-j)] = r_i q_j[1-C(n, m)], \qquad i, j\in\{0, 1, 2, 3\}$$

for an unsuccessful contention.

We solved this two-dimensional Markov chain by the Gauss-Seidel iteration method to obtain the equlibrium state probabilities $\pi(n, m)$. Let $a(k)$ = Prob [K channels are in use], and $b(j)$ = Prob [the contention section and the channel section have a total of j messages].

Then

$$a(k)=\begin{cases}\sum_{n=0}^{N}\pi(n, k) & k=0, 1, \cdots, M-2\\ \sum_{m=M-1}^{2M-2}\sum_{n=0}^{N}\pi(n, m) & k=M-1\end{cases}$$

$$b(j)=\sum_{n+m=j}\pi(n, m) \qquad j=0, 1, 2, \cdots, N.$$

The average number of channels $\tilde{m}$ in use is $\Sigma_{k=0}^{M-1} ka(k)$, and the average number of stations in contention and channel sections N_s is $\Sigma_{j=0}^{N} jb(j)$. For a closed queue network, the rate of message generation $\mathbb{R}$ is equal to the throughput at the channel section, or $\mathbb{R} = \tilde{m}v$. Hence, the average delay T by Little's formula is $T = N_s/\mathbb{R} = N_s/\tilde{m}v$ and the throughput is $S = \tilde{m}/M$.

IV. Results and Discussions

We obtain the delay throughput characteristics of a network with a total channel capacity of 10 Mbits/s. This total capacity is evenly divided by the M channels. Fig. 4 shows the normalized message transmission delay as a function of the network throughput for various values of M. Also shown are the simulation results using GPSS-V. As expected, we found that a steady maximum throughput is maintained under very heavy loading condition. This maximum throughput, moreover, is independent of the number of stations connected to the network.

Figs. 5 and 6 compare the throughput and delay performance of CLAR to M-CSMA [2] on a multichannel network and to the CSMA on a single channel with the same total channel capacity. Here we choose the normalized propagation delay a to be 0.075, and use a fixed message size of 50 slots. The results show that for $M = 10$, CLAR gives a maximum throughput of 0.9, while the throughput of the single-channel CSMA protocol is limited to 0.6. The average delay of CLAR is similar to M-CSMA at low utilization. But when the offered traffic increases beyond the network capacity, stable maximum throughput is maintained for CLAR only. The CLAR

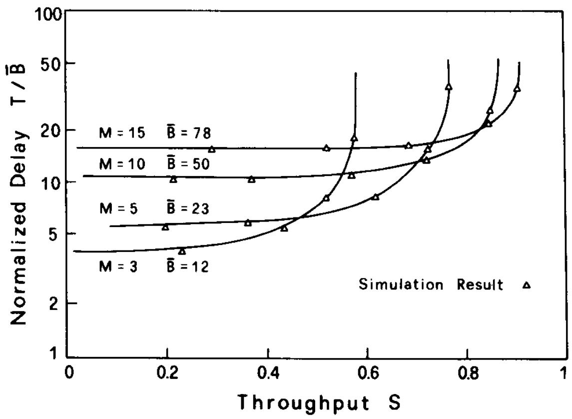

Fig. 4. Delay-throughput characteristics of CLAR: comparison of analytical and simulation results for geometrically distributed message lengths.

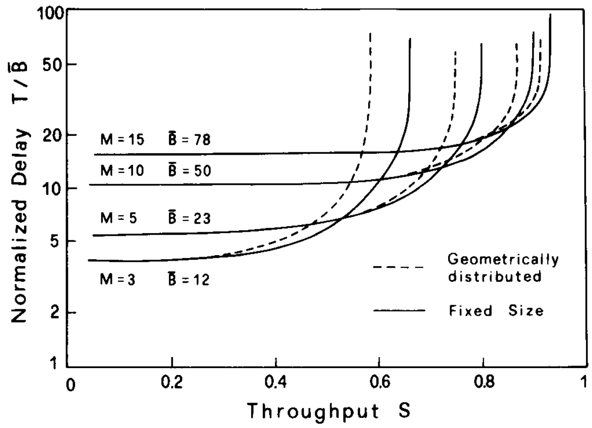

Fig. 7. Delay-throughput characteristics of CLAR as function of M and $\bar{B}$ for fixed and geometrically distributed messages.

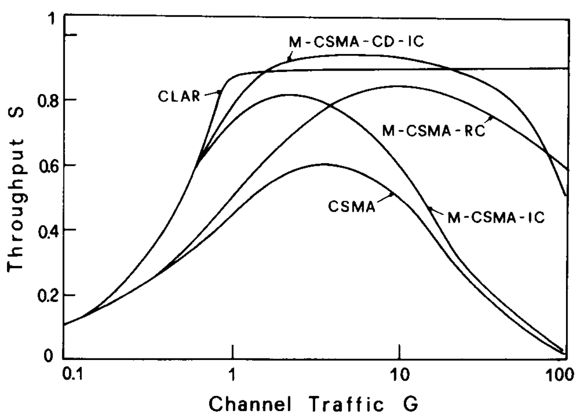

Fig. 5. Throughput comparisons: CLAR, M-CSMA, and CSMA with M = 10.

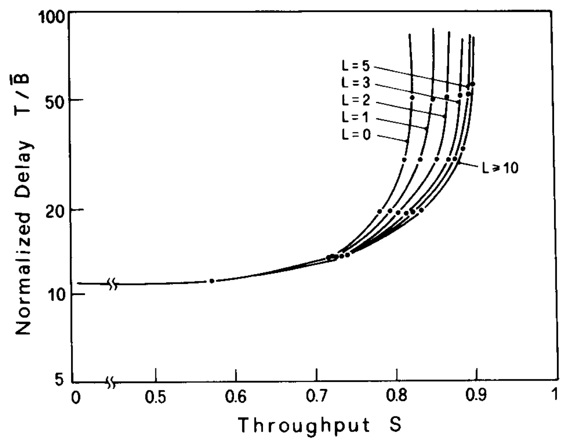

Fig. 8. The effect of lookahead interval with fixed message size B = 50 slots, M = 10.

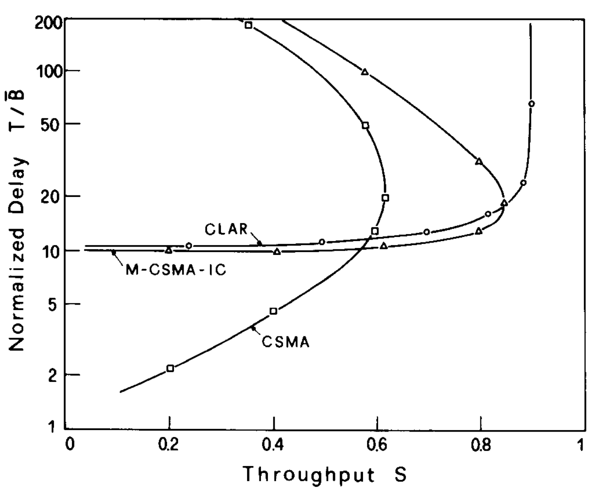

Fig. 6. Throughput comparisons: CLAR, M-CSMA, and CSMA with M = 10.

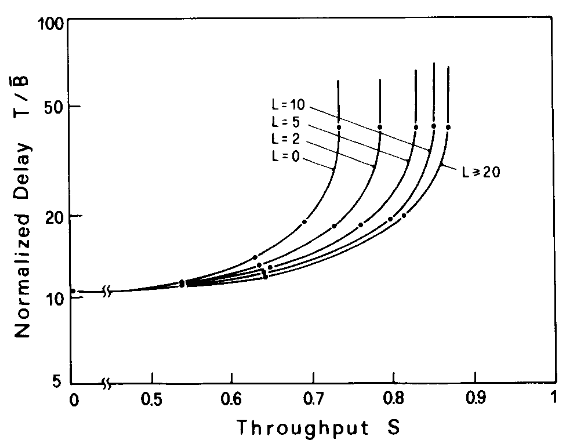

Fig. 9. The effect of lookahead interval with geometric message size $\bar{B}$ = 50 slots, M = 10.

protocol therefore can be used even for a larger network under heavy loading conditions.

Fig. 7 shows the throughput-delay characteristics of CLAR with M and a as parameters for both fixed and geometrically distributed message lengths using simulation. The maximum throughput for the geometrically distributed messages is slightly smaller than that for the fixed size messages. This is because for geometrically distributed messages, some will have sizes smaller then L slots (the lookahead interval size), and so the transmission may terminate before a successful reservation can be made and results in the data channel being idle for a few slots. The normal operation of the protocol, however, is not affected.

Fig. 8 shows the effect of L on delay and throughput for fixed size messages. We see that any value of L greater than ten slots can be chosen to obtain maximum throughput. Comparing to the same protocol without lookahead reservation (i.e., $L = 0$), we found that lookahead reservation can increase the throughput by 9 percent. Fig. 9 shows the same for the geometrically distributed messages. Here, $L \geq 20$ is needed for maximum throughput. Again, comparing to the case without lookahead reservation, a throughput increase of 19 percent is observed.

References

[1] T. S. Yum and P. C. Wong, "An integrated services local area network," in *Proc. INFOCOM '87,* San Francisco, CA, Mar. 1987, pp. 999–1008.

[2] M. Ajmone Marsan and D. Roffinella, "Multichannel local area network protocols," *IEEE J. Select. Areas Commun.*, pp. 885–897, Nov. 1983.

[3] "Local area network standards: Token-passing bus access methods and physical layer specifications," ANSI/IEEE Standard 802.4, 1985.

[4] "Local area network standards: Token-passing ring access methods and physical layer specifications," ANSI/IEEE Standard 802.5, 1985.

[5] "Local area network standards: Carrier sense multiple access with collision detection (CSMA/CD) and physical layer specifications," ANSI/IEEE Standard 802.3, 1985.

[6] F. A. Tobagi, "Multiaccess protocols in packet communication systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 468–488, Apr. 1980.

[7] W. R. Crowther *et al.,* "A system for broadcast communication: Reservation-ALOHA," in *Proc. 6th Hawaii Int. Syst. Sci. Conf.,* Jan. 1973.

[8] S. S. Lam, "Packet broadcast networks: A performance analysis of the *R*-ALOHA protocol," *IEEE Trans. Commun.*, vol. COM-29, pp. 596–603, July 1980.

[9] A. S. Tanenbaum, *Computer Networks.* Englewood Cliffs, NJ: Prentice-Hall, 1981.

[10] W. Stallings, "Local network performance," *IEEE Commun. Mag.*, pp. 27–36, Feb. 1984.

A Method to Dramatically Improve Subcarrier Tracking

WILLIAM J. HURD AND SERGIO AGUIRRE

Abstract—**A method is presented for achieving a dramatic improvement in phase tracking of square wave subcarriers or other square waves. The method is to set the amplitude of the phase quadrature reference signal to zero, except near the zero crossings of the input signal. Without changing the loop bandwidth, the variance of the phase error can be reduced to approximately $W\sigma_0^2$ where σ_0^2 is the phase error variance without windowing and W is the fraction of cycle in which the reference signal has a nonzero value. Simulation results confirm the analysis and establish minimum W versus SNR. Typically, the window can be made so narrow as to achieve a phase error variance of 1.5 σ_0^4.**

Paper approved by the Editor for Signal Design, Modulation, and Detection of the IEEE Communications Society. Manuscript received November 7, 1986; revised August 25, 1987. This work was supported under contract by the National Aeronautics and Space Administration.

The authors are with the Jet Propulsion Laboratory, Pasadena, CA 91103.

IEEE Log Number 8718636.

I. Introduction

In deep space telemetry, the data bits are normally channel encoded for low error rate transmission from the spacecraft to earth. The binary symbol waveforms are modulated onto a square-wave subcarrier, which in turn phase modulates a sinusoidal carrier. To aid carrier acquisition, a residual carrier component is transmitted by controlling the modulation index. Use of subcarriers originated from the need to move the modulated spectrum away from the carrier frequency, especially at very low data rates [1].

The receiving system must track the phases of the carrier and subcarriers, and errors in tracking these phases cause losses in effective telemetry bit signal-to-noise ratio (SNR). The loss is often more severe for subcarrier tracking than for carrier tracking because the subcarriers are square waves, whereas the carriers are sinusoidal. The SNR loss varies approximately as the mean-square phase error for sinusoids, but only as the rms phase error for square waves.

Subcarrier tracking loss is most significant in low-rate telemetry systems where the subcarrier loop bandwidth cannot be made narrow enough to reduce the rms phase error to a small enough value. For example, current subcarrier demodulators in the Deep Space Network (DSN) track the Pioneer 10 spacecraft with an average loss in symbol SNR of 0.4 dB at 16 bits/s and 0.6 dB at 8 bits/s when the symbol SNR is 0 dB. The actual loss in decoder SNR threshold is as much as several decibels because of the correlation of phase error over the code block lengths or constraint lengths.

These losses motivated the analysis and simulation of the improved subcarrier tracking method presented here. The method is capable of reducing the loss in average symbol SNR (SSNR) to under 0.1 dB for the Pioneer example without reducing the loop bandwidth.

II. Method and Performance

The improvement in subcarrier tracking is achieved by windowing one of the subcarrier channel reference signals as done in a digital data transition tracking loop (DTTL) bit synchronizer [2]. A theoretical basis for this method was presented by Layland [3] who concluded that, for a first-order phase-locked loop and high-loop SNR, the optimum reference signals needed to track square waves resemble alternating trains of narrow pulses.

A maximum likelihood estimation strategy of the subcarrier phase which is not data aided (NDA) needs to average the likelihood function (of the received signal conditioned on the phase to be estimated) over the random data. The derivative of the log-likelihood function suggests a Costas loop where a tanh $(\cdot)$ nonlinearity is incorporated into the upper arm of the NDA loop [4], while the other arm should have as a reference signal the derivative of the square wave. The derivative is approximated here by the gating function, while the tanh $(\cdot)$ nonlinearity is approximately equal to its argument if the argument is small. The loop considered in this paper is then a practical low SNR implementation of the NDA maximum likelihood estimator of the subcarrier phase.

Fig. 1 shows the windowed quadrature phase reference waveform and its relationship to the subcarrier and the standard reference waveform. Let W be the fraction of each cycle of the reference signal which has a nonzero value. The reference signal looks like a square wave multiplied by zero, except for the regions within plus or minus $WT_{sc}/4$ of the zero crossings as illustrated in Fig. 1. The theoretical improvement in loop SNR is approximately a factor of $1/W$, provided that the phase error is small enough that the loop is in the linear region. Based on simulation results, values of W from 1/16 to 1/64 appear practical in cases for which the loop SNR would otherwise be low enough to cause significant symbol SNR