TR-86-45

ON THE STABILITY OF INTERACTING QUEUES IN A MULTIPLE ACCESS SYSTEM

by

R. Rao and A. Ephremides

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **1986**	2. REPORT TYPE	3. DATES COVERED **00-00-1986 to 00-00-1986**
4. TITLE AND SUBTITLE **On the STability of Interacting Queues in a Multiple Access System**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Maryland,Electrical Engineering Department,College Park,MD,20742**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT **see report**		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**32**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

ON THE STABILITY OF INTERACTING QUEUES IN A MULTIPLE ACCESS SYSTEM

by

R. Rao

University of California, San Diego

and

A. Ephremides

University of Maryland

Abstract

We consider the standard discrete-time slotted ALOHA system with a finite number of buffered terminals. The stability (ergodicity) region for this system is known for the case of two terminals and for the case of infinite, but symmetric, terminals. In this paper we introduce a new approach of studying the stability of this system by means of a simple concept of dominance. As a result we show that the stability region for the case of two terminals can be obtained in a very simple way. Furthermore, we obtain lower (inner) bounds for the stability region of the system with an arbitrary finite number of terminals. These bounds are superior to the ones already known. Finally we point out a similarity betwen these stability results and the achievable region of the no-feedback collision channel that may suggest a connection between the two problems.

1. Introduction

Interacting queueing systems occur naturally in multiple access channel models and shared computer processor systems. As is the case with most non-standard queueing systems, these interacting queues are difficult to analyze. Their study has received much attention lately owing to their importance in applications, as well as to their theoretical interest. In [1] Fayolle and Iasnogordski displayed the inherent difficulty of the analysis of such systems. In [2] the importance of interfering queues in multiple access systems was recognized. In [3] Saadawi and Ephremides introduced an approximate model for the analysis of the slotted ALOHA system. In [4,5] Sidi and Segall introduced different approximations that led to an exact analysis of a simple 2-user system. Szpankowski in [6,7] has considered the ergodicity region of the slotted ALOHA system and obtained lower bounds. Tsybakov and Mikhailov [8] obtained sufficient or necessary conditions for the stability of this system for the case of M users and, based

*This work was supported in part by ONR Grant N00014-84K-614 and by NSF Grant 85-00108.

on a result by Malyshev [9], were able to obtain both necessary and sufficient conditions for ergodicity. Finally, in a number of papers that have appeared in the literature since the early seventies it has been established that in the classical, bufferless, symmetric, infinite-user ALOHA system the stability threshold is e^{-1} [10,11,12].

In this paper we consider the case of discrete-time, slotted ALOHA system with M users each of which has a buffer of infinite capacity to store incoming packets. The assumptions of the model are the usual ones. Time is slotted and the transmission time of a packet is equal to one slot. Each user receives (generates) packets according to a Bernoulli process. The rate of arrivals is λ_i for the ith user. Arrivals to different users are independent. In each slot user i attempts to transmit the head-of-the-line packet with probability p_i, provided the buffer is not empty. Based on instantaneous ternary feedback (collison, idle, success) each user determines the outcome of the attempted transmission. Simultaneous transmission attempts by two or more users result in collision.

We are interested in determining the region of values of the arrival rates $\lambda_i, i = 1, ..., M$ for which this system of M queues is stable.* We introduce a new approach in the study of the stability question. This approach consists of considering hypothetical, auxiliary systems of queues that closely parallel the operation of the system of interest but dominate it in a well-defined sense. As a result, we are able to determine the stability region for the case of M=2, that was obtained in [8], in a very simple way that illuminates the relationship and interaction between the two queues. This is done in Section 2. Furthermore, by using the same approach in a slightly more elaborate form we are able to obtain improved sufficient conditions for the stability of the system of M queues, for any finite M. This is done in stages. In Section 3 a first set of bounds is obtained from the direct extension of the dominance concept as used in Section 2 and in section 4 the more elaborate set of dominating systems is used to lead to the derivation of the final set of bounds. Finally in section 5 we relate these results to the achievable region of the collision channel without feedback [13,14]. Some of the technical proofs have been moved to appendices for the sake of clarity and continuity in the main text.

2. The Dominant System for the Case of 2 Queues

Let us denote by S the system of queues that is the object of our study as described in

* the precise definition of stability is provided in the next section

the introduction. Let Q_i denote the queue size at the ith terminal, $i = 1, 2$. Clearly, the queue sizes evolve as a 2-dimensional random walk in the first quadrant as shown in Fig. 1. The transition probabilities of the underlying Markov chains can be easily obtained, based on the rules of operation of the ALOHA system. In fact this chain is irreducible and aperiodic. What makes the analysis difficult is of course the infinite size of the state space and the fact that the probabilities of transition on the boundaries are different from those of their counterparts in the interior of the quadrant. This is shown, for example, in the figure for the transition $(Q_1, Q_2) \rightarrow (Q_1 - 1, Q_2)$. If these transition probabilities were identical, regardless of whether $Q_2 = 0$, the system of two queues would decompose nicely* and each queue could be separately analysed. This observation motivates the introduction of an auxiliary hypothetical system that has precisely this property. Consider a new system, S^2, consisting of copies of the two queues of S, with the following properties:

1) arrivals at queue i in the new system occur at exactly the same instants as in the original system, $i = 1, 2$;

2) the "coin-tosses" that determine transmission attempts at queue $i, i = 1, 2$, have exactly the same outcomes in both systems;

3) whenever $Q_i = 0, i = 1, 2$, terminal i continues to transmit "dummy" packets with the same probability p_i, thus continuing its interference with the other terminal, whether it is empty or not.

It is clear that the queue sizes, at either terminal, in the new system will never be less than their counterparts in the original system, provided they start from identical initial conditions in both systems. In the new system S^2 the "service" rate seen by, say, terminal 1 is always equal to $p_1(1 - p_2)$ while in the original system it oscillates between the values of p_1 and $p_1(1 - p_2)$ depending on whether terminal 2 is empty or not.

Definition. The stability region of either system is the set of values of $\lambda_i, i = 1, 2$ for which the underlying irreducible, aperiodic Markov chain (Q_1, Q_2) is ergodic.

Clearly, the strong stochastic dominance of S^2 on S implies that the stability region of S^2 inner-bounds that of S. The stability region of S^2 is easy to determine. Each terminal

* Note, however, that the queues would still be statistically dependent since the transition $(Q_1, Q_2) \rightarrow (Q_1 - 1, Q_2 - 1)$ has zero probability.

operates as a discrete $M \mid M \mid 1$ system and is therefore stable if and only if

$$\lambda_i < p_i \prod_{j \neq i} (1 - p_j)$$

Thus the stability region of S^2 is the rectangle shown in Fig. 2. This is of course a poor bound for the stability region of S because S^2 dominates the original system too strongly. We would like, to draw attention to one observation before proceeding any further. Note that each queue behaves identically in both systems so long as neither queue empties. Thus if both systems get started with identical, non-zero queue sizes, they will evolve following exactly identical trajectories at least until one of the queues empties. This observation plays a crucial role in some of the proof arguments that will follow.

Now, let us consider another hypothetical, auxiliary system S^1, in the hope of obtaining a dominant system that does not dominate the original system as strongly and that can lead, therefore, to a closer tracking of the original system's behavior and to a better bound of its ergodicity region. The system S^1 is identical to both S and S^2 with respect to arrivals and attempted transmissions (that is, it has properties (1) and (2)) but differs with respect to the behavior when the queues empty. In this system queue 2 behaves as in S while queue 1 behaves as in S^2; that is, only queue 1 continues the irresponsible transmission of "dummy" packets when it empties. As a result, queue 1 sees a "service" rate that oscillates between p_1 and $p_1(1 - p_2)$ depending on whether queue 2 is empty or not, while queue 2 always sees a worst-case service rate of $p_2(1 - p_1)$ regardless of the status of queue 1. It is clear that S^1 dominates S since either queue will have a successful departure in S whenever it has one in S^1, but not necessarily vice-versa. It is also clear that in S^1 the stability of queue 2 is easily determined by the fact that it operates as a discrete-time $M \mid M \mid 1$ system. Thus it will be stable iff

$$\lambda_2 < p_2(1 - p_1) \tag{1}$$

Let us therefore assume that (1) is satisfied and let us determine the criterion for the stability of queue 1. Terminal 1 sees a "service" rate that has the value p_1 when terminal 2 is empty (and that happens with probability $1 - \frac{\lambda_2}{p_2(1-p_1)}$ according to the well-known $M \mid M \mid 1$

formula) and the value $p_1(1-p_2)$ when terminal 2 is non-empty (which occurs with probability $\frac{\lambda_2}{p_2(1-p_1)}$). In appendix it is shown that a necessary and sufficient condition for the stability of queue 1 is that the arrival rate λ_1 be less than its average service rate, thus extending the well-known result for a $G \mid G \mid 1$ queue [15]. Applying this result here yields

$$\lambda_1 < p_1(1 - \frac{\lambda_2}{p_2(1-p_1)}) + p_1(1-p_2)\frac{\lambda_2}{p_2(1-p_1)}$$

which can be written as

$$\lambda_1 < p_1(1 - \frac{\lambda_2}{1-p_1}) \tag{2}$$

Thus the stability region of S^1 is as shown in Figure 3.

What is important is that the boundary of the stability region given by Eq. (2) is not only a bound for the stability region of S but coincides with it. The proof relies on the observation made earlier that, so long as the queues do not empty, systems S and S^1 are indistinguishable. Here is the argument: Given that $\lambda_2 < p_2(1-p_1)$, if for some λ_1 queue 1 is stable in S^1, it is also stable in S by virtue of the dominance. Conversely, if for some λ_1 queue 1 is unstable in S^1, then the queue size $Q_1(n)$ will grow to ∞ without emptying with finite (non-zero) probability.* Thus, not all sample paths of $Q_1(n)$ correspond to transient behavior with infinitely often visits to 0. However, so long as queue 1 does not empty, S^1 and S behave identically if started from the same initial conditions. Thus the same sample paths that go to ∞ without visiting 0 belong to the evolution of queue 1 in system S. Therefore queue 1 is unstable in system S as well.

Notice that by reversing the roles of the two queues in system S^1, that is by assuming that queue 2 is the one that transmits dummy packets when it empties, we obtain a stability region given by

$$\lambda_1 < p_1(1-p_2)$$

* This follows from the fact that queue 2 is stable in its own right and decoupled from queue 1; nevertheless an independent proof can be supplied as in [16].

and

$$\lambda_2 < p_2(1 - \frac{\lambda_1}{1 - p_2}) \tag{3}$$

The branch of the boundary that corresponds to Eq. (3) can be shown to be part of the stability region of S in a similar fashion. Thus the union of the branches corresponding to Eqs. (2) and (3) defines the stability region for S as shown in Fig. 4. This is identical to the region obtained by Tsybakov and Mikhailov in [8] through use of Malyshev's lemma. Note that by taking the envelope of these regions as p_1 and p_2 vary in $[0, 1]$ we obtain the curve C, shown in Fig. 4, which is analytically described by

$$\sqrt{\lambda_1} + \sqrt{\lambda_2} = 1 \tag{4}$$

or parametrically, by

$$\lambda_1 = p_1(1 - p_2)$$

$$\lambda_2 = p_2(1 - p_1) \tag{5}$$

$$p_1 + p_2 = 1$$

The idea of this simple derivation as well as the choice of dominant systems was first briefly reported in [16].

3. The Dominant System for the case of M queues

Now consider the case of $M > 2$. The approach described in section 2 is insufficient to yield the ergodicity region of the system. It does help however in producing bounds that are superior to the ones known to date. In this section we will develop this dominance approach and show how these bounds can be obtained.

Let

$Q_i(n)$ = queue size at terminal i at time n

$D_i(n)$ = number of departures from terminal i at time $n, i = 1, ..., M, n = 1, 2, ...$

and let

$$P_{i,k}^{(n)} = Pr[Q_i(n) = k] \tag{6}$$

$$\mu_i(n) = Pr[D_i(n) = 1] \tag{7}$$

Thus $\mu_i(n)$ is the "service" rate at terminal i at time n. It clearly depends on the status of the queues at the different terminals and assumes values between p_i, when all other queues are empty and $p_i \prod_{j \neq i}(1 - p_j)$, when none of the other queues are empty. Let $\tilde{\mu}_i(n)$ denote the average value of $\mu_i(n)$. We further define two types of events of interest:

$$\xi_k(n) = \{Q_k(n) = 0\}, \text{ i.e. the event that at time n the kth queue is empty} \tag{8}$$

and

$$E_{k_1 k_2 ... k_e}(n) = \bigcap_{i \in \{k_1, ..., k_e\}} \xi_i(n) \bigcap_{i \notin \{k_1, ..., k_e\}} \overline{\xi}_i(n) \tag{9}$$

i.e. the event that at time n the only empty queues are the ones corresponding to terminals $k_1, ..., k_e$. Also note that the family of events E forms a partition of the sample space. Furthermore we have

$$\xi_k(n) = E_k(n) \bigcup \{ \bigcup_{k_1 \neq k} E_{kk_1}(n)\} \bigcup ... \bigcup \{ \bigcup_{\substack{k_1, ..., k_e \\ \text{all distinct}}} E_{kk_1 ... k_e}(n)\} \bigcup ... \bigcup E_{1,2,3,...,k,...M}(n), \tag{10}$$

In Appendix 1 it is established that the limits of $\mu_i(n)$ as $n \to \infty$ exist; we denote them by μ_i and their average values by $\tilde{\mu}_i$. Let us consider also the system S^1, defined in a manner similar to that of the preceeding section. Namely, S^1 is identical to S in all respects (arrival instants, transmission attempts, etc.) except that terminal 1 continues to transmit "dummy" packets, even when it is empty, with the same probability p_1 as when it is not empty. All other

terminals behave as in S. We define the quantities Q^1, D^1, P^1, and μ^1 as above but with the use of the superscript "1" to denote reference to system S^1. Similarly we also consider for reference purposes the system S^M, in which all terminals continue to transmit when empty. This system is the generalization of S^2 as defined in Section 2.

Our goal is to derive sufficient conditions for the stabilityof S^1 and to conclude, by virtue of the dominance relationship, that these are also sufficient conditions for the stability of S. We shall focus on queue 1. To "decouple" its stability status from that of the other queues we need to assume that the other queues are stable. Unfortunately, unlike the case of $M = 2$, we do not know what is the stability criterion for the set of queues $2, ..., M$ and thus we are forced to assume a "strong" stability status for them, namely that

$$\lambda_i < \mu_i^M \triangleq p_i \prod_{j \neq i} (1 - p_j), \qquad i = 2, ..., M \tag{11}$$

The validity of Eq. (10) implies stability of this group of queues under system S^M and thus under S^1 and S as well.

Let $B_1^1(n)$ denote the event that in system S^1, terminal 1 at time n successfully transmits a packet, whether real of "dummy"; its probability is $\mu_1^1(n)$. We shall presently show that a sufficient condition for the stability of queue 2 in S^1 is

$$\lambda_1 < \phi(\lambda_2, ..., \lambda_M; p_1, ..., p_M) \tag{12}$$

where ϕ is a specific function of the other arrival rates and all retransmission probabilities, that lower-bounds the asymptotic service rate μ_1^1 of terminal 1. In Appendix 2 we show that this bound is already an improvement over these derived in [6-8]. We present the results in the form of two lemmas that lead to a theorem. The first lemma lower-bounds $\mu_1^1(n)$, the second lemma establishes the bound as $n \to \infty$, and the theorem establishes the sufficient condition by making use of the results of Appendix 1.

Lemma 3.1

In the system S^1, the probability of a successful transmission by queue 1 at the instant n satisfies the following lower bound inequality for all $n \geq 1$:

$$\mu_1^1(n) \geq p_1 \prod_{j=2}^{M}(1-p_j) + p_1 \sum_{j=2}^{M} Pr[\xi_i^1(n)] \cdot p_i \prod_{\substack{j=2 \\ j \neq i}}^{M}(1-p_j) \tag{13}$$

Proof

By definition:

$$\mu_1^1(n) = Pr[B_1^1(n)] \tag{14}$$

We apply the theorem of total probability to Eq. (14) conditioning the event $B_1^1(n)$ on the 2^{M-1} mutually exclusive events of Eq. (9). This results in the following equation:

$$\mu_1^1(n) = Pr[B_1^1(n)] = Pr[B_1^1(n)/\{E_\phi^1(n)] \cdot P_r[E_\phi^1(n)] +$$

$$+ \sum_{k_1=2}^{M} Pr[B_1^1(n)/E_{k_1}^1(n)] \cdot Pr[E_{k_1}^1(n)] + \ldots$$

$$\ldots + \sum_{\substack{k_1,\ldots,k_e \in \{2,\ldots,M\} \\ \text{all } k_i \text{ distinct}}} Pr[B_1^1(n)/E_{k_1\ldots k_e}^1(n)] \cdot Pr[E_{k_1,\ldots,k_e}(n)] + \ldots$$

$$+ Pr[B_1^1(n) \mid E_{2,\ldots,M}^1(n)] Pr[E_{2,\ldots,M}^1(n)] \tag{15}$$

The following further decomposition is now possible.

$$Pr[B_1^1(n)/E_{k_1,\ldots,k_e}^1(n)]$$

$$= Pr[B_1^1(n)/\{Q_1^1(n) > 0\}, E_{k_1,\ldots,k_e}^1(n)] \cdot Pr[\{Q_1^1(n) > 0\}/E_{k_1,\ldots,k_e}^1(n)] +$$

$$+ Pr[B_1^1(n)/\{Q_1^1(n) = 0\}, E_{k_1,\ldots,k_e}^1(n)] \cdot Pr[\{Q_1^1(n) = 0\}/E_{k_1,\ldots,k_e}^1(n)] \tag{16}$$

Now, if queue 1 is non-empty and of the remaining ones only queues k_1 through k_e are empty then the probability of a successful transmission by terminal 1 is $p_1 \prod_{i \neq \{k_1,\ldots,k_e\}}(1-p_i)$. Note that in this case the transmitted packet is a real information packet. On the other hand, if queue 1 is empty and the status of the other queues is as before then the probability of a successful transmission is still given by $p_1 \prod_{i \notin \{k_1,\ldots,k_e\}}(1-p_i)$. The difference is that in this case the transmitted packet is a dummy packet. However, we are not concerned yet with this distinction since at this stage we seek a lower bound for the probability of successful transmission of information or dummy packets $\mu_1^1(n)$. Besides, near saturation, all packets tend to be information packets since, then, the queues are rarely empty.

Based on the above comments we deduce that

$$P_r[B_1^1(n)/E_{k_1,\ldots k_e}^1(n)] = p_1 \prod_{i \notin \{k_1,k_2,\ldots k_e\}} (1-p_i) \tag{17}$$

Now, note that the following inequalities can be easily verified for any set of indices $k_1, k_2, \ldots k_e$:

$$1 = \prod_{k \epsilon \{k_1,k_2,\ldots k_e\}} (p_k + (1-p_k)) = (p_{k_1} + 1 - p_{k_1}) \prod_{k \epsilon \{k_2,k_3,\ldots k_e\}} (p_k + (1-p_k)) =$$

$$= [p_{k_1} + (1-p_{k_1})] \prod_{k \epsilon \{k_2,k_3,\ldots k_e\}} (1-p_k)(1+\frac{p_k}{1-p_k}) \geq p_{k_1} \prod_{k \epsilon \{k_2,k_3,\ldots k_e\}} (1-p_k) +$$

$$+ \prod_{k \epsilon \{k_1,k_2,\ldots k_e\}} (1-p_k) \tag{18}$$

$$1 = \prod_{k=2}^{M} [p_k + (1-p_k)] \geq \prod_{k=2}^{M} (1-p_k) + \sum_{k_1=2}^{M} p_{k_1} \prod_{\substack{k=2 \\ k \neq k_1}}^{M} (1-p_k) \tag{19}$$

Using eqns. (17), (18) and (19), we rewrite Eq. (15) as follows:

$$\mu_1^1(n) \geq p_1 \prod_{k=2}^{M}(1-p_k) \cdot Pr[E_\phi^1(n)] +$$

$$+ \sum_{k_1=2}^{M} [\{p_1 \prod_{k=2}^{M}(1-p_k)\} + \{p_1 p_{k_1} \prod_{\substack{k=2 \\ k \neq k_1}}^{M}(1-p_k)\}] \cdot Pr[E_{k_1}^1(n)] + \ldots$$

$$\ldots + \sum_{\substack{\{k_1,k_2,\ldots k_e\} \subset \{2,3,\ldots M\} \\ \text{all } k_i \text{ distinct}}} [\{p_1 \prod_{k=2}^{M}(1-p_k)\} + \{p_1 p_{k_1} \prod_{\substack{k=2 \\ k \neq k_1}}^{M}(1-p(k)\}] \cdot Pr[E_{k_1,k_2,\ldots k_e}^1(n)] + \ldots$$

$$\ldots + [\{p_1 \prod_{k=2}^{M}(1-p_k)\} + \{\sum_{k_1=2}^{M} p_1 p_{k_1} \prod_{\substack{k=2 \\ k \neq k_1}}^{M}(1-p_k)\}] Pr[E_{2,3,\ldots M}^1(n)] \tag{20}$$

Rearranging Eq. (20) and using eqs. (9) and (10) we obtain the following inequality

$$\mu_1^1(n) \geq p_1 \prod_{j=2}^{M}(1-p_j) + \sum_{i=2}^{M} p_1 \{p_i \prod_{\substack{j=2 \\ j \neq i}}^{M}(1-p_j)\} Pr[\xi_i^1(n)] \tag{21}$$

This concludes the derivation.

In the next result, we derive the asymptotic value of the bound of Lemma 3.1. The bound is calculated on the assumption that queues 2 through M are stable.

Lemma 3.2

If

$$\lambda_i < \mu_i^M = p_i \prod_{j \neq i}(1-p_j), \qquad i = 1, \ldots, M$$

then

$$\lim_{n \to \infty} \mu_1^1(n) \geq p_1 \prod_{j=2}^{M}(1-p_j) + \sum_{i=2}^{M} p_1 p_i (1 - \frac{\lambda_i}{\mu_i^M}) \prod_{\substack{j=2 \\ j \neq 1}}^{M}(1-p_j)$$

Proof:

It is sufficient to show that

$$\lim_{n\to\infty} Pr[\xi_i^1(n)] \geq 1 - \frac{\lambda_i}{\mu_i^M} \qquad 2 \leq i \leq M$$

Note that by definition

$$Pr[\xi_i^1(n)] = Pr[Q_i^1(n) = 0]$$

Since S^M dominates S^1 we know that

$$P[Q_i^1(n) = 0] \geq Pr[Q_i^M(n) = 0]$$

Therefore, taking limits and using the decoupling properties of S^M we obtain

$$\lim_{n\to\infty} Pr[\xi_i^1(n)] \geq \lim_{n\to\infty} Pr[Q_i^M(n) = 0] = (1 - \frac{\lambda_i}{\mu_i^M}) \tag{22}$$

Taking limits on eq. (21) and employing the inequality (22) we obtain the desired result, namely

$$\lim_{n\to\infty} \mu_1^1(n) \leq p_i \prod_{j=2}^{M}(1-p_j) + p_1 \sum_{i=2}^{M}(1 - \frac{\lambda_i}{\mu_i^M})(1-p_j) = \mu_1^1 \tag{23}$$

The existence of these limits is again justified in Appendix 1.

Now we proceed to establish the main theorem.

Theorem 3.1

The system S is stable if

$$\lambda_1 < p_1 \prod_{j=2}^{M}(1-p_j) + p_1 \sum_{j=2}^{M} p_j(1 - \frac{\lambda_j}{\mu_j^M}) \prod_{\substack{k=2 \\ k\neq j}}^{M}(1-p_k) \tag{24}$$

and

$$\lambda_j < \mu_j^M, \qquad 2 \le j \le M \tag{25}$$

where

$$\mu_j^M = p_j \prod_{\substack{k=1 \\ k \ne j}}^{M} (1 - p_k)$$

Proof

From Lemma 3.2 we know that the R.H.S. of Eq. (24) is a lower bound for μ_1^1, hence $\lambda_1 < \mu_1^1$ and $\lambda_1 < \tilde{\mu}_1^1$. According to the results in Appendix 1, this is sufficient to ensure that

$$Pr[Q_1^1 = 0] > 0$$

The inequalities (25) are also sufficient to imply that

$$Pr[Q_i^1 = 0] > 0, \qquad 2 \le i \le M$$

Thus under the hypothesis of the theorem the system S^1 is stable and so is S since it is dominated by S^1.

Comments

1. Recall that in System S^1 we required that terminal 1 transmit dummy packets when empty; the choice of terminal was clearly arbitrary. If we interchange the roles of terminal 1 and terminal i and repeat the proof we will derive a similar result for a different ordering of the terminals. In fact the union of the regions that are obtained in this manner for all possible orderings produces a better bound.
2. Szpankowski's results on the stability of system S [6,7] are of similar nature. In appendix 2 we show that the stability region established in [6,7] is a strict subset of the region obtained here. We should like to record that Szpankowski's results apply to systems for which the arrival rates are not necessarily Bernoulli, and in that sense his results are more general. Nevertheless, we believe that our results can also be easily generalized to accomodate non-Bernoulli arrival processes because

the dominating systems S^1 and S^M, crucial to our analysis, retain their essential properties for arbitrary arrival processes.

3. The bound obtained in Theorem 3.1 is not satisfactory. As shown in Fig. 5, it fails to provide values for regions of values of the λ_i's that are outside the strict strong stability ones implied by Eq. (25). These regions correspond to the areas marked by question marks in Fig. 5. Also it is clearly not very tight since on the λ_i-axes the actual stability surface intersects at the points p_1, p_2, p_3 etc. In the case of $M = 3$, as shown in Fig. 5, the bound is below these values by an amount equal to $p_1 p_2 p_3$.

4. A Series of Dominant Systems that Yield Improved Bounds

In the preceeding section we obtained an inner bound to the stability region of S by essentially using the same approach as in section 2 in which we were successful in determining exactly the stability region for the 2-user system. In this section we refine our approach by considering a series of dominant systems that are able to track the behavior of S a little closer then S^1 could. These systems permit us to by-pass the shortcomings of the bound of Theorem 3.1 and to relax the sufficient conditions for the stability of S.

Retaining the notation we have used so far we introduce now an extended set of systems. For $j = 1, 2, ..., M$ we define S^j as follows:

1) Arrivals at the ith queue of $S^j, i = 1, ..., M$, are identical to those at the ith queue of S.

2) "Coin tosses" that determine transmission attempts at the ith queue of $S^j, i = 1, ..., M$, have identical outcomes to those at the ith queue of S.

3) For $i > j$, terminal i behaves exactly as in S, that is it does not attempt to transmit "dummy" packets when empty.

4) For $i \leq j$, terminal i attempts to transmit "dummy" packets when empty according to the following rules: with the aide of a "genie", terminal i is informed whether any terminal k, with $k < i$, will attempt a transmission in the slot; if yes, terminal i refrains from attempting to transmit; if no, it attempts to transmit a "dummy" packet with probability p_i.

As a result we have a series of systems in each of which terminal 1 is afforded the first choice to transmit a "dummy" packet, if empty. If it fails to transmit, "dummy" or real

packet, terminal 2, if empty, is afforded the same choice. If it doesn't transmit, terminal 3 is given the chance, and so on, up to terminal j in the system S^j. Terminals with identity number greater than j do not attempt to transmit "dummy" packets in S^j.

As a result, this series of systems has the following properties.

i) S^j dominates $S^{j-1}, j = 2, ..., M$, and, of source, S^1 dominates S. Note that S^1 coincides with the system S^1 considered in section 3 and so does S^M with its counterpart in that section. The reason for the dominance is this: Terminal j in S^{j-1} is always silent when empty, while it may transmit a "dummy" packet in S^j. Furthermore, all other terminals in S^{j-1} will have a successful transmission whenever they do have one in S^j, but not vice-versa. Thus, it is clear that

$$Q_i^j(n) \geq Q_i^{j-1}(n) \geq Q_i(n), \qquad n = 1, 2, ...; j = 2, ..., M; i = 1, ..., M$$

ii) The joint queue sizes of the M users in each of these systems evolve according to an M-dimensional, aperiodic, irreducible Markov chain.

iii) In system S^j the cummulative contention from terminals 1,2,...,j for the use of the channel remains fixed and equal to

$$1 - \prod_{i=1}^{j} (1 - p_i)$$

irrespective of whether any of these queues are empty. Thus any terminal $k, k > j$, faces competition from the group of the first j terminals that stays always fixed. This is ensured by the rules of transmission explained above in (4).

We present the derivation of the new bounds in the form of some lemmas that lead to a theorem. We start by introducing some notation similar to that of Eqs. (8), (9), and (10).

For any set of distinct indices $k_i, (j+1) \leq k_i \leq M$, let

$$\xi^j_{k_1}(n) = \{Q^j_{k_1}(n) = 0\}$$

$$E^j_{k_1}(n) = \{Q^j_{k_1}(n) = 0\} \bigcup_{\substack{i=1+j \\ i \neq k_1 2}}^{M} \{Q^j_i(n) > 0\}$$

$$E^i_{k_1,k_2,\ldots k_e}(n) = \bigcap_{i \epsilon \{k_1,k_2,\ldots k_e\}} \{Q^j_i(n) = 0\} \bigcap_{\substack{i \epsilon \{j+1,\ldots M\} \\ i \notin \{k_1,k_2,\ldots k_e\}}} \{Q^j_i(n) > 0\} \tag{26}$$

$$E^j_{\phi}(n) = \bigcap_{i=j+1}^{M} \{Q^j_i(n) > 0\}$$

Thus $\xi^j_{k_1}(n)$ is the event that the k_1th queue of S^j is empty at time n. $E^j_{k_1}(n)$ is the event that only the k_1th queue of S^j in the group from $j+1$ to M is empty at time n (all others being non-empty). $E^j_{k_1,k_2,\ldots k_e}(n)$ is the event that only queues $k_1, k_2, \ldots k_e$ are empty at time n and all the others in the group are non-empty. Since the group of terminals from 1 to j, collectively, present the same behavior in competing for the channel, regardless of the status of their queues, there is no interest in whether they are empty or not. Thus we consider only the events that pertain to the group $\{j+1, \ldots, M\}$.

There are $\binom{M-j}{1}$ events of the type $E^j_{k_1}(n)$ and $\binom{M-j}{e}$ events of the type $E^j_{k_1,k_2,\ldots k_e}(n)$. The total number of events is 2^{M-j}, and they are all mutually exclusive. Their union exhausts the sample space. The event $\xi^j_{k_1}$ may be decomposed into mutually exclusive events as follows.

$$\xi^j_{k_1}(n) = E^j_{k_1}(n) \bigcup \{ \bigcup_{\substack{k_2 \epsilon \{j+1,\ldots M\} \\ k_2 \neq k_1}} E^j_{k_1,k_2}(n)\} \ldots \bigcup E^j_{j+1,\ldots M}(n) \tag{27}$$

We now prove a lemma that generalizes lemma 3.1.

Lemma 4.1

In the system S^j, the probability of a successful transmission by terminal j at time n satisfies the following lower bound inequality for all $n \geq 1$,

$$\mu_j^j(n) \geq p_j \prod_{\substack{i=1 \\ i \neq j}}^{M} (1-p_i) + p_j \sum_{i=j+1}^{M} Pr[E_i^j(n)] p_i \prod_{\substack{k=1 \\ \neq i,j}}^{M} (1-p_k)$$

Proof

By definition,

$$\mu_j^j(n) = Pr[B_j^j(n)] \tag{28}$$

where $B_j^j(n)$ is the event that at time n terminal j has a successful transmission of a real or a "dummy" packet in system j. We apply the theorem of total probability to Eq. (28) conditioning the event $B_j^j(n)$ on the 2^{M-j} mutually exclusive events of Eq. (26). This results in the following equation.

$$Pr[B_j^j(n)] = Pr[B_j^j(n)/\{E_\phi^j(n)\}] \cdot Pr[E_\phi^j(n)] + \sum_{k_1=j+1}^{M} Pr[B_j^j(n)/E_{k_1}^j(n)] \cdot Pr[E_{k_1}^j(n)] + \ldots$$

$$\ldots + \sum_{\substack{k_1,k_2,\ldots k_e \in \{j+1,\ldots M\} \\ k_i \text{ all distinct}}} Pr[B_j^j(n)/E_{k_1,k_2,\ldots k_e}^j(n)] \cdot Pr[E_{k_1,k_2,\ldots k_e}^j(n)] + \ldots$$

$$+ Pr[B_j^j(n) \mid E_{j+1,\ldots,M}^j(n)] \cdot Pr[E_{j+1,\ldots,M}^j(n)] \tag{29}$$

A successful transmission by terminal j is impossible if any of the stations 1 through j-1 transmit. Let us define by $T_i^j(n)$ the event that terminal i attempts a transmission in slot n in system j. Then we decompose the general term of Eq. (29) according to the following:

$$Pr[B_j^j(n)/E_{k_1,k_2,\ldots k_e}^j(n)] = Pr[B_j^j(n)/\bigcap_{i=1}^{j-1}\{T_i^j(n)\}^c \bigcap \{Q_j^j(n) > 0\} \bigcap E_{k_1,k_2,\ldots k_e}^j(n)] \cdot$$

$$\cdot Pr[\bigcap_{i=1}^{j-1}\{T_i^j(n)\}^c / \{Q_j^j(n) > 0\} \bigcap E_{k_1,\ldots k_e}^j(n)] \cdot Pr[Q_j^j(n) > 0 / E_{k_1,k_2,\ldots k_e}^j(n)] +$$

$$+Pr[B_j^j(n)/\bigcap_{j-1}^{j-1}\{T_i^j(n)\}^c\bigcap\{Q_j^j(n)=0\}\bigcap E_{k_1,k_2,\ldots k_e}^j(n)]\cdot$$

$$\cdot Pr[\bigcap_{i=1}^{j-1}\{T_i^j(n)\}^c/\{Q_j^j(n)=0\}\bigcap E_{k_1,\ldots,k_e}^j(n)]\cdot Pr[\{Q_j^j(n)=0\}/E_{k_1,k_2,\ldots k_e}^j(n)] \tag{30}$$

Since the competition from the first group of j-1 terminals is fixed we have

$$Pr[\bigcap_{i=1}^{j-1}\{T_i^j(n)\}^c/\{Q_j^j(n)=0\}\bigcap E_{k_1,\ldots k_e}(n)]=\prod_{k=1}^{j-1}(1-p_k) \tag{31}$$

The same result holds for the case of $Q_j^j(n)>0$. Now let us consider the first term of the first summand of Eq. (30). The conditioning event implies that

i) terminals 1 through j-1 remain silent at time n.

ii) terminal j has a non-empty queue and hence will independently attempt transmission of an information packet with probability p_j.

iii) of the remaining terminals only k_1 through k_e are empty and hence they will remain silent (according to the rules of system S^j).

Based on this we deduce that

$$Pr[B_j^j(n)/\bigcap_{i=1}^{j-1}\{T_i^j(n)\}^c\bigcap\{Q_j^j(n)>0\}\bigcap E_{k_1,k_2,\ldots k_e}^j(n)]$$

$$=p_j\prod_{\substack{i\epsilon(j+1,\ldots,M)\\ i\notin\{k_1,k_2,\ldots k_e\}}}(1-p_i) \tag{32}$$

Consider now the first term of the second summand of Eq. (30). The conditioning event implies that

i) terminals 1 through j-1 remain silent at time n.

ii) terminal j has an empty queue; therefore, by the rules of in S^j, it will broadcast a dummy packet with probability p_j.

iii) Of the remaining terminals only k_1 through k_e are empty and hence they will remain silent. Based on this we have

$$Pr[B_j^j(n)/\bigcap_{i=1}^{j-1}\{T_j^j(n)\}^c \bigcap \{Q_j^j(n)=0\} \bigcap E_{k_1,\ldots k_e}^j(n)] = p_j \prod_{\substack{i=j+1,\ldots,M \\ i\notin\{k_1,k_2,\ldots k_e\}}} (1-p_i) \tag{33}$$

Note that when the jth queue is empty terminal j broadcasts a dummy packet, whereas when the queue is non-empty it broadcasts an information packet. However, this distinction is again of no concern since we seek a lower bound for the probability of successful transmission of information or dummy packets, and since the distinction becomes meaningless near saturation.

Substituting Eqs. (31), (32) and (33) in Eq. (30) we obtain

$$Pr[B_j^j(n)/E_{k_1,\ldots k_e}^j(n)] = p_j \prod_{\substack{i\notin\{k_1,\ldots,k_m\} \\ i\neq j}} (1-p_i) \tag{34}$$

Also note that

$$Pr[B_j^j(n)/E_\phi^j(n)] = p_j \prod_{\substack{1\leq i\leq M \\ i\neq j}} (1-p_i) \tag{35}$$

Note that the special nature of S^j enabled us to obtain eqns. (34) and (35) without having to estimate $Pr\{Q_j^j(n)>0/E_{k_1,\ldots k_e}^j(n)]$ or $Pr[Q_j^j(n)=0/E_{k_1,\ldots k_e}^j(n)]$, which would not be possible.

Now we invoke the inequality (18) and a slightly modified version of inequality (19), namely,

$$1 = \prod_{k=j+1}^{M} [p_k + (1-p_k)] \geq \prod_{k=j+1}^{M} (1-p_k) + \sum_{k_1=j+1}^{M} p_{k_1} \prod_{\substack{k=j+1 \\ k\neq k_1}}^{M} (1-p_k) \tag{36}$$

Substituting eqns. (34), (35) and using inequalities (18) and (36) in eqn. (29) we obtain

$$\mu_j^j(n) \geq p_j \prod_{\substack{k=1\\k\neq j}}^{M}(1-p_k)\cdot Pr[E_\phi^j(n)] + \sum_{k_1=j+1}^{M}\{p_j \prod_{\substack{k=1\\k\neq j}}^{M}(1-p_k) + p_j p_{k_1} \prod_{\substack{k=1\\k\neq k_1,j}}^{M}(1-p_k)\}Pr[E_{k_1}^j(n)] +$$

$$+ \sum_{\substack{\{k_1,k_2,\ldots k_e\}\in\{j+1,\ldots M\}\\ \text{all distinct}}} (\{p_j \prod_{\substack{k=1\\k\neq j}}^{M}(1-p_k)\} + \{p_j p_{k_1} \prod_{\substack{k=1\\k\neq k_1,j}}^{M}(1-p_k)\})\cdot Pr[E_{k_1,k_2,\ldots k_e}^j(n)] + \ldots$$

$$\ldots + (\{p_j \prod_{\substack{k=1\\k\neq j}}^{M}(1-p_k)\} + \{\sum_{k_1=j+1}^{M} p_j p_{k_1} \prod_{\substack{k=1\\k\neq k_1,j}}^{M}(1-p_k)\})\cdot Pr[E_{j+1,\ldots M}^j(n)] \qquad (37)$$

or, by further rearranging and using Eq. (27),

$$\mu_j^j(n) \geq p_j \prod_{\substack{j=1\\i\neq j}}^{M}(1-p_i) + \sum_{i=j+1}^{M} p_j\{p_i \prod_{\substack{k=1\\k\neq i,j}}^{M}(1-p_k)\}Pr[\xi_i^j(n)] \qquad (38)$$

This concludes the derivation.

By letting $n \to \infty$ in Eq. (38) we obtain the following asymptotic bound.

$$\tilde{\mu}_j^j > \mu_j^j = \lim_{n\to\infty}\mu_j^j(n) \geq p_j \prod_{\substack{i=1\\i\neq j}}^{M}\overline{p}_i + \sum_{i=j+1}^{M}(p_j p_i \prod_{\substack{k=1\\k\neq i,j}}^{M}\overline{p}_k)Pr[Q_i^j = 0] \qquad (39)$$

It is clear now that lower bounds for μ_j^j can be found if lower bounds for $Pr[Q_i^j = 0]$ can be determined. The latter are derived in the next lemma by examination of the departure statistics at queue i in S^j for $j < i$, and assuming that $\lambda_i < \tilde{\mu}_i^i$, for $i > j$, where $\tilde{\mu}_i^i$ is the average asymptotic service rate of queue i in system S^i.

Lemma 4.2

If $\lambda_i < \tilde{\mu}_i^i$ for all i, then in steady state $Pr[Q_i^j > 0] \leq \frac{\lambda_i}{\tilde{\mu}_i^i}$ for any $j < i$. *

* This result is similar to the one for the $G \mid G \mid 1$ queue, see Kleinrock [15], with average arrival rate λ and average service time $\overline{x}$. If the queue is stable for $\lambda < \frac{1}{\overline{x}}$, then $Pr[Q > 0] = \lambda\overline{x}$.

Proof:

Let $X_i^j(n)$ denote the number of packets in the ith queue of S^j after the nth departure.

Let $A_i^j(n+1)$ denote the number of packets that arrive at queue i of S^j during the $(n+1)^{st}$ service interval. Then:

$$X_i^j(n+1) = X_i^j(n) + A_i^j(n+1) - 1[X_i^j(n)] \tag{40}$$

where

$$1[x] = \begin{matrix} 1, & x > 0 \\ 0, & \text{else} \end{matrix}$$

We know from Appendix 1 that if $\lambda_i < \tilde{\mu}_i^j$ then the ith queue of S^j is stable. Hence there exists a steady state distribution for the system satistics. Furthermore, the results of Appendix 1 imply that the average queue sizes are finite in that case. Taking expectations in Eq. (40) with respect to this distribution we obtain the following:

$$E[X_i^j(n+1)] = E[X_i^j(n)] + E[A_i^j(n+1)] - Pr[X_i^j(n) > 0] =$$

$$= E[X_i^j(n)] + \lambda_i E[\text{ length of } (n+1)^{th} \text{ service interval for queue i}]$$

$$-Pr[X_i^j(n) > 0] \tag{41}$$

Now, in steady state, $E[X_i^j(n+1)] = E[X_i^j(n)]$; hence

$$Pr[X_i^j > 0] = \lambda_i E[\text{length of service interval of queue i in } S^j]$$

Since for $\lambda_i < \tilde{\mu}_i^j$ the ith queue is stable in S^j, we have

$$1 \geq Pr[X_i^j > 0] = \lambda_i f^j(\lambda_i) \tag{42}$$

where $f^j(\lambda_i)$ denotes the expected length of the service interval of queue i in S^j. Obviously this service time is also a function of the other arrival rates $\lambda_k, k > j$, but not of the ones for $k \leq j$, since arrivals to the first j queues have no bearing on how often there are transmission attempts from the group of the first j queues in S^j (they occur with probability $1-\prod_{k=1}^{j}(1-p_k)$ in every slot). In any event, the function f^j is non-decreasing in λ_i; this is quite obvious since more frequent arrivals will create more frequent collisions and less frequent successes by all queues, including the ith one. Thus, for $\lambda_i < \tilde{\mu}_i^j$,

$$f^j(\lambda_i) \leq f^j(\tilde{\mu}_i^j) \tag{43}$$

Using inequality (43) in Eq. (42) we obtain

$$1 \geq \tilde{\mu}_i^j f^j(\tilde{\mu}_i^j)$$

or

$$f^j(\tilde{\mu}_i^j) \leq \frac{1}{\tilde{\mu}_i^j}$$

Hence

$$Pr[X_i^j > 0] = \lambda_i f^j(\lambda_i) \leq \lambda_i f^j(\tilde{\mu}_i^j) \leq \frac{\lambda_i}{\tilde{\mu}_i^j} \tag{44}$$

Now we argue that

$$Pr[X_i^j > 0] = Pr[Q_i^j > 0] \tag{45}$$

Recall that Q_i^j is the steady state size of queue i in S^j at any arbitrary instant in time, whereas X_i^j is the steady state queue size at any arbitrary departure point in time. They could conceivably be different random variables. However since the joint queue size $[Q_i^j(n) - - - Q_n^j(n)]$ is an irreducible aperiodic Markov Chain, ergodicity implies that eq. (45) is valid.

Finally, by virtue of the dominance of S^j by S^{j+1} we conclude that

$$\tilde{\mu}_i^j \geq \tilde{\mu}_i^{j+1}, \qquad 1 \leq j \leq M-1$$

Consequently

$$Pr[Q_i^j > 0] \leq \frac{\lambda_i}{\tilde{\mu}_i^j} \leq \frac{\lambda_i}{\tilde{\mu}_i^i}, \qquad j+1 \leq i \leq M$$

This completes the proof.

Now we state the main theorem of this section.

Theorem 4.1

The system S is stable if

$$\lambda j < b_j^j, \qquad 1 \leq j \leq M$$

where

$$b_j^j \triangleq p_j \prod_{\substack{i=1 \\ i \neq j}}^{M} (1-p_i) + \sum_{i=j+1}^{M} p_j (1 - \frac{\lambda i}{b_i^i}) p_i \prod_{\substack{k=1 \\ k \neq i,j}}^{M} (1-p_k) \tag{46}$$

and

$$b_M^M \triangleq \mu_M^M = p_M \prod_{i=1}^{M-1} (1-p_i) \tag{47}$$

Proof

By lemma 4.2 we know that if $\lambda_i < \tilde{\mu}_i^i$ we have

$$Pr[Q_i^j = 0] \geq 1 - \frac{\lambda i}{\tilde{\mu}_i^i}, \qquad i > j$$

Substituting this in Eq (39) we obtain

$$\tilde{\mu}_j^j \geq p_j \prod_{\substack{i=1 \\ i \neq j}} \bar{p}_i + \sum_{i=j+1}^{M} (p_j p_i \prod_{\substack{k=1 \\ k \neq i,j}}^{M} \bar{p}_k)(1 - \frac{\lambda_i}{\tilde{\mu}_i^i}) \tag{48}$$

Note that the b_j^j's, as defined in Eqs (46), (47), are less than or equal to the $\tilde{\mu}_j^j$'s. Therefore substituting the b_j^j's in the RHS of inequality (48) we strengthen the bound. Thus if $\lambda_j < b_j^j$, it is also true that $\lambda_j < \tilde{\mu}_j^j$ and, hence, the queue is stable in S^j and, by the dominance property, in S as well. This completes the proof.

Notes:

1. This is an improved bound as easily seen from the comparison of b_1^1 to the value of the RHS of Eq. (24). In fact, it is easily verified that $b_j^j > \mu_j^M$.
2. The ordering of the terminals has been arbitrary, but the bound depends on that ordering. The union of the bounds of the stability regions over all possible orderings yields an improved bound.

5. The Connection to Information Theory

The results obtained so far for the stability of the slotted-ALOHA system present a striking similarity to those concerning the achievalbe region of the no-feedback collision channel [13,14]. First of all let us refer to the problem we studied here as the "stability" problem and to that considered in [13,14] as the "capacity" problem. Let us discuss their similarities and differences:

"Stability" problem - Determine the regions of values of the arrival rates λ_i of the Bernoulli arrival processes for which the discrete-time, slotted-ALOHA system of M buffered users is stable in the sense of ergodicity of the underlying chain, or finiteness of the average queue size (the two are equivalent here); that is, find the values of λ_i for which $\exists$ values of p_i, such that the system is stable. There is no concern about Shannon capacities, and feedback is assumed to be provided to the terminals.

"Capacity" problem - The object of study is the same; however, a different question is asked about it, namely at what rates R_i can the terminals simultaneously transmit to the common receiver reliably (in the Shannon sense). There is no concern about stability of queues; in fact the queues may very well be assumed to be infinite in order to provide an inexhaustible supply of packets, when needed, for transmission. There is no feedback, and the terminals are

not constrained to use the ALOHA protocol for transmission, but only some (any) protocol sequence of packet transmission and retransmission.

To summarize:

Both problems consider the same channel (the collision channel), that requires a single packet transmission to ensure successful reception. However they differ in terms of the question of interest (queue stability vs. reliable transmission rate), and in terms of the assumptions on feedback, arrivals, and transmission protocol.

Interestingly the results are remarkably similar for the two problems. In fact,

1) for $M = \infty$ (and $\lambda_1 = \lambda_2 = ...$) the "stability" problem requires that $\lim_{\substack{M\to\infty \\ \lambda_i\to 0}} M\lambda_i < e^{-1}$ and the "capacity" problem states that $\lim_{\substack{M\to\infty \\ \lambda_i\to 0}} MR_i < e^{-1}$ is achievable.

2) for $M = 2$ the "stability" problem, as shown here and in [8], requires that

$$\sqrt{\lambda_1} + \sqrt{\lambda_2} \leq 1$$

and the "capacity" problem as shown in [13,14] states that

$$\sqrt{R_1} + \sqrt{R_2} \leq 1$$

defines an achievable region.

3) for $2 < M < \infty$ the "stability" problem leads to the bounds derived here and elsewhere [6,7]. These bounds can be shown to be consistent with the results of the "capacity" problem, which state that

$$R_i = p_i \prod_{j\neq i}(1 - p_j)$$

$$\sum_{i=1}^{M} p_i = 1$$

defines an achievable region (capacity region, actually, under the constraint of using a transmission protocol sequence); namely that region contains the region of inner bounds obtained for the stability problem (as can be shown quite easily).

These observations lead to the following two questions.

1) Is there a non-superficial connection between the informationtheoretic and the queueing theoretic versions of the problem, or is the identity of the results a mere coincidence?

2) In either case, is the conjecture that for $2 < M < \infty$ the stability region is given by

$$\lambda_i = p_i \prod_{j \neq i} (1 - p_j)$$

$$\sum_{i=1}^{M} p_i = 1$$

true?

Whatever the answers to these questions we may close with two, hopefully, useful remarks.

i) Feedback is known to, in general, enlarge the capacity region of a multi-user channel; the results seem to imply that, at least for M=2, the ALOHA transmission protocol makes poor use of this feedback in that the achievable region is not enlarged, but, rather, stays puzzlingly the same.

ii) The meaning of the requirement that $\sum_{i=1}^{M} p_i = 1$ in the parametric description of the capacity region remains unclear, in connection with the role of the p_i's in the transmission protocol of the "stability" problem. The intriguing nature of the sum of the values of the p_i's was first noticed by Szpankowski in [6,7]. Here we simply note that for the "stability" problem one can easily show that if

$$\sum_{i=1}^{M} p_i < 1$$

all users can improve their delay (or queue-size) performance if they uniformly increase their respective p_i's by an amount given by

$$\Delta p_i = p_i(1 - p_i) \cdot \Delta t \tag{49}$$

where Δt is a common auxiliary variable, until the sum hits the value 1. Specifically one can show that any increment in the values of the p_i's according to Eq. (49) produces a system dominated by the previous one in the dominance sense that was used throughout this paper. We do not know how to show the improvement by decreasing the p_i's when $\sum_{i=1}^{M} p_i > 1$.

6. Conclusions

In this paper we have used a new technique involving stochastic dominance to study the ergodicity region of the discrete-time slotted-ALOHA system with a finite number of buffered users. We have obtained the region exactly for M=2 and derived inner bounds to it for $M > 2$, that improve upon earlier ones. The results look remarkably similar to those of the no-feedback collision channel and thus lead to speculation about a deeper relationship between the queueing-theoretic and information-theoretic approaches to the problem. They also suggest a conjecture about the precise boundaries of the stability region for $M > 2$. The dominance approach introduced in this paper represents a natural and simple concept that may be of further value to the study of other related problems.

Appendix 1

In this appendix we establish the dynamics of the ith queue in system S^j, as defined in the general formulation of section 4, and obtain the stability criterion that is used repeatedly in all sections of the paper. Specifically the objective is to show that in S^j, if $\lambda_i < \tilde{\mu}_i^j$, the ith queue is stable in the sense that

$$P_{i,0}^j = Pr[Q_i^j = 0] > 0.$$

Consider the ith queue of S^j at time n and the system equations for the queue-size probabilities. These are

$$P_{i,0}^j(n) = \overline{\lambda}_i P_{i,0}^j(n-1) + \overline{\lambda}_i \mu_{i,1}^j(n-1) P_{i,1}^j(n-1) \tag{A1.1}$$

$$\begin{aligned} P_{i,1}^j(n) = {} & [\overline{\lambda}_i \overline{\mu}_{i,1}^{-j}(n-1) + \lambda_i \mu_{i,1}^j(n-1)] P_{i,1}^j(n-1) \\ & + \lambda_i P_{i,0}^j(n-1) \\ & + \overline{\lambda}_i \mu_{i,1}^j(n-1) P_{i,2}^j(n-1) \end{aligned} \tag{A1.2}$$

$$\begin{aligned} P_{i,k}^j(n) = {} & \overline{\lambda}_i \overline{\mu}_{1,k-1}^j(n-1) P_{i,k-1}^j(n-1) \\ & + \overline{\lambda}_i \overline{\mu}_{i,k}^j(n-1) + \lambda_i \mu_{i,k}^j(n-1) P_{i,k}^j(n-1) \\ & + \overline{\lambda}_i \mu_{i,k+1}^j(n-1) P_{i,k+1}^j(n-1) \end{aligned} \tag{A1.3}$$

Obviously the "death-rates" $\mu_{i,k}^j(n)$ are functions of the status of the other queues, which is why the queues are coupled and difficult to anlayze. However, the above equations are valid conditioned on the status of the other queues. There is, in our notation, an explicit dependence shown on the size of the ith queue. Thus

$$\mu_{i,k}^j(n) = Pr[D_i^j(n) = 1 \mid Q_i^j(n) = k]$$

Suppose now that $P_{i,k}^j(n)$ and $\mu_{i,k}^j(n)$ possess limits as $n \to \infty$; that they indeed do will be established a little later. By considering the limits of the system equations we obtain

$$P^j_{i,0} = P^j_{i,0}\overline{\lambda}_i + \overline{\lambda}_i\mu^j_{i,1}P^j_{i,1} \tag{A1.4}$$

$$\begin{aligned} P^j_{i,1} &= \lambda_i P^j_{i,0} \\ &+\overline{\lambda}_i\overline{\mu}^{-j}_{i,1} + \lambda_i\mu^j_{i,1}P^j_{i,1} \\ &+\overline{\lambda}_i\mu^j_{i,2}P^j_{i,2} \end{aligned} \tag{A1.5}$$

$$\begin{aligned} P^j_{i,k} &= \lambda_i\overline{\mu}^j_{i,k-1}P^j_{i,k-1} \\ &+\overline{\lambda}_i\mu^j_{i,k+1}P^j_{i,k+1} \end{aligned} \tag{A1.6}$$

These equations can be solved to yield

$$P^j_{i,1} = P^j_{i,0}\frac{\lambda_i}{\overline{\lambda}_i\mu^j_{i,1}} \tag{A1.7}$$

$$P^j_{i,k} = P^j_{i,0}\frac{\lambda^k_i}{\overline{\lambda}^k_i}\frac{1}{\mu^j_{i,k}}\prod_{\ell-1}^{k-1}\frac{\overline{\mu}^j_{i,\ell}}{\mu^j_{i,\ell}} \tag{A1.8}$$

and hence

$$P^j_{i,0} = \frac{1}{1 + \frac{\lambda_i}{\overline{\lambda}_i\mu^j_{i,1}} + \sum_{k=2}^{\infty}\frac{\lambda_i}{\overline{\lambda}_i\mu^j_{i,k}}\prod_{\ell=1}^{k-1}(\frac{\overline{\mu}^j_{i,\ell}\lambda_i}{\mu^j_{i,\ell}\overline{\lambda}_i})} \tag{A1.9}$$

Based on Eq. (A1.9) we conclude that:

1. Queue i is stable in the sense that

$$Pr[Q^j_i = 0] = P^i_{i,0} > 0$$

if

$$\lambda_i < \lim_{k\to\infty} \mu^j_{i,k} \tag{A1.10}$$

2. Queue i is unstable in the sense that

$$Pr[Q^j_i = \infty] = \lim_{k\to\infty} P^j_{i,k} = 1$$

if

$$\lambda_i > \lim_{k\to\infty} \mu^i_{i,k} \tag{A1.11}$$

Note also that Eqs. (A1.8) and (A1.9) imply that under the stability condition the moments of the queue sizes exist as well.

We now wish to show that $\lambda i < \tilde{\mu}^j_i$ implies stability. If $\lambda_i < \tilde{\mu}^j_i$, one of the following is true:

Case 1 - $\lambda_i < \tilde{\mu}^j_i \leq \lim_{k\to\infty} \mu^j_{i,k}$

Case 2 - $\lambda_i < \lim_{k\to\infty} \mu^j_{i,k} \leq \tilde{\mu}^j_i$

Case 3 - $\lambda_i = \lim_{k\to\infty} \mu^j_{i,k} < \tilde{\mu}^j_i$

Case 4 - $\lim_{k\to\infty} \mu^j_{i,k} < \lambda_i < \tilde{\mu}^j_i$

Cases 1 and 2 satisfy Eq. (10) and hence the queue is stable.

Case 4 cannot occur, because if it does then by Eq. (A1.11) $\lim_{k\to\infty} P^j_{i,k} = 1$, which imples that

$$\tilde{\mu}^j_i = \sum_{k=0}^{\infty} \mu^j_{i,k} P^j_{i,k} = \lim_{k\to\infty} \mu^j_{i,k}$$

which contradicts the inequality

$$\lim_{k\to\infty} \mu^j_{i,k} < \tilde{\mu}^j_i.$$