A decentralized control strategy for multiaccess broadcast networks

Jessy W. GRIZZLE * and Steven I. MARCUS *
Department of Electrical Engineering, University of Texas at Austin, Austin, TX 78712, U.S.A.

Kai HSU **
Scientific Systems, Inc., San Antonio, TX 78216, U.S.A.

Received 24 March 1981
Revised 15 January 1982
Communicated by J. Meditch

A decentralized control strategy, based on the one step delay sharing information pattern, is proposed for controlling terminal access in a multi-access broadcast network. A Markov chain model of the system is calculated and the infinite time problem of maximizing average throughput is formulated and analyzed using decentralized control theory. Optimal control policies are shown to be extreme and stationary. For comparison purposes, several other schemes that have appeared in the literature are briefly described. All schemes are then numerically compared and their delay versus throughput curves presented.

1. Introduction

There are two fundamental approaches to allocation of communication bandwidth: pre-allocation and dynamic allocation [15]. In pre-allocation, a fixed bandwidth is pre-allotted for a single call, from source to destination, and is not released until the call is completed. Conversely, with dynamic allocation, a message is dynamically allotted bandwidth on a link to link basis; one never attempts to schedule bandwidth over the entire source to destination path. Packet switching takes this one step further; a message is divided into small segments, called packets, which are individually routed through the network. The first major network to prove the viability of this method was the ARPANET which began operation in 1969; many more have since followed.

In this paper, a special class of packet switching networks, termed multiaccess broadcast networks, is considered in a control framework. Such networks are characterized by many users sharing a single communication channel which is assumed to provide a fully connected topology; by this it is meant that each terminal can transmit to and receive from every other terminal in the network. In addition, it is assumed that transmissions over the channel are successful if, and only if, they do not overlap in time (are disjoint); coupled with the assumption of a fully connected topology this guarantees that all messages are sent directly to their destinations since nothing can be gained by transmitting to an 'intermediate' terminal. Note how this simple configuration contrasts to a more general network (such as the ARPANET) in which terminals are only indirectly linked and packets are propagated through the network in a similar manner as a letter through the U.S. Postal System.

* Research supported in part by the Air Force Office of Scientific Research under Grant AFOSR-79-0025.
** Research supported in part by the Joint Services Electronics Program at the University of Texas at Austin under Contract F49620-77-C-0101.

Many reasons for studying multiaccess networks have been offered in the literature; two are stressed here. First, local communication needs could easily be served by such a network. Since local distribution is by far the most expensive portion of a communication network, ground radio techniques are of considerable interest to the extent that they can replace wire for local distribution [15]. Second, the more general (partially connected) networks are very difficult to analyze. It is often the case that to find an optimal control policy it is necessary to search over a set of possible solutions that 'exceeds the number of atoms in the universe' [8]. Thus suboptimal procedures, often relying on heuristics, must be used. The insight and intuition necessary to do this effectively may first be gained by considering simpler problems, such as multiaccess broadcast networks, which retain many of the features of the original problem.

For multiaccess networks, the control problem is essentially one of deciding how individual terminals are to access the channel so as to maximize the performance index. Obviously, the solution to this problem is intimately related to the assumed underlying information structure. An access scheme based on a nonclassical information pattern – *one step delay sharing* (OSDS) – is the major topic of this paper. The work of [6], [7], and [18] extending the theory of controlled Markov chains to the OSDS case allows a rigorous formulation and solution of the multiaccess problem within the domain of decentralized control theory. This is in contrast to other schemes which have appeared in the literature, such as Controlled Aloha [11] and Urn [10] where the problem is formulated and solved in a centralized framework and then the constraints are 'relaxed' to admit a distributed implementation of the control policy. The 'relaxation' usually involves substituting estimated quantities for exact values; i.e., a certainty equilvalence principle is assumed to hold. The optimality of such a solution is not at all clear.

The formulation and analysis of the OSDS protocol is presented in detail. First the general description of a multiacess network is given, independent of an underlying information pattern; this is done to facilitate the comparison of the OSDS protocol to some other access techniques later in the paper. The OSDS information pattern, which results in the OSDS access scheme, is then introduced. From this a Markov chain model is formed and the performance measures, throughput and delay, are expressed in terms of the model variables. Next the Dynamic Programming and Policy Iteration Algorithms for the OSDS case are summarized in Appendix A. Combining these tools with some standard results in the literature, optimal policies are shown to exist and to be non-randomized, stationary, and extreme (for the infinite horizon problem). They can be calculated using the Policy Iteration Algorithm, but it becomes computationally impractical when the number of terminals is large. An alternate (possibly sub-optimal) procedure is discussed and shown to give good results. The OSDS scheme is numerically compared to four other access schemes which are briefly described in the paper. It is shown to perform nearly as well as the Perfect Information scheme. This is due to the use of a control channel for the interchange of (delayed) state information between the terminals. The final numerical analysis studies the performance of the OSDS scheme when this additional bandwidth is accounted for.

For previous work on access schemes utilizing delayed sharing information patterns, see [4,5,17,19].

2. Description of the basic model

In the following, a multiaccess broadcast radio network consisting of a population of *N packet radio units* [1] (PRUs) $PR_1, PR_2, \ldots, PR_N$ which share a single radio channel for communication is considered. It is assumed that the local topography is such that the radio channel is a fully connected communication medium; i.e., each PRU can communicate with every other PRU in the network. Message arrivals at the PRUs are modeled as independent Bernoulli processes with success parameters p_i; for simplicity, it is assumed that all messages are uniform in length (specifically one packet). Each terminal is assumed to have only a single buffer, and when it is full, arriving messages are refused; these are assumed to be lost. The time axis is supposed to be divided into slots of duration equal to the transmission time of a single packet. All transmissions are in the form of packets of fixed length and are synchronized to initiate at the slot

[1] This terminology was introduced in [21].

boundaries. Furthermore, a packet is successfully transmitted if and only if exactly one terminal sends a packet in any particular slot; all collisions are mutually destructive. A copy of a message is retained at a terminal until it is successfully sent. Finally, in the case of the OSDS scheme, an additional radio channel, similar to the message channel, will be assumed to exist for the dissemination and collection of control data.

2.1. OSDS protocol

Let $x(t)=(x_1(t), x_2(t),\ldots,x_N(t))^T$ denote the state of the network, where $x_i(t)$ equals one if PR_i has a message at time t and equals zero otherwise. Also, let $u(t)=(u_1(t), u_2(t),\ldots,u_N(t))$ where $u_i(t)$ is the decision or control variable of PR_i; at each time t, PR_i transmits a message according to a Bernoulli probability distribution having success parameter $u_i(t)$. The distributions for the individual PRUs are assumed to be mutually independent. Note that the control $u_i(t)$ should not be confused with the event of transmission of a message by PR_i for which $u_i(t)$ is the probability.

At the beginning of every time slot the ith PRU sends to all other PRUs the two-tuple $(x_i(t), u_i(t))$ over the control channel. [2] Note that this is not another random acess problem because the amount and source of the data is known a priori. Hence, this subcommunication problem is assumed to have been solved using standard communication techniques. Due to finite bandwidth restrictions this information will not be available to the terminals until the next time slot. The ith PRU thus knows:

(a) the 'augmented state' (x, u) of the network perfectly with one step delay, and

(b) whether or not it has a message ready for transmission at the present time (i.e. $x_i(t)$).

In summary, a PRU's observation at time t is given by $(x_i(t), x(t-1), u(t-1))$. Finally, each transmission probability $u_i(t)$ is constrained to be a function of at most $x_i(t)$ and all past state information $(x(s), u(s); s \leqslant t-1)$.

Remark 2.1. At least theoretically the above access scheme could be implemented as described. The desirability or practicallity of doing so will be left as an open question until Section 5. The use of an additional channel to increase the utilization efficiency of the message channel is not a new concept. Reservation schemes often employ a reservation channel on which terminals make requests for time slots in which to transmit packets over the message channel [14,16]; however, the reservation channel may become another multiaccess problem – unlike the control channel proposed here.

2.2. Markov model

A controlled-vector-Markov chain model of the above network can be constructed. To aid in writing down the one step state transition probabilities, the following definitions are made:

- $\{\eta_1, \eta_2,\ldots,\eta_n\}$, where $n=2^N$ is the set of N-vectors with entries consisting of 0 or 1;
- $S=\{1, 2,\ldots,N\}$ is the set of terminals labeled 1 through N;
- $M^t=\{k\in S \mid x_k(t)=1\}$ is the set of PRUs that have messages at time t;
- $Q^t=M^t-M^{t+1}=\{q_1, q_2,\ldots,q_r\}$ is the set of PRUs that had a message at time t but do not have a message at time $t+1$;
- $|\cdot|$ denotes the number of elements (cardinality) in a set;
- $p_{ij}(u(t))=P\{x(t+1)=\eta_j \mid x(t)=\eta_i, u(t)\}$.

With the above notation the one-step state transition probabilities are

$$p_{ij}(u(t))=\begin{cases} 0, & |Q^t|\geqslant 2, \\ \alpha_1\beta, & |Q^t|=1, \\ \left(1-\sum\limits_{k\in M^t}\alpha_k\right)\beta, & |Q^t|=0, \end{cases} \tag{2.1}$$

[2] An alternate implementation would be to embed the control channel in the message channel as in [20]. This is discussed later.

where

$$\alpha_r = u_{q_r}(1-p_{q_r}) \prod_{l \in M^t - q_r} (1-u_l), \tag{2.2}$$

$$\beta = \prod_{l \in M^{t+1} - M^t} p_l \times \prod_{m \in S - M^{t+1}} (1-p_m). \tag{2.3}$$

The reasoning is as follows. If $|Q^t| \geq 2$, then at least two terminals have successfully sent a message. This event is clearly impossible, so the probability of occurrence must be zero. If $|Q^t| = 1$, then exactly one of the terminals (PR_{q_1}) sent a packet and did not receive a 'new' one. In addition, $\{PR_i | i \in M^{t+1} - M^t\}$ received new packets and $\{PR_i | i \in S - M^{t+1}\}$ did not. Lastly, $|Q^t| = 0$ implies that no terminal with a message at time t successfully sent it and also received a new packet. As before, $\{PR_i | i \in M^{t+1} - M^t\}$ received new messages and $\{PR_i | i \in S - M^{t+1}\}$ did not.

Recalling the independence assumption on the transition probabilities (see beginning of Section 2.1), the marginal conditional probabilities for the individual terminals can be easily calculated; they are

$$\begin{aligned}
&P\{x_i(t+1)=1 \,|\, x_i(t)=0,\, u_i(t);\, x_j(t),\, u_j(t), j \in S-\{i\}\} = p_i, \\
&P\{x_i(t+1)=0 \,|\, x_i(t)=0,\, u_i(t);\, x_j(t),\, u_j(t), j \in S-\{i\}\} = 1-p_i, \\
&P\{x_i(t+1)=1 \,|\, x_i(t)=1,\, u_i(t);\, x_j(t),\, u_j(t), j \in S-\{i\}\} = \\
&\quad = 1 - u_i(t)\left[\prod_{j \in S-\{i\}} (1 - x_j(t)u_j(t))\right](1-p_i), \\
&P\{x_i(t+1)=0 \,|\, x_i(t)=1,\, u_i(t);\, x_j(t),\, u_j(t), j \in S-\{i\}\} = \\
&\quad = u_i(t)\left[\prod_{j \in S-\{i\}} (1 - x_j(t)u_j(t))\right](1-p_i)
\end{aligned} \tag{2.4}$$

for all $i \in S$. It can be shown that if the controls $u_i(t)$ are restricted to the set $\{0, 1\}$, then the states of the terminals are conditionally independent, and thus the joint probabilities can be calculated by multiplication of the marginal probabilities (it will later be shown that optimal controls *are* extreme).

2.3 Performance measures

The objective function to be extremized is

$$J = \lim_{T \to \infty} \frac{1}{T} E\left\{\sum_{t=1}^{T} g(\cdot)\right\} \tag{2.5}$$

where $g(\cdot)$ is the cost-per-stage. Two types of cost-per-stage are defined here, throughput and delay.

The instantaneous throughput at time t is defined to be expected number of successful packet transmissions at time t. Since a message is delivered successfully if and only if exactly one terminal with a message transmits and since the transmission probabilities are assumed to be mutually independent, instantaneous throughput is given in terms of the model variables by

$$\sum_{i \in S} x_i(t)u_i(t) \prod_{j \in S-\{i\}} (1 - x_j(t)u_j(t)). \tag{2.6}$$

Average throughput is defined as

$$\text{Avg. throughput} = \lim_{T \to \infty} \frac{1}{T} E\left\{\sum_{t=1}^{T} \text{Throughput}(t)\right\}. \tag{2.7}$$

Assuming the instantaneous throughput is a stationary process, [3] average throughput equals the probability of a successfull packet transmission in any given time slot.

[3] This will be true whenever both the message arrival rates and the access protocol are stationary.

The instantaneous delay is defined as the random process, taking values in $\{1, 2, \ldots, N\}$, which equals j if there are j packets in the system at time t (either waiting to be, or in the process of being, transmitted); it is given simply by

$$\sum_{i \in S} x_i(t). \tag{2.8}$$

Average delay is by definition the expected time a packet resides in the system, as measured from the time it arrives through the time it is successfully delivered. Assuming that the limit in (2.7) exists and that the

$$\lim_{T \to \infty} \frac{1}{T} E\left\{ \sum_{t=1}^{T} \text{Delay}(t) \right\} \tag{2.9}$$

exists, Little's result [9] implies that

$$\text{Avg. delay} = \lim_{T \to \infty} \frac{E\left\{ \sum_{t=1}^{T} \text{Delay(t)} \right\}}{E\left\{ \sum_{t=1}^{T} \text{Throughput}(t) \right\}}. \tag{2.10}$$

Henceforth, when no confusion will occur, throughput (delay) and average throughput (delay) will be used interchangeably. The exact meaning should be clear from the context.

3. Solution with OSDS information

In this section the problem of determining a policy which maximizes the average expected throughput is investigated. The main tools applied are the Dynamic Programming Algorithm (DPA) and the Policy Iteration Algorithm (PIA) which have been recently worked out for the OSDS information pattern [6,7,18]. The results are similar to the centralized case with the 'state at time t' in the algorithms being effectively given by $(x(t-1), u(t-1))$. In addition, there is a constraint which enforces the decentralized nature of the admissible control laws. A brief summary of the relevant results is included in Appendix A.

This section is structured as follows. First the optimization problem is stated and it is established that a nonrandomized stationary optimal policy exists. Then it is shown that the optimal policy can be chosen to be extreme, i.e. $u_i(t) \in \{0, 1\}$. This permits the application of the PIA given in [6,7]. Next, an equivalence relation is introduced on the 'state space' of the PIA which reduces it in size form 4^N to 2^N elements. By using the PIA on the resulting equivalence classes, the optimal control law for a 2-PRU network is explicitly calculated. It is seen to be similar to the optimal control law of the corresponding centralized problem. On the bases of this result and several numerical calculations of optimal policies for larger networks, a conjecture is made on the form of the optimal policy for a general N-terminal network.

3.1. Preliminary notions

The objective is to determine a control policy which maximizes (2.5) with cost-per-stage (2.6) subject to the one-step state transition probabilities (2.1) and the OSDS information pattern. That the problem is well posed will now be established.

Let $\delta_t = \{x_i(s), u_i(s) \mid s \leqslant t, i \in S\}$ where $\delta_0 = \emptyset$, the empty set. Then the information available to PR_i at time t is $\delta_{t-1} \cup x_i(t)$. Let a control policy be denoted by $\pi = \{\mu(1), \mu(2), \ldots\}$ where $\mu(t) = \{\mu_1(t), \mu_2(t), \ldots, \mu_N(t)\}$ and $\mu_i(t)$ is a random variable mapping $\{\delta_{t-1} \cup x_i(t)\}$ to $[0,1]$. By combining the techniques of [6] with Corollary 9.18.1 of [2], it can be shown (a) that an optimal policy exists and (b) that it can be chosen to be both stationary and nonrandomized. Moreover, from the DPA it follows that $\mu_i(t) = \mu_i(x_i(t), x(t-1), u(t-1))$; i.e. $\mu_i(t)$ is merely a functional mapping $x_i(t)$, $x(t-1)$, $u(t-1)$ to $[0,1]$.

3.2. Reduction and solution

The PIA, as derived for the OSDS information pattern in [6,7], applies only to problems with both a finite state space and a finite control space. The model as stated in Section 2 has a finite state space ($n=2^N$) but an uncountable number of control values are followed ($u_i(t)\in[0, 1]$). Of course, the interval [0, 1] could be discretized and the algorithm applied. However, it will be shown that optimal policies are extreme ($u_i(t)\in\{0, 1\}$) and thus that the control space can be restricted to $\{0, 1\}\times\{0, 1\}\times\cdots\times\{0, 1\}$. (Note that this is *not* a consequence of the nonrandomness of the control policy; for example, the Controlled Aloha scheme [11] has a nonrandomized nonextreme retransmission policy.) Before proceeding some preliminary results are needed.

Definition 3.1. Let F_N be the set of *multiaffine functionals in N variables*; i.e. F_N consists of all functions f_N such that

$$f_N(x_1, x_2,\ldots,x_N)=a_0+\sum_{i=1}^{N} a_i x_i+\sum_{i=1}^{N}\sum_{j=i+1}^{N} a_{ij}x_i x_j+\cdots$$
$$+\sum_{i=1}^{N}\sum_{j=i+1}^{N}\cdots\sum_{l=k+1}^{N} a_{ij\ldots kl}x_i x_j\cdots x_l+\cdots+b_0 x_1 x_2\cdots x_N \tag{3.1}$$

for some a_0, b_0, $a_{l_1 l_2\ldots l_k}\in\mathbb{R}$, $k=1, 2,\ldots,l_k\in\{1, 2,\ldots\}$. Furthermore denote by F the class of all multiaffine functional F_N, $N=1, 2,\ldots$.

Lemma 3.2. *Let* $C^N=[0, 1]\times[0, 1]\times\cdots\times[0, 1]$ *and* $D^N=\{0, 1\}\times\{0, 1\}\times\cdots\times\{0, 1\}$. *For each functional* $f_k\in F$,

$$\min_{x\in C^k} f_k(x)=\min_{x\in D^k} f_k(x) \quad \textit{and} \quad \max_{x\in C^k} f_k(x)=\max_{x\in D^k} f_k(x). \tag{3.2}$$

Proof. The proof is by induction on k and uses only standard arguments. For details see [3].

Lemma 3.3. *Let* $u_j(t)$, $j=1, 2,\ldots,N$, *be the independent variables in* (2.1) *and* (2.6) *such that* Γ (*see Appendix A*) *is satisfied; then* (2.1) *and* (2.6) *are elements of* F_N, *the set of multiaffine functionals in* $u_1, u_2,\ldots,u_N$.

Proof. Γ is the constraint which imposes the decentralized nature of the control laws in the DPA and PIA algorithms. By inspection, it does not affect the multiaffineness of (2.1) or (2.6).

Using the multiaffine property of the cost-per-state and the one-step-state transition probabilities with the DPA, it is now possible to show that optimal policies are extreme.

Theorem 3.4. *Assume the dynamic equations* (2.1), *objective function* (2.5) *with cost-per-stage* (2.6), *and one-step delay sharing information pattern. Then the optimal control policy is extreme, i.e.,* $\mu_i(x_i(t), x(t-1), u(t-1))\in\{0, 1\}$.

Proof. Consider the T-step truncated problem. Applying the dynamic programming algorithm (see Appendix A),

$$J_T^*(x(T-1), u(T-1))=\max_{\substack{v_l(T)\in\Gamma_l\\ l=1,2,\ldots,N}}\left\{\sum_{i=1}^{n} g\big(\eta_i, v^i(T)\big)\alpha_i(T)\right\} \tag{3.3}$$

where

$$v_l^i(t)=\mu_l(x_l(\eta_i), x(t-1), u(t-1)),\quad i=1, 2,\ldots,N,$$
$$v^i(t)=\left[v_1^i, v_2^i,\ldots,v_N^i\right],\qquad v_l(t)=\left[v_l^1, v_l^2,\ldots,v_l^n\right]^{\mathrm{T}},$$
$$\alpha_j(t)=P\{x(t)=\eta_j\,|\,x(t-1), u(t-1)\}.$$

Note that $v_i^i(t)$ is a probability and therefore $v^i(t) \in C^N\ \forall i=1,2,\ldots,N$, $t=1,2,\ldots,T$. Since $\alpha_i(T)$ is not a function of $v^i(T)$, by Lemmas 3.2 and 3.3 it follows that $\mu^*(T) \in D^N$.

From the recursion relation of the dynamic programming algorithm,

$$J_{T-1}^*(x(T-2), u(T-2)) = \max_{\substack{v_i(T-1) \in \Gamma_i \\ i=1,2,\ldots,N}} \left\{ \sum_{i=1}^{n} \left[g(\eta_i, v^i(T-1)) + J_T^*(\eta_i, v^i(T-1)) \right] \alpha_i(T-1) \right\}.$$

$\alpha_i(T-1)$ is not a function of $v^i(T-1)$. By Lemma 3.3, $g(\eta_i, v^i(T-1))$ is an element of F_N. $J_T^*(\eta_i, v^i(T-1))$ depends on $v^i(T-1)$ only through $\alpha_i(T)$ which, by Lemma 3.3, is an element of F_N. Thus $\mu^*(T-1) \in D^N$. By using backward induction, $\mu^*(t) \in D^N\ \forall t=T, T-1,\ldots,1$; hence there always exists an extreme policy $\pi^* = \{\mu^*(1), \mu^*(2),\ldots,\mu^*(T)\}$ such that

$$E_{\pi^*} \sum_{t=1}^{T} g(x(t), u(t)) \geqslant E_{\pi} \sum_{t=1}^{T} g(x(t), u(t)) \tag{3.4}$$

where $\pi = \{\mu(1), \mu(2),\ldots,\mu(T)\}$ is any policy. Multiplying on both sides of (3.4) by $1/T$ and taking the limit as $T \to \infty$ gives

$$\lim_{T\to\infty} \frac{1}{T} E_{\pi^*}\left\{ \sum_{t=1}^{T} g(x(t), u(t)) \right\} \geqslant \lim_{T\to\infty} \frac{1}{T} E_{\pi}\left\{ \sum_{t=1}^{T} g(x(t), u(t)) \right\} \tag{3.5}$$

which completes the proof.

The 'state space' of the OSDS scheme has been reduced from $\{\eta_1,\ldots,\eta_n\} \times [0, 1]^N$ to $\{\eta_1,\ldots,\eta_n\} \times \{0, 1\}^N$. Further reduction is possibly by introducing the following relation on $\{\eta_1,\ldots,\eta_n\} \times \{0, 1\}^N$. Define the relation R by

$$(\eta_i, u) R (\eta_j, \hat{u}) \quad \text{iff} \quad p_{k_i}(u) = p_{k_j}(\hat{u}) \quad \forall k=1,2,\ldots,n. \tag{3.6}$$

It is easily verified that R is in fact an equivalence relation. For the two-user case the equivalence classes are

$$\begin{aligned}
&\left[\begin{bmatrix}0\\0\end{bmatrix}, 0, 0\right] = \left\{ \left(\begin{bmatrix}0\\0\end{bmatrix}, 0, 0\right), \left(\begin{bmatrix}0\\0\end{bmatrix}, 0, 1\right), \left(\begin{bmatrix}0\\0\end{bmatrix}, 1, 0\right), \left(\begin{bmatrix}0\\0\end{bmatrix}, 1, 1\right), \right.\\
&\qquad \left. \left(\begin{bmatrix}0\\1\end{bmatrix}, 0, 1\right), \left(\begin{bmatrix}0\\1\end{bmatrix}, 1, 1\right), \left(\begin{bmatrix}1\\0\end{bmatrix}, 1, 0\right), \left(\begin{bmatrix}1\\0\end{bmatrix}, 1, 1\right) \right\},\\
&\left[\begin{bmatrix}0\\1\end{bmatrix}, 0, 0\right] = \left\{ \left(\begin{bmatrix}0\\1\end{bmatrix}, 0, 0\right), \left(\begin{bmatrix}0\\1\end{bmatrix}, 1, 0\right), \left(\begin{bmatrix}1\\1\end{bmatrix}, 1, 0\right) \right\},\\
&\left[\begin{bmatrix}1\\0\end{bmatrix}, 0, 0\right] = \left\{ \left(\begin{bmatrix}1\\0\end{bmatrix}, 0, 0\right), \left(\begin{bmatrix}1\\0\end{bmatrix}, 0, 1\right), \left(\begin{bmatrix}1\\1\end{bmatrix}, 0, 1\right) \right\},\\
&\left[\begin{bmatrix}1\\1\end{bmatrix}, 0, 0\right] = \left\{ \left(\begin{bmatrix}1\\1\end{bmatrix}, 0, 0\right), \left(\begin{bmatrix}1\\1\end{bmatrix}, 1, 1\right) \right\}.
\end{aligned} \tag{3.7}$$

The generalization to N-users is obvious.

In Appendix A the PIA of [6,7] is adapted to the equivalence classes. This reduces the number of simultaneous equations that must be solved in the algorithm from $4^N + 1$ to $2^N + 1$. It was with this form of the PIA that the following theorem was proved in [3].

Theorem 3.5 *Consider a network of two PRUs,* PR_1 *and* PR_2*, with message arrival rates p and q, respectively,* *where $p \geqslant q$. Then the control policy which maximizes the average expected throughput is given by*

$$\begin{aligned}
&\mu_1\left(1,\left[\begin{bmatrix}0\\0\end{bmatrix},0,0\right]\right)=1, \qquad \mu_2\left(1,\left[\begin{bmatrix}0\\0\end{bmatrix},0,0\right]\right)\in\{0,1\},\\
&\mu_1\left(1,\left[\begin{bmatrix}0\\1\end{bmatrix},0,0\right]\right)=0, \qquad \mu_2\left(1,\left[\begin{bmatrix}0\\1\end{bmatrix},0,0\right]\right)=1,\\
&\mu_1\left(1,\left[\begin{bmatrix}1\\0\end{bmatrix},0,0\right]\right)=1, \qquad \mu_2\left(1,\left[\begin{bmatrix}1\\0\end{bmatrix},0,0\right]\right)=0,\\
&\mu_1\left(1,\left[\begin{bmatrix}1\\1\end{bmatrix},0,0\right]\right)=1, \qquad \mu_2\left(1,\left[\begin{bmatrix}1\\1\end{bmatrix},0,0\right]\right)=0,
\end{aligned} \tag{3.8}$$

where

$$\mu_2\left(1,\left[\begin{bmatrix}0\\0\end{bmatrix},0,0\right]\right)=\begin{cases}0 & \text{if } (p^4-p^3-2p^2+p)q^2-(p^4+p^3+-p^2-2p+1)q+p^3>0,\\ 1 & \text{otherwise.}\end{cases}$$

Proof. The proof is a straightforward though tedious application of the PIA. The symbolic manipulation program MACSYMA was used to great advantage in carrying out the required computations. For details see [3].

Notice that the form of the optimal control law is very similar to the optimal policy for the perfect centralized information case: within the set of terminals having messages with certainty, the one with the highest message arrival rate is given exclusive transmission rights (see Section 4). This similarity is exploited in the following conjecture.

Conjecture 3.6. *Let the N PRUs be labeled such that $p_1 \geqslant p_2 \geqslant \cdots \geqslant p_N$.* Define *$M'(\eta_i, \hat{u})$ to be the set of terminals at time t that have messages with certainty conditioned on $x(t-1)=\eta_i$, $u(t-1)=\hat{u}$; i.e.,*

$$M'(\eta_i,\hat{u})=\{m_1,m_2,\ldots,m_r\}=\left\{j \mid P\left[x_j(t)=1 \mid x(t-1)=\eta_i,\ u(t-1)=\hat{u}\right]=1,\ j=1,2,\ldots,N\right\}.$$

Then the control policy which maximizes the average throughput is characterized by

$$\begin{aligned}
&\mu_1(1,[\eta_i,\hat{u}])=1,\\
&\mu_l(1,[\eta_i,\hat{u}])=\begin{cases}1 & \text{if } M'(\eta_i,\hat{u})\neq\emptyset,\ m_1=l,\\ 0 & \text{if } M'(\eta_i,\hat{u})\neq\emptyset,\ m_1\neq l,\end{cases}\\
&\mu_k(1,[\eta_i,\hat{u}])\geqslant\mu_j(1,[\eta_i,\hat{u}]) \quad \text{if} \quad M'=\emptyset,\ k>j;\ k,j=1,2,\ldots,N.
\end{aligned} \tag{3.9}$$

In words, the optimal policy is: if based on the shared information only, the PRUs can determine that a set of terminals have messages with probability one then the PRU with the highest message arrival rate in that set is given exclusive transmission rights; otherwise, PRUs 1 through L are given transmission rights and will send if a message is possessed. L is not specified but is a function of the message arrival rates p_i (as in the 2-user case).

Remark. 3.7. A proof analogous to the case $N=2$ does not seem possible. The difficulty lies in symbolically performing the PIA for an arbitrary N. The conjecture has been verified numerically for $N \leqslant 5$ over a wide range of input rates p_i. Furthermore, the numerical results presented in Section 5 for larger networks indicate that even if the conjecture does not characterize the optimal policy correctly, it does provide a good suboptimal solution. This is important since the amount of computation required to carry out the PIA grows exponentially with increasing N.

4. Some other access schemes

In this section a few other access schemes are described, They will be used in Section 5 to indicate the relative performance of the OSDS scheme. Their control policies span the range from closed loop to full perfect state feedback to reservation.

4.1. TMDA

The traditional method of sharing a channel among several users is time division multiple access (TDMA), where at each slot a single PRU is chosen to have access rights to the channel. Two methods used to choose a PRU are: sequential and random.

With the sequential TDMA protocol (STDMA) the PRUs are consecutively given access rights to the channel; first PR_1, then $PR_2,\ldots,PR_N$, $PR_1,\ldots$,etc. When a PRU's turn comes up, if it has a message it sends, else the slot is wasted. The average throughput T and delay D are easily calculated [3] and are given by

$$T=\frac{1}{N}\sum_{i=1}^{N}\left[1-(1-p_i)^N\right] \tag{4.1}$$

and

$$D=N+1-\frac{1}{NT}\sum_{i=1}^{N}\frac{1-(1-p_i)^N(1+Np_i)}{p_i}. \tag{4.2}$$

In the special case of a homogeneous network, these equations reduce to

$$T_{\mathrm{H}}=1-(1-p)^N \tag{4.3}$$

and

$$D_{\mathrm{H}}=1+N/T_{\mathrm{H}}-1/p \tag{4.4}$$

where $p=p_1=p_2=p_N$, and the subscript H denotes 'homogeneous'.

With the random TDMA protocol (RTDMA), at the beginning of each slot a PRU is randomly selected to have transmission rights. As with STDMA, when chosen, if a PRU has a message it sends; otherwise the slot is wasted. Let c_i denote the probability that, in any given slot, PR_i is chosen to have transmission rights. Then the average throughput and delay are given by [3]

$$T=\sum_{i=1}^{N}\frac{c_ip_i}{p_i+c_i-p_ic_i} \tag{4.5}$$

and

$$D=\frac{1}{T}\sum_{i=1}^{N}\frac{p_i}{p_i+c_i-p_ic_i}. \tag{4.6}$$

When the network is homogeneous ($p_i=p$ and $c_i=1/\mathrm{N}$) these equations reduce to

$$T_{\mathrm{H}}=\frac{Np}{(N-1)p+1} \tag{4.7}$$

and

$$D_{\mathrm{H}}=N. \tag{4.8}$$

4.2. Perfect information

The most information that could be made available to a terminal is the identity of every PRU that has a message at the beginning of each slot. This is called the Perfect Information case. The control law which maximizes throughput is: If M^t is the set of PRUs at time t that have messages, then give exclusive transmission rights to the PRU $\in M^t$ that has the largest message arrival rate p_i.

The Perfect Information (PI) case, for obvious reasons, is never implemented. Its value lies in the use as a benchmark for all other techniques.

4.3. IFFO – A reservation protocol

In the IFFO (*interleaved frame flush-out*) scheme [20], the channel slots are grouped into frames, each consisting of a reservation slot followed by a sequence of reserved message slots. The reservation slot consists of N minislots which are assigned in a (contention-free) STDMA fashion. The terminals are assumed to have infinite buffers; during each reservation slot, the terminals reserve message slots for all packets received since the last reservation slot. To allow for propagation delay, these reservations will not be 'honored' until the next frame. Thus after a reservation slot, the terminals sequentially transmit their messages in slots that were reserved in the previous frame. For further details the reader is referred to [20].

Remark 4.1. The IFFO scheme was included for comparison with the OSDS scheme because of the similarities in the information patterns of the two schemes. However, since it assumed the PRUs have infinite buffers, its performance is not directly comparable to the other schemes considered. When the network is lightly loaded ($\ll 1$ new messages arriving each slot) the difference between having one buffer and many buffers will be negligible. For higher message arrival rates, a network with only one buffer will appear to have better delay versus throughput properties than one with many buffers because messages arriving at 'busy' PRUs are rejected and thus do not contribute to the network delay calculations. Of course, this difference is artificial because in practice rejected messages must be reinserted into the stream of arriving messages and eventually be transmitted.

5. Numerical results and comparisons

In Fig. 1, the delay versus throughput curves of the OSDS and PI schemes are presented for homogeneous networks consisting of 5, 10, and 50 PRUs. (Note that here the bandwidth dedicated as the control channel is *not* accounted for. The throughput and delay are simply calculated using (2.7) and (2.10) respectively.) Nonhomogeneous networks are considered in [3]; the results are very similar to those of Fig. 1. For $N=5$, the OSDS's curve was calculated using the PIA given in the appendix. For $N=10$ and 50, Conjecture 3.6 was assumed valid and the 'best' policy determined by combining simulation with a simple search procedure. If in fact this method does not result in optimal policies, Fig. 1 indicates they are nearly so.

It is readily observed that the performance of the OSDS protocol is nearly as good as that of the PI scheme. This should not be surprising since their information structures (i.e. information available to each PRU) are very similar. An important difference, however, is that the OSDS information pattern is implementable.

Clearly the OSDS scheme does well because of the wealth of information provided to each PRU. In practice this would be achieved at the cost of dedicating part of the available communication channel bandwidth as a control channel. An important question is how efficiently the OSDS scheme utilizes the spectrum relative to other information patterns when this additional bandwidth is accounted for.

With the aim of addressing this question, consider the OSDS scheme with the control channel embedded in the message channel as in [20]. Specifically, divide the channel into frames consisting of a control slot followed by a message slot. Constrain messages to arrive only at the beginning of a frame (this introduces an additional $\frac{1}{2}$ slot delay) so that $x(t)$ is the state of the network at the beginning of the tth frame. During the control slot each PRU sends its present state $x_i(t)$ and the control $u_i(t)$ which it will implement at the beginning of the message slot. Due to transmission delays these values will not be received until the beginning of the next frame.

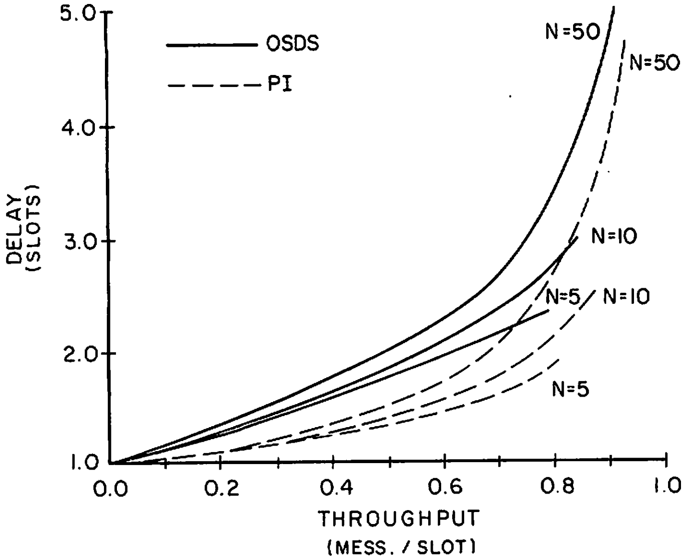

Fig. 1. Comparison of one-step delay sharing and perfect information schemes.

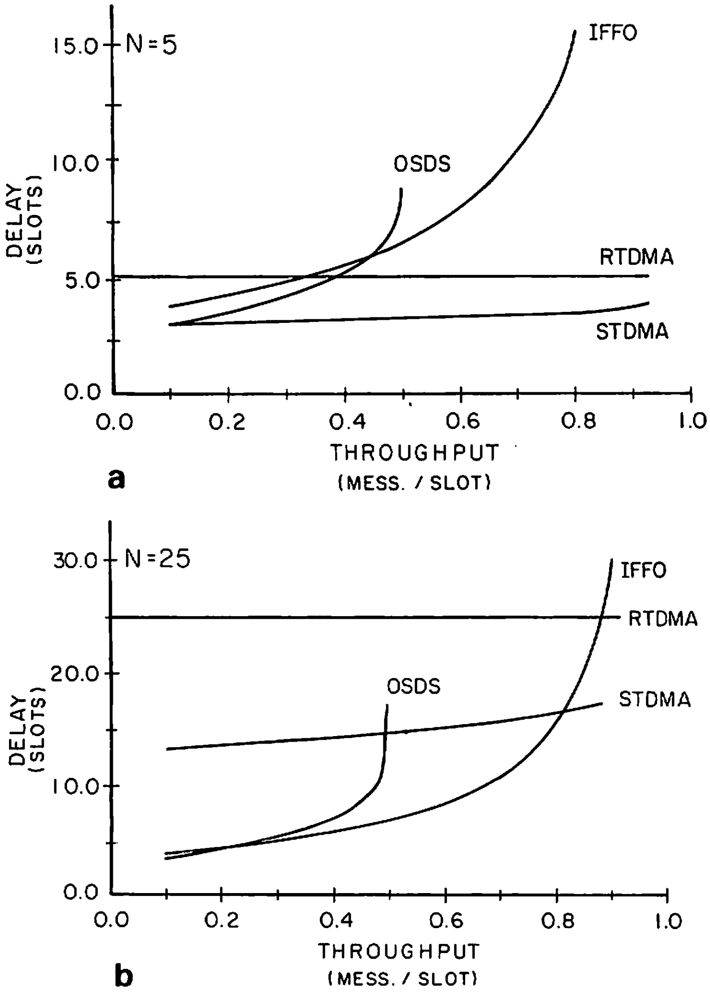

Fig. 2. Comparison of implementable access control schemes: (a) 5 users; (b) 25 users.

In Fig. 2 the delay versus throughput curves of the OSDS scheme, when implemented as above, are shown (Note that here the control channel bandwidth *is* accounted for. This is accomplished by using (2.7) and (2.10) and the following transformation: Throughput → $\frac{1}{2}$ Throughput and Delay → 2 Delay + $\frac{1}{2}$). Also displayed are the corresponding curves for the IFFO, RTDMA, and STDMA schemes. If the network is lightly loaded the OSDS and IFFO schemes perform almost identically. This is because the average number of message slots reserved in each frame in the IFFO scheme is less than one; this results in an information pattern similar to the OSDS case. When the network is more heavily loaded, the IFFO scheme adapts by lengthening the interval between successive reservation slots, resulting in throughputs greater than $\frac{1}{2}$, the upper limit for the OSDS case when the control channel is embedded in the message channel. Of course, for small networks or very heavy traffic the TDMA schemes are superior.

6. Discussion and conclusion

A packet broadcasting system based on a one step delay sharing information pattern has been analyzed. A nonrandomized stationary extreme optimal control policy was shown to exist. Optimal policies could be calculated using a policy iteration algorithm; however, even after reducing the size of the state space via an equivalence relation, this was seen to be impractical when the network consisted of more than just a few terminals. A class of suboptimal policies (conjectured to be optimal) were proposed and shown to give excellent results. Furthermore, choosing the best policy in this class was computationally feasible even for networks consisting of more than 50 terminals.

The results of this analysis were both good and bad. Strong points of the model are (i) that nonhomogeneous networks (i.e. the terminals each have different message arrival rates) are easily handled, (ii) the control law is decentralized, and (iii) excellent delay versus throughput performance was demonstrated. On the negative side is the large amount of control data that must be interchanged among the terminals. When the required bandwidth to accomplish this is accounted for it was shown that an existing reservation scheme could give better overall performance.

The problem that needs to be solved is the following. Let $y(t)$ equal 1 if any terminal successfully sends a message at time t and 0 otherwise. Let the ith terminal observe the pair $(x_i(t), y(t-1))$ at time t. (In practice this may be all that can reasonably be provided to a terminal.) Then based on this information pattern, calculate the policy (if it exists) which maximizes the average throughput. The difficulty is that such an information pattern involves imperfect delayed observations of the state and, furthermore, is not partially nested. The few general results [12,13] known for this class of problems [4] are not computationally feasible. For the model considered in this paper, two variations of the above information pattern do admit solutions by presently available techniques. One possibility is, at time t, to let the terminals observe only $y(t-1)$. Through state augmentation this problem can be cast into the classical framework. But any control policy which does not allow a terminal to observe whether or not it possesses a message would seem to be very impractical and inefficient. A second possibility is to use one step delay sharing so that at time t the ith terminal observes $(x_i(t), x(t-1), u(t-1), y(t-1))$. But one can then easily show that $y(t-1)$ is redundant given $x(t-1)$ and $u(t-1)$, so that this situation reduces to the case considered in this paper. The conclusion is that for the type of model considered in this paper, practical control policies cannot be calculated until further results in the area of nonclassical information patterns are developed.

Acknowledgment

Many of the calculations necessary to prove Theorem 3.5 were carried out by the symbolic manipulation program MACSYMA. The work of the Mathlab group is currently supported, in part, by the United States Department of Energy and the National Aeronautics and Space Administration.

[4] In [4,5] Hajek obtains, by means of some nice approximations, significant results along these lines for the Aloha scheme.

Appendix A

In this appendix the Dynamic Programming Algorithm and the Policy Iteration Algorithm for the OSDS information pattern are given. All notations are assumed to be as defined in the main body of the paper.

A.1. Dynamic Programming Algorithm

Theorem. A.1. *If $p_{ij}(u)$ and $g(\eta_i, u)$ are continuous functions of u for all $0 \leq i, j \leq n$, then the optimal cost to go at time t for the finite horizon problem $J = E\sum_{t=1}^{T} g(x(t), u(t))$ is given recursively by*

$$J_t^*(\eta_i, u(t-1)) = \max_{\substack{v_l \in \Gamma_l \\ l=1,2,\ldots,N}} \left\{ \sum_{j=1}^{n} p_{ij}(u(t-1))\left[g(\eta_j, v^j) + J_{t-1}^*(\eta_j, v^j)\right]\right\} \tag{A.1}$$

for $1 = 1, 2, \ldots, n$, $t = 1, 2, \ldots, T$, $n = 2^N$, *where*
- $J_{T+1}^*(x(T), u(T)) \equiv 0$, $v_l^j \in [0, 1]$ *for fixed* $x(t-1) = \eta_i$,
- $u(t-1)$ *is the control applied by* PR_l *at time t if* $x(t) = \eta_j$,
- $v_l = [v_l^1, \ldots, v_l^n]^{\mathrm{T}}$, $v^j = [v_1^j, v_2^j, \ldots, v_N^j]$,
- $\Gamma_l = \{v_l \mid v_l^j = v_l^k$ *whenever* $x_l(t, \eta_j) = x_l(t, \eta_k)\}$, *and*
- $x_l(t, \eta_i)$ *equals* $x_l(t)$ *when* $x(t) = \eta_i$ (*i.e. l-th component of* η_i).

Remarks. A.2. (a) The restriction to Γ_l is to ensure that the functional $\mu_l(t)$ only depends on $x_l(t)$, $x(t-1)$, $u_1(t-1), \ldots, u_N(t-1)$, and not on $x_k(t)$, $k \neq l$.

(b) In this algorithm, for each fixed $x(t-1) = \eta_i$, $u(t-1)$, the control values v_l, $l = 1, 2, \ldots, N$, for all possible observations at time t are found simultaneously, thus insuring the constraint Γ. As in [18], the optimal cost-to-go J_t^* is a function only of the 'shared information' $(x(t-1), u(t-1))$, which thus represents an 'augmented state at time t' for this dynamic programming algorithm. The same concept is used in the policy iteration algorithm stated later.

(c) Clearly the continuity assumptions of this theorem are satisfied by (2.1) and (2.6).

(d) This is a generalization of the DPA given in [6,7] where the controls are assumed to take values in a finite set. The extension to the present case where $u_i(t) \in [0, 1]$ (or any other compact set) is straightforward and no proof is given here.

A.2. Policy Iteration Algorithm

Let $C = \{\pi = (\mu, \mu, \ldots,) \mid \mu_i(x_i(t), x(t-1), u(t-1)) = \mu_i(x_i(t), [x(t-1), u(t-1)]\}$. It is easy to show (see Remark A.3) that an optimal policy can be chosen to be a member of C and thus that the PIA can be performed on the equivalence classes. The PIA of [6,7] so modified is as follows.

Step 1. Guess an initial stationary policy $\pi^0 \in C$.

Step 2. For the given stationary policy $\pi^k \in C$ compute λ^k and h^k from

$$\lambda^k + h^k([\eta_i, u]) = \sum_{j=1}^{n} p_{ij}(u)\left[g(\eta_j, \bar{u}^j) + h([\eta_j, \bar{u}^j])\right], \quad h^k([0, 0]) = 0 \tag{A.2}$$

where for fixed (η_i, u), $\bar{u}_l^j = u_l^k(x_l(\eta_j), [\eta_i, u])$ and $\bar{u}^j = [\bar{u}_1^j, \bar{u}_2^j, \ldots, \bar{u}_N^j]$. This is a set of $2^N + 1$ linear equations. If $\lambda^k > \lambda^{k-1}$, go to Step 3; else stop since π^k is optimal.

Step 3. Obtain a new (improved) policy $\pi^{k+1} \in C$ satisfying

$$\sum_{j=1}^{n} p_{ij}(u)\left[g(\eta_j, \hat{u}^j) + h^k([\eta_j, \hat{u}^j])\right] =$$

$$= \max_{\substack{v_l \in \Gamma \\ l=1,2,\ldots,N}} \left\{ \sum_{j=1}^{n} p_{ij}(u)\left[g(\eta_j, v^j) + h^k(\eta_j, v^j)\right]\right\} \tag{A.3}$$

where for fixed (η_i, u), $\hat{u}^j=[\hat{u}_1^j, \hat{u}_2^j, \ldots, \hat{u}_N^j]$ and $\mu_l^{k+1}(x_l(\eta_j), [\eta_i, u])=\hat{u}_l^j$, $\pi^{k+1}=(\mu^{k+1}, \mu^{k+1}, \ldots)$.
Return to Step 2.

Remark A.3. (a) Specific sufficient conditions which assure the convergence of the algorithm are discussed in [6,7]. They are analogs of the usual centralized conditions [1].

(b) Suppose as in [6,7] that the function h was defined for each (η_i, u) instead of on the equivalence classes. Then in Step 3 a policy $\pi^{k+1}\in C$ would still obtain the maximum in (A.3) since $p_{ij}(u)=p_{kj}(u)$ if $[\eta_i, u]=[\eta_k, u]$. Now, in Step 2, it is easy to show that if $\pi^k\in C$, then $h(\eta_i, u)=h(\eta_k, \hat{u})$ if $[\eta_i, u]=[\eta_k, \hat{u}]$. Thus the PIA can be performed only on the equivalence classes as given above.

References

[1] D. Bertekas, Dynamic Programming and Stochastic Control (Academic Press, New York, 1976).
[2] D. Bertsekas and S. Shreve, Stochastic Optimal Control: The Discrete Time Case (Academic Press, New York, 1978).
[3] J. Grizzle, An analysis of centralized and decentralized control strategies for multiaccess broadcast networks, Masters Thesis, Electrical Engineering Dept., The University of Texas at Austin, Austin, TX, 1980.
[4] B. Hajek, Dynamic decentralized estimation and control in a multi-access broadcast channel, 19th IEEE CDC Proceedings (12) (1980) 618–623.
[5] B. Hajek and T. vanLoon, Decentralized dynamic control of a multi-access broadcast channel, Preprint.
[6] K. Hsu, On discrete time centralized and decentralized estimation and stochastic control, Ph. D. Dissertation, Electrical Engineering Dept., The University of Texas at Austin, Austin, TX, 1979.
[7] K. Hsu and S. Marcus, Decentralized control of finite state Markov processes, IEEE Trans. Automat. Control 23 (2) 1982 426–431.
[8] L. Kleinrock, Principles and lessons in packet communications, Proc. IEEE 66 (11) (1978) 1320–1329.
[9] L. Kleinrock, Queueing Systems: Vol. 1, Theory (Wiley, New York, 1975).
[10] L. Kleinrock and Y. Yemini, An optimal adaptive scheme for multiaccess broadcast communications, Conf. Rec. ICC '78, Toronto (1978) 7.2.1–7.2.5.
[11] S. Lam and L. Kleinrock, Packet-switching in a multiaccess broadcast channel: Dynamic control procedures, IEEE Trans. Comm. 23 (9) (1975) 891–904.
[12] J. Levine, Principle d'optimalité et principe de séparation en contrôle stochastique à information incomplète non classique, C.R. Acad. Sci. Paris 292 (1981) 877–880.
[13] J. Levine, Incomplete information in differential games and team problems: Necessary and sufficient optimally conditions, 8th I.F.A.C. World Congress, Kyoto, 1981.
[14] S. Rappaport, Demand assigned multiple access systems using collision type request channels: Traffic capacity comparisons, IEEE Trans. Comm. 27 (9) (1979) 1325–1331.
[15] L. Roberts, The evolution of packet switching, Proc. IEEE 66 (11) (1978) 1307–1313.
[16] I. Rubin, Access-control disciplines for multi-access communication channels: Reservation and TDMA schemes, IEEE Trans. Inform. Theory 25 (9) (1979) 516–536.
[17] F. Schoute, Decentralized control in packet switched satellite communication, IEEE Trans. Automat. Control 23 (2) (1978) 362–371.
[18] P. Varaiya and J. Walrand, On delayed sharing patterns, IEEE Trans. Autom. Control 23 (3) (1978) 443–445.
[19] P. Varaiya and J. Walrand, Decentralized control in packet switched satellite communication, IEEE Trans. Automat. Control 24 (5) (1979) 794–796.
[20] J. Wieselthier and A. Ephremides, A new class of protocols for multiple access in satellite networks, IEEE Trans. Automat. Control 25 (5) (1980) 865–879.
[21] Y. Yemini and L. Kleinrock, On a general rule for access control or, Silence is Golden, Preprint.