SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
	AD-A086655	

4. TITLE (and Subtitle)
(14) Bounds on the Utilization of Aloha-Like Multiple-Access Broadcast Channels

5. TYPE OF REPORT & PERIOD COVERED
Paper-Technical

6. ~~PERFORMING ORG.~~ REPORT NUMBER
(14) LIDS-P-1000

7. AUTHOR(s)
(10) Pierre A. Humblett

8. ~~CONTRACT OR GRANT~~ NUMBER(s)
ARPA Order No. 3045/5-7-75
ONR/N00014-75-C-1183

9. PERFORMING ORGANIZATION NAME AND ADDRESS
M.I.T.
Laboratory for Information and Decision Systems
Cambridge, MA 02139

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Program Code No. 5T10
ONR Identifying No.049-383

11. CONTROLLING OFFICE NAME AND ADDRESS
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

12. REPORT DATE
(12) June 1980

13. NUMBER OF PAGES
15

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
Office of Naval Reseearch
Information Systems Program
Code 437
Arlington, Virginia 22217

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

LEVEL II

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(11) June 1980

DTIC ELECTE JUL 14 1980 D C

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(15) N00014-75-C-1183
ARPA Order-3045

18. SUPPLEMENTARY NOTES

(9) Technical Report

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Pippenger's model for synchonous protocols that resolve conflicts among message transmissions on a multiple-access broadcast channel is extended to the case where the transmission times depend on the outcomes of the transmissions An information theoretic method is used to provide new bounds on the utilization of the channel.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

410950

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

80 7 11 014

June 1980 LIDS-P-1000

BOUNDS ON THE UTILIZATION OF ALOHA-LIKE MULTIPLE-ACCESS BROADCAST CHANNELS*

by

Pierre A. Humblet**

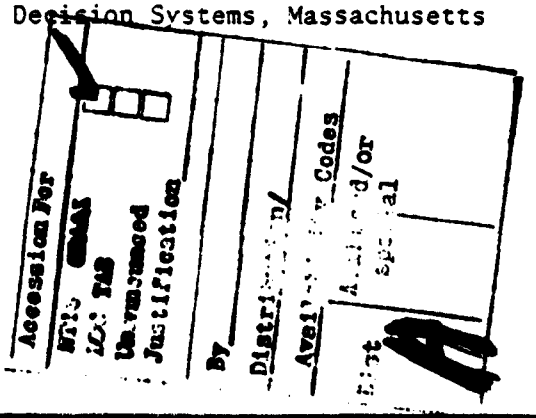
ABSTRACT

Pippenger's model for synchonous protocols that resolve conflicts among message transmissions on a multiple-access broadcast channel is extended to the case where the transmission times depend on the outcomes of the transmissions. An information theoretic method is used to provide new bounds on the utilization of the channel.

---

*This research was conducted at the M.I.T. Laboratory for Information and Decision Systems with partial support provided by NSF under Grant ENG-77-19971 and by ARPA under Grant ONR/N00014-75-C-1183.

**Room. 35-203, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139.



Introduction

Consider the following model of the generalized ALOHA system. Geographically separated but time synchronized transmitters send and receive messages on a common channel. If no transmitter is active, this fact is recognized by all within t_0 seconds. If exactly one transmitter sends a message, the message is received successfully and this is known to all within t_1 seconds. Finally, if two or more transmitters are active simultaneously, then a collision is said to occur and it is detected by all within t_2 seconds. All messages involved in the collision must be retransmitted at a later time.

This model represents a variety of systems. The slotted ALOHA channel [1] has $t_0 = t_1 = t_2$. Carrier sense multiple access radio systems [2] can detect idles quickly (carrier not present) while they distinguish between collisions and successes by using error detecting codes. Thus they have $t_0 << t_1 = t_2$. Some broadcast cable systems (e.g., the Ethernet [3]) have a "listen while transmit" feature that allows the quick abortion of transmission when a collision is detected. Thus typically $t_0 = t_2 << t_1$. Finally "reservation" systems use short messages to reserve time for longer data messages. The short messages can be seen as an idle/collision detection mechanism, and again $t_0 = t_2 << t_1$ [4].

We define the utilization of a channel access scheme as the fraction of the time during which messages are successfully transmitted. We define the "capacity" of this channel as the supremum, over all schemes, of the utilization. If the number of transmitters if finite, then the capacity is 1. Simple schemes like synchronous time division multiplexing or round robin transmission (cyclic polling) avoid collisions and can achieve this capacity. Unfortunately

they cause relatively long message delays when the generation rate of the messages is much smaller than $1/t_1$. In that case "random" transmission schemes are preferred. They allow collisions in the hope of reducing delay. Such random schemes are customarily analyzed assuming that they are infinitely many transmitters, each generating at most one message during its life-time, and that the global generation process of the messages is Poisson with rate λ.

The capacity of the channel under those conditions is still unknown. An early scheme, the slotted Aloha [1] strategy, has been said to have an utilization of 1/e (when $t_o = t_1 = t_2$), but has been shown to be unstable, i.e., with probability one its utilization decreases to 0 as time goes by. A new class of protocols has recently been proposed [5], [6]. Each of those has a maximum utilization λ_o with the property that the number of messages which have been generated but not yet successfully transmitted will be bounded with probability 1 as long as $\lambda t_1 < \lambda_o$. If $\lambda t_1 \geq \lambda_o$, the utilization of the channel is λ_o, but the expected message delay is infinite. The largest λ_o found to this day is .4877 [7].

Note that the definition of capacity given above is not the only one that has been proposed. Pippenger [8] defines capacity as the supremum of the λ's for which message delays can remain finite with probability one It is clear that the value of the capacity under this definition is not larger than under the original definition, and we conjecture that they are equal.

Pippenger [8] has shown that the capacity is bounded away from 1, in fact is not more than .744 ($t_o = t_1 = t_2$). He also generalized the model to include channels when the number of transmitted messages can be determined up to some maximum d, and has found a bound on the utilization that is strictly increasing function of d, converging to 1. Moreover he showed the existence of strategies achieving utilization arbitrary close to 1 when $d = \infty$.

This paper generalizes Pippenger's results. An upperbound on the capacity is derived for the case of different t_i's, its value is .704 when d=2 and $t_o = t_1 = t_2$, and it increases rather slowly with d.

Before proceeding with the precise description of the model and the derivation of the bound we will examine the implication of these results. First, an algorithm that is efficient for infinitely many sources will also be efficient for $M < \infty$ sources as long as the typical intergeneration time at a source (M/λ for symmetric systems) is longer than the typical message delay. In that case, each transmission at a source is independent of the previous one, and one might as well assume that all messages have distinct sources.

Secondly, the previously mentioned results show the existence of some number C, (Pippenger's Capacity) $.4877 \leq C \leq .704$ such that if $\lambda t_1 < C$, the average message delay can remain bounded no matter the value of M. However, if $\lambda t_1 > C$, the average message delay must increase with M. It is readily seen that the increase is linear for synchronous time division multiplexing and cyclic polling.

Determining the values of C and of the capacity remains a challenging proposition.

2. The Precise Model

To understand the following model, note that a conflict resolution protocol is a sequential decision process, thus it can be described as a tree. Every node corresponds to an "experiment", i.e., the transmission of messages. Branches correspond to outcomes, i.e., numbers of messages transmitted. Associated with each experiment is a set of times, typically a time

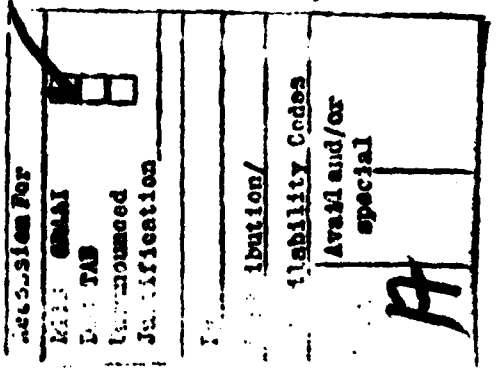
interval. Only those messages generated during the set corresponding to an experiment are transmitted when the experiment is made. Other conflict resolution algorithms rely on random choices, both ways are probabilistically equivalent when the generation times are Poisson.

A protocol for [0,T] is an infinite d-ary tree in which there is an initial node called the root, and in which each node k is connected by branches to offsprings k^i, i = 0,1,2,...d, that can be other nodes or leaves. Every node k is labelled with a measurable subset y(k) of [0,T]. $\mu(k)$ denotes the Lebesque measure of y(k) divided by T.

Let the random variables E denote a set of Poisson message arrival times in [0,T], with expected cardinality ν. The execution of a protocol with respect to a finite set E in [0,T] is a path through the tree defined as follows. Let k_o, the first node on the path, be the root and let E_o be E. Suppose that k_m and E_m have been determined, then $k_{m+1} = k_m^j$ where j is the minimum of d and the cardinality of $y(k_m) \cap E_m$; $E_{m+1} = E_m$ if $k_{m+1} = k_m^i$, and $E_{m+1} = E_m/y(k_m)$ otherwise. In other words, j is the number of non-transmitted messages whose arrival times are in $y(k_m)$, or d if there are more than d-1 such messages, and E_{m+1} is E_m minus any successfully transmitted message.

The set of nodes k in an execution ℓ that have offsprings $k^{(1)}$ is denoted by S_ℓ, the set of successful experiments in ℓ.

A protocol will be called valid if, for almost every subset E of [0,T], the execution ℓ of the protocol with respect to E terminates after finitely many steps with $E \subset \bigcup_{k \in S_\ell} y(k)$, i.e., if every message has been successfully transmitted.



The set of nodes k in an execution ℓ that have offsprings $k^{(0)}$ or $k^{(1)}$ is denoted by T_ℓ, the set of experiments in ℓ not resulting in collisions.

A valid protocol will be called <u>minimal</u> if for all executions ℓ, $y(k) \cap y(k') = \phi$, $k \neq k'$, $k, k' \in T_\ell$. Thus, in a minimal protocol, a subset of [0,T] is never tested again once it has been determined not to contain a message, or when the only message present has been successfully transmitted. Any valid protocol can be made minimal by iteratively changing the y(k)'s, starting from the root, so as to satisfy the null intersection property. The execution of the protocol with respect to a set E is not affected by the change.

The execution of a protocol is a random path through the tree. P(k) denotes the probability that node k is included in an execution, and q(k,i) denotes the conditional probability that $k^{(i)}$ follows k in the execution of a protocol.

The expected number of experiments, σ, in an execution of a protocol has value

$$\sigma = \sum_k P(k)$$

The expected fraction q_i of experiments resulting in outcome i is given by (assuming $\sigma < \infty$)

$$q_i = \frac{1}{\sigma} \sum_k P(k)\, q(k,i) \qquad (1)$$

Note that $\sum_{i=0}^{d} q_i = 1$.

For valid protocols

$$q_1 = \frac{\nu}{\sigma} \tag{2}$$

We will denote $(q_0, q_1, \ldots, q_d)$ by $\underline{q}$.

The efficiency e of a protocol is simply

$$e = \frac{\nu t_1}{\sigma \sum_{i=0}^{d} q_i t_i}$$

where $t_i \geq 0$ is the time it takes to observe outcome i. Note that for valid protocols

$$e = \frac{\nu t_1}{\nu t_1 + \sigma \sum_{i \neq 1} q_i t_i} = \frac{t_1}{t_1 + f}$$

where f is defined by $f = \frac{\sigma}{\nu} \sum_{i \neq 1} q_i t_i$ and can be thought of as the expected time overhead per message. Note that efficiencies close to 1 are achieved when $t_1 >> f$.

The previous relation between e and f allows us to lowerbound f (which does not depend on t_1) in order to upperbound e. This is the object of the next section.

3. Derivation of the Results

Our goal is to lowerbound f for any valid protocol.

As mentioned in the previous section, it is enough to consider minimal protocols. We will show that $\underline{q}$ lies in some closed convex region S of the unit simplex. The minimum over that region of f considered as a function of $\underline{q}$ will be our lower bound.

We first note that for any execution ℓ of a minimal protocol $\sum_{k \in T_\ell} \mu(k) \leq 1$. Averaging over ℓ yields

$$\sum_k P(k)\ \mu(k)(q(k,o) + q(k,1)) \leq 1 \qquad (3)$$

Next the entropy h (i.e., minus the mean of the log of the probabilities) of the executions of a protocol can be written

$$h = \sum_k P(k)\ H(\underline{q}(k)) \qquad (4)$$

where $H(\underline{q}(k)) = -\sum_{i=0}^{d} q(k,i) \log q(k,i)$.

The probability of an execution ℓ of a minimal protocol is no more than $e^{-\nu} \prod_{b \in S_\ell} \nu\mu(b)$, as one arrival must have occurred in every $y(b)$, $b \in S_\ell$, which are disjoint, and no arrival could have occurred outside such a subset.

Thus

$$h \geq E(-\log \prod_{k \in S_\ell} \mu(k)e^{-\nu}) = \nu \log e - \sum_k P(k)q(k,1)\log(\mu(k)) \qquad (5)$$

The right hand side of (5) is not less that

$$\nu \log e - \sum_k P(k)q(k,1) \log \frac{q(k,1)}{q(k,0) + q(k,1)}$$

as can be seen by using the inequality $\ln(x) \leq x-1$ and (3).

Subtracting this last expression from the right hand side of (4), dividing by σ and using (2) one obtains

$$\frac{1}{\sigma} \sum_k P(k)\, g(\underline{q}(k)) \geq 0,$$

where $g(\underline{x}) = - x_1 \log (x_o+x_1) - \sum_{\substack{i=0 \\ i\neq 1}}^{d} x_i \log(x_i) - x_i \log e$

It is shown in the appendix that g is a strictly concave function, thus by Jensen's inequality and (1),

$$g(\underline{q}) \geq 0$$

To obtain a lowerbound α on f, we find

$$\alpha = \min_{\underline{q}\in S} f = \min_{\underline{q}\in S} \frac{\sum_{i\neq 1} q_i t_i}{q_1}$$

where $S = \{\underline{q} \in R^{d+1},\ q_i \geq 0,\ \sum q_i = 1,\ g(\underline{q}) \geq 0\}$

α is finite, as S contains the point $(\frac{1}{d+1},\ldots,\frac{1}{d+1})$. Note that $\underline{q}$ achieving the above minimum also achieves

$$0 = \min_{\underline{q}\in S} \left(\sum_{i\neq 1} q_i t_i - q_1 \alpha \right) \qquad (6)$$

A q^* in S minimizing (6) must satisfy $g(q^*) = 0$, as if $g(q^*) > 0$, then $q_1^* \neq 1$ and one can decrease the objective function by increasing q_1^* and decreasing some q_i^* $(i \neq 1)$ by the same amount. A solution exists as S is compact in $\mathbb{R}^{d+1}$ and the objective function (6) is continuous. The solution is unique as g is strictly concave. The problem of minimizing (6) subject to equality constraints is standard. Necessary and sufficient conditions for the optimality of $\underline{q}^*$ are

$$g(\underline{q}^*) = 0$$

$$\sum q_i^* = 1 \tag{7}$$

$$\ln q_o + \frac{q_1}{q_o+q_1} + \mu t_o + (\lambda + 1) = 0 \tag{8}$$

$$\ln (q_o+q_1) + \frac{q_1}{q_o+q_1} - \mu\alpha + (\lambda + 1) = 0 \tag{9}$$

$$\ln q_i - \mu t_i + (\lambda + 1) = 0 \quad i = 2,3\ldots d. \tag{10}$$

$$\sum_{i \neq 1} t_i q_i - \alpha q_1 = 0 \tag{11}$$

where λ and μ are Lagrange multipliers.

One checks that $g(\underline{q}^*) = 0$ implies $\lambda+1 = 0$. To find $\underline{q}^*$ and μ, one must proceed numerically. An iterative way is to first guess a value of μ, thus determining q_i^* $2 \leq i \leq d$ by (10). $q_o^*+q_1^*$ is obtained from (7), then q_o^* and q_1^* from (8). If (9) and (11) are not verified, the value of

u should be changed and the process repeated. Figure 1 showing the upper bound on the overhead when d=2, as a function of t_1/t_2 was so obtained.

The solution is simpler when $t_i = t$ $(i = 0,2,\ldots,d)$. Defining $L = \frac{q_1}{q_0+q_1}$, one obtains from (8) and (10) that $q_i = q_0 e^{L}$ $2 \leq i \leq d$, and from (7) that $q_0 = (\frac{1}{1-L} + (d-1)\, e^{L})^{-1}$. Subtracting (8) from (9) yields $\log (1-L) + u(t+\alpha) = 0$. As from (8) and (11), $u(t+\alpha) = \frac{1-L}{L\, q_0}(-L-\ln q_0)$, L satisfies the equation:

$$\ln (1-L) = \frac{(1+(1-L)(d-1)e^{L})(L + \ln(1-L) - \ln(1+(1-L)(d-1)e^{L}))}{L} \qquad (12)$$

Once L is computed, one can find q_1^*, which here is an upperbound on the relative frequency of experiments resulting in a success, by the formula $q_1^* = \frac{L}{1+(1-L)(d-1)e^{L}} \cdot \frac{t(1-q_1^*)}{q_1^*}$ is equal to α, our lowerbound on the time overhead per message, f. Numerical results appear in Table 1. One sees that they are not very sensitive to d. Indeed one can derive from (12) that $1-L = \frac{1}{e(d-1)\ln d}\ (1 + 0\ (\frac{1}{(\ln d)^2}))$,

where $(\ln d)^2\ 0\ (\frac{1}{\ln\ (d))^2})$ is bounded for $d \geq 1$, and consequently

$q_1^* = \frac{\ln d}{1 + \ln d}\ (1 + 0\ ((\ln(d))^{-2})$. This result indicates that determining how many messages are involved in a collision does not greatly pay off.

d	α	q_1^*
2	.4198	.7043
3	.3164	.7596
4	.2753	.7841
5	.2518	.7988
10	.2029	.8313
100	.1298	.8851
1000	.0972	.9114
10000	.0781	.9275

TABLE 1

Lowerbound α on the time overhead per message and upperbound q_1^* on the relative frequency of success as d varies ($t_i = 1$, $i = 0,2,\ldots,d$)

Appendix

We will show here that the function

$$g(\underline{x}) = - x_o \log x_o - x_1 \log (x_o+x_1) - \sum_{i=2}^{d} x_i \log x_i - x_1 \log e$$

is strictly concave for $\underline{x} \geq 0$. By inspection, this is immediately true for all terms, except the second one. We prove now that $- x_o \ln x_o - x_1 \ln (x_o+x_1)$ is strictly concave by showing that the matrix of second partial derivatives is negative definite. This matrix is equal to

$$\begin{pmatrix} \dfrac{-(x_o^2+x_o x_1+x_1^2)}{x_o(x_o+x_1)^2} & \dfrac{-x_o}{(x_o+x_1)^2} \\ \dfrac{-x_o}{(x_o+x_1)^2} & \dfrac{-(2x_o+x_1)}{(x_o+x_1)^2} \end{pmatrix}$$

The upper diagonal term is always negative, while the determinant, $\dfrac{1}{x_o(x_o+x_1)}$, is always positive. Thus Sylvester's test is verified.

References

[1] Roberts, L.G., "Aloha Packet System With and Without Slots and Capture," ASS, Note 8, June 26, 1972.

[2] Kleinrock, L. and Tobagi, F., "Carrier Sense Multiple Access for Packet Switched Radio Channels", Proc. Int. Conf. Communications, Minneapolis, Minn., June 1974.

[3] Metcalfe, R.M. and Boggs, D.R., "Ethernet: Distributed Packet Switching for Local Computer Networks", CACM, July, 1976.

[4] Roberts, L.G., "Dynamic Allocation of Satellite Capacity Through Packet Reservation", 1973 Nat. Comput. Conf., AFIPS Conf. Proc., Vol. 42.

[5] Capetanakis, J., "The Multiple Access Broadcast Channel: Protocol and Capacity Considerations", Ph.D. Thesis, Dept. of Electrical Engineering and Computer Science, M.I.T., 1977.

[6] Gallager, R.G., "Conflict Resolution in Random Access Broadcast Networks", Proceedings AFOSR Workshop in Communication Theory and Applications, Sept. 17-20, 1978, Provincetown, Mass. pp. 74-76.

[7] Mosely, J., "An Efficient Contention Resolution Algorithm for Multiple Access Channels", Report LIDS-TH-918, Laboratory for Information and Decisions Systems, M.I.T., May 1979.

[8] Pippenger, N., "Bounds on the Performance of Protocols for a Multiple Access Broadcast Channel", Report RC 7742, Math. Science Dept., IBM Thomas J. Watson Research Center, Yorktown Heights, N.Y., June 1979.

3 %
2.5
2
1.5
1
.5
.42
1.05
.2
.4
.6
.8
1.0