

# UC Irvine

## UC Irvine Previously Published Works

### Title

A continuous state space model of multiple service, multiple resource communication networks

### Permalink

<https://escholarship.org/uc/item/63k7t9m1>

### Journal

IEEE Transactions on Communications, 43(2/3/4)

### Author

Jordan, Scott

### Publication Date

1995-02-01

### DOI

10.1109/26.380065

Peer reviewed

# A Continuous State Space Model of Multiple Service, Multiple Resource Communication Networks

Scott Jordan, *Member, IEEE*

**Abstract** - The merging of telephone and computer networks is introducing multiple resources into networks, and information is becoming increasingly distributed across the network. Related services are being integrated onto a single network rather than being offered on separate uncoordinated networks. We focus upon communication networks that integrate multiple services using multiple resources.

In previous work, such networks have been modeled by multi-dimensional Markov chains with product form distributions. In this paper, we approximate the distribution on the original discrete state space by a similar product form distribution on a continuous state space. We consider access control of such a system and prove that the resulting optimal coordinate-convex control policy is convex. Based on this result, we suggest an algorithm for finding a near-optimal policy for the discrete problem that has much less complexity than existing methods for finding optimal or near-optimal policies.

## I. INTRODUCTION

In this paper, we focus upon communication networks that integrate multiple services using multiple resources. We investigate resource allocation strategies and try to capture the nature of controlling such a system. In particular, we present a continuous state space model and characterize the optimal coordinate convex access policy for this model. We show that this characterization allows for simpler numerical computation of a near-optimal policy than existing methods.

This work is motivated by several trends in networks. The merging of telephone and computer networks is introducing multiple resources into networks, and information is becoming increasingly distributed across the network. Related services are being integrated onto a single network rather than being offered on separate uncoordinated networks.

These trends are made possible by the availability of fiber and of inexpensive electronic storage, and by the introduction of greater intelligence into the signaling system. Furthermore, these trends are made profitable by the proliferation of desktop computers and the increased demand for better information transfer.

Paper approved by Jim F. Kurose, the Editor for Networking Systems of the IEEE Communications Society. Manuscript received: March 8, 1991; revised August 25, 1993, and June 30, 1994. This work was supported by the National Science Foundation under grants ECS-8719779 and ECS-8719298, the California State MICRO Program, and the Fannie and John Hertz Foundation. This paper was presented in part at IEEE ICC '91.

S. Jordan is with the Department of Electrical Engineering & Computer Science, Northwestern University, Evanston, IL 60208-3118 (scott@eecs.nwu.edu)

IEEE Log Number 9410079.

Proposals for implementing services in these *multiple service, multiple resource (MSMR)* networks abound. A few examples of these *services* might be electronic/voice mail, mixed media telephone calls, video conferencing, distributed databases, hypertext systems, electronic catalogues, electronic yellow pages, and collaborative editors.

Our premise is that each *service* relies upon a number of underlying *resources* in the network. Examples of these *resources* might be communication links, databases, switches, storage devices, special purpose hardware and software. Although the precise meaning of "service" and "resource" and the relationship between them is a topic for future research, we assume in this paper that we have identified each service and the set of resources on which it depends.

Integrated services will share resources both for functionality and to decrease cost. Since these resources are limited, there will be interaction among the services. What types of interaction might we see? If you are the manager of a multiple service, multiple resource system, what requests for service do you accept? Based on what? If you base these decisions on maximizing revenue, what prices do you charge? And what resources should you acquire? The purpose of this research effort is to address such *resource allocation problems*.

In [1], we investigated the nature of this interaction, and in [2], we investigated control of a MSMR system. In this paper, we introduce a continuous state space version of the model used in the previous two papers in order to obtain a stronger characterization of the optimal access control policy and to suggest a simpler algorithm to numerically obtain it.

The MSMR model, and a simpler related multiple service single resource (MSSR) model, were initially introduced and shown to have a product form solution under a wider variety of scenarios in Aein [3] and Kaufman [4].

Research pertinent to searches for optimal or near-optimal policies for these systems fall into two categories: that concerned with efficient numerical evaluation of a single policy, and that concerned with algorithms to search through various candidates.

For a single policy, the straightforward approach of solving the balance equations to find the stationary distribution of the Markov chain has a complexity that is exponential in the number of services. A good deal of recent literature has made significant progress on reducing this complexity, for MSSR systems, if the policy is of a threshold type (and hence convex). See Kaufman [4], Kraimeche [5], Ross [6], and Tsang [7] for various recursion schemes, and Dziong [8] and Mitra [9] for

approximations. In addition, Whitt [10] and Kelly [11] investigate reduced-load approximations for MSMR systems.

Research on characterization of optimal control policies and algorithms to find them has been more scarce. Ross [12] characterizes optimal policies for a system with 2 service types. Zachary [13], Ross [14], and Key [15] suggest dynamic programming approaches. Kelly [11] suggests approximating a MSMR system as a collection of MSSR systems.

In [2], we characterized the optimal coordinate convex policy for a MSMR system, and noted, as had earlier researchers, that it is not guaranteed to be a convex policy. This lack of convexity complicates both numerical evaluation of a single policy, and the search for the optimal policy. Section II displays the model originally presented in [1] and the optimal control policies derived in [2]. Section III proposes a continuous state space model and proves that the optimal access control for this model is convex. In sections IV and V, we discuss algorithms suggested by the theoretical results and present a few examples.

## II. THE MSMR MODEL

Consider the following model for resource allocation in multiple service multiple resource (MSMR) communication networks, originally presented in [1]:

Consider a system that offers  $n$  types of services using  $m$  types of resources. Each service requires a set of resources, dependent upon the service type, to process. If these resources are available then the system manager may choose to accept a service request, and then processing starts immediately; if the necessary resources are unavailable, or if the system manager denies the request, then the request is lost to the system.

Service requests arrive as independent Poisson processes. Each request occupies each resource that it needs for the same amount of time, and releases these resources simultaneously upon service completion. This amount of time is exponentially distributed, and independent of other service times.

Adopt the following notation:

$A$  = a  $m \times n$  matrix, with column  $i$  indicating the number of each of  $m$  resources used by service  $i$ .

$B$  = a vector of length  $m$  indicating the number of each resource type in the system.

$\lambda = (\lambda_1, \dots, \lambda_n)$ , the rates of incoming service requests.

$\mu = (\mu_1, \dots, \mu_n)$ , the rates of service.

$\rho = (\rho_1, \dots, \rho_n)$ , the loads, given by  $\rho_i = \lambda_i / \mu_i$ .

$L = (L_1, \dots, L_n)$ , the rates of *accepted* service requests (throughput).

$x = (x_1, \dots, x_n)$ , the state of the system, where  $x_i$  is the number of type  $i$  requests being processed.

$Z = \{x \mid Ax \leq b, \text{ i.e. } x \text{ can be simultaneously processed with available resources}\}$ .

$F_i = \{x \mid x \in Z \text{ but } (x_1, \dots, x_i + 1, \dots, x_n) \notin Z\}$ , the full set w.r.t. service type  $i$ .

$E_i = \{x \mid x \in Z \text{ but } (x_1, \dots, x_i - 1, \dots, x_n) \notin Z\}$ , the empty set w.r.t. service type  $i$ .

$\pi(x)$ , the steady state probabilities.

$N = \{1, \dots, n\}$ , the set of all service types.

$x \uparrow I = (y_j), y_j = \begin{cases} x_j, j \notin I \\ 0, j \in I \end{cases}$  for  $I \subset S$ , a projection function.

$C_Z(x \uparrow I) = \{y \in Z \mid y \uparrow I = x \uparrow I\}$ , an  $|I|$  dimensional cross section of  $Z$ .

$r_i$  = rate of revenue generated by servicing request type  $i$ , per unit of time.

$r(x)$  = rate of revenue generated while in state  $x = \sum_i r_i x_i$

$R$  = the average revenue per time unit generated by the system =  $E r(X) = r(EX)$ .

$Z$  is *coordinate convex* (c.c.) = if  $x \in Z$  and  $x_i \geq 1$  then  $(x_1, \dots, x_{i-1}, \dots, x_n) \in Z$

$E_V r(X) = E[r(X) \mid X \in V]$  where  $V \subseteq Z$

$V$  is *annexable* to  $Z = V \cap Z = \emptyset$  and  $V \cup Z$  is c.c.

$V$  is *removable* from  $Z = V \subset Z$  and  $Z - V$  is c.c.

$ss(V) = \{x \notin V \mid 0 \leq x_i \leq v_i \forall i \text{ for some } v \in V\}$ , the supporting set of  $V$ .

Our assumptions regarding the arrival and departure processes give us a Markov chain on state space  $Z$  with transition rates:

$$r_{xy} = \begin{cases} \lambda_i, & \text{if } x \notin F_i \text{ and } y = (x_1, \dots, x_i + 1, \dots, x_n) \\ x_i \mu_i, & \text{if } x \notin E_i \text{ and } y = (x_1, \dots, x_i - 1, \dots, x_n) \\ 0, & \text{else} \end{cases}$$

Assume that service completion is never blocked. This implies that the state space  $Z$  is coordinate convex. We will generally assume  $Z = \{x \mid Ax \leq b\}$ , so that  $Z$  is also convex.

The uncontrolled Markov chain is time reversible with stationary distribution:

$$\pi(x) = \pi(0) \prod_{i=1}^n \frac{\rho_i^{x_i}}{x_i!} \quad \pi(0) = \frac{1}{\sum_{x \in Z} \prod_{i=1}^n \rho_i^{x_i} / x_i!} \quad (1)$$

We now assume that the system manager can choose to deny a service request even if the corresponding resources are available. Since resources are limited, accepting a request of one type may preclude the possibility of accepting a request of another type in the near future. We restrict ourselves to control policies that are coordinate convex (c.c.). In this class, admission decisions depend on *the state the system would enter if the request is granted*. This is equivalent to *restricting the state to some subset of the original space*. Since service completion can not be realistically blocked, this subset, like the original state space, must be coordinate convex. This class has the property that the controlled system is the restriction of the uncontrolled time reversible Markov chain to a subset, and is thus itself time reversible, maintaining its product form stationary distribution.

Our goal is to maximize expected revenue. Our objective is thus to choose the c.c. subset  $Z^* \subseteq Z$ , that maximizes  $E[r(X) \mid X \in Z^*]$ .

Consider the removal of a set  $V$  from  $Z$ . Since the Markov chain is time reversible, the removal of a set  $V$  from  $Z$  affects the distribution on  $Z - V$  only through the normalization constant  $\pi(0)$ . If  $E_V r(x) < R$ , then removing  $V$  would proportionally increase the rewards from other states because the new normal-

ization constant would rise. Furthermore, the removal of  $V$  from  $Z$  increases the average revenue on  $Z$  if and only if  $E_V r(x) < R$ . We can therefore characterize an optimal c.c. policy as a subset of  $Z$  to which nothing above average can be added and nothing below average removed [16], i.e. a c.c. set  $Z^* \subseteq Z$  is optimal iff

$$\begin{aligned} \exists V \subseteq Z \ni V \text{ is annexable to } Z^* \ \& \ E_V r(x) > R^* \\ \exists V \ni V \text{ is removable from } Z^* \ \& \ E_V r(x) < R^* \end{aligned}$$

where  $R^* = E_{Z^*} r(X)$

In [2], we characterized the optimal c.c. policy for our MSMR model:

**Theorem 2.1:** There exist a set of constants  $\{c_j^k, I \subset N\}$  and a set of constants  $\{\alpha_{i,k,I}\}$  such that the optimal coordinate convex policy,  $Z^*$ , for the MSMR model, can be represented as:

$$x \notin Z^* \text{ iff } x \notin Z \text{ or } \exists k \ni \sum_{i \in I} \alpha_{i,k,I} x_i \geq c_j^k \ \forall I \subset N \quad (2)$$

### III. A CONTINUOUS STATE SPACE MODEL

This characterization is weak, principally because the optimal access control policy is *not* guaranteed to be a convex policy. This lack of convexity makes numerical evaluation of policies more difficult, and makes the search for an optimal policy extraordinarily difficult. For convex policies, (2) has a much simpler form:

$$x \in Z^* \text{ iff } x \in Z \ \& \ x \uparrow I \in G_I \ \forall I \subset N \quad (3)$$

where  $G_I$  are convex sets.

Without any characterization, the optimal policy is found by a exhaustive search among all c.c. subsets of the state space  $Z$ . With (2), the set of candidate policies is reduced, but still extremely large. If the optimal policy is also known to be convex, then the set of candidate policies is much smaller.

The optimal c.c. policy of a similar MSSR queueing problem has been shown to always be convex [2]. We are thus lead to investigate why convexity is not guaranteed in the MSMR model. Consider a system with 2 services and 1 resource in which service #1 requires 2 of the common resource and service #2 requires 3 (Fig. 1).

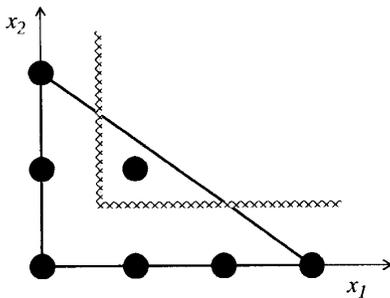


Fig. 1. A MSSR system with a non-convex optimal c.c. policy

Assume services 1 and 2 generate rates of revenue of 2 and 3, respectively and that the loads are proportional to the revenues but very high. State (1,1) generates a rate of revenue of 5. The overall average rate of revenue in the system, however, is approximately 6, since almost all the probability is in states

(0,2) and (3,0). Thus the optimal control policy would exclude state (1,1), namely  $Z^* = Z - \{(1,1)\}$ . Note that  $Z^*$  is not convex.

The convexity of the optimal policy is caused by the non-concavity of expected revenue on cross sections of type 2,  $E_{C(x \uparrow \{2\})} r(X)$ . It is *not* true that

$$\begin{aligned} E_{C[(2,0) \uparrow \{2\}]} r(X) \geq \\ \frac{1}{2} E_{C[(1,0) \uparrow \{2\}]} r(X) + \frac{1}{2} E_{C[(3,0) \uparrow \{2\}]} r(X) \end{aligned}$$

This is in turn caused by the non-concave height of the state space with respect to  $x_1$  (Fig. 2), which is due to the non-integer multiple of resource usage.

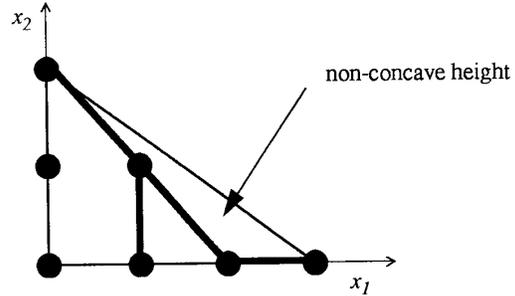


Fig. 2. Non-concavity of  $E_{C(x \uparrow \{2\})} r(X)$

This lack of concavity is a local effect, and of magnitude given by the granularity of the resource usage. In systems which can accommodate a moderate number of every service, this should have a small effect on the optimal revenue. We therefore propose approximating the discrete state space by a continuous space, and consequently approximating the discrete product form distribution by a continuous distribution:

$$\pi(x) = K \prod_{i=1}^n f_i(x_i) \quad f_i(x_i) \text{ continuous on } x_i \geq 0 \quad (4)$$

The state space is now a continuous region given by  $Ax \leq b$ . The state is  $x$ , a vector of length  $n$  of non-negative real numbers;  $A$  is a  $m \times n$  matrix of non-negative real numbers representing resource usage for each service type;  $b$  is a vector of length  $m$  of positive real numbers representing total numbers of each resource type.

We require that each additional resource added to the system produces a non-negative but decreasing return to the optimal revenue. This is equivalent to the following concavity property on parallel cross sections of the optimal subset of the state space [17]:

**Property DR:** For any two cross sections  $C_1$  and  $C_2$ ,  $C_1 = C_{Z^*}(x \uparrow I)$  and  $C_2 = C_{Z^*}((x + \beta e_j) \uparrow I), j \in I$ ,  $R_{[\alpha C_1 + (1-\alpha)C_2]} \geq \alpha R_{[C_1]} + (1-\alpha) R_{[C_2]} \ \forall i \in I$ , where  $\alpha C_1 + (1-\alpha)C_2 = C_{Z^*}((x + (1-\alpha)\beta e_j) \uparrow I)$  and  $e_j$  is a vector with all zeroes except for a 1 in the  $j^{\text{th}}$  position.

The optimal c.c. access policy for the continuous model given by (4) satisfying (DR) can now be shown to be convex, and hence characterized in the form of (3). Furthermore, we can characterize the convex sets  $G_I$  that define the policy.

Adopt the following notation:

$$M_I(R) = \{x \in Z \mid r(E_{C_Z(x \uparrow I)} X) > R\}$$

$$BM_I(R) = M_I(R) + ss(M_I(R))$$

$M_I(R)$  is the set of states in "I" type cross-sections generating above average rates of revenue.  $BM_I(R)$  consists of  $M_I(R)$  and those states supporting  $M_I(R)$ .  $BM_I(R)$  is thus coordinate convex.

**Theorem 3.1:** If  $X$  has a distribution on  $Ax \leq b$  given by (4) satisfying (DR), then the optimal c.c. subset  $Z^* \subseteq Z$  is defined by:

$$x \in Z^* \text{ iff } x \in Z \text{ and } x \uparrow I \in G_I^* \forall I \subset N \quad (5)$$

where  $G_I^* = BM_I(R^*) \uparrow I$  and  $R^* = E_{Z^*} r(X)$

The proof can be found in the Appendix.

Theorem 3.1 implies that we should cut back on services of lesser value (the  $N-I$  services) to get increased throughput of some services of high value (the  $I$  services). More precisely, it states that if extreme points of the state space generate a below average rate of revenue, then we can do better by restricting the system from entering the regions near these points. We investigate the algorithmic implications in the next section.

#### IV. ALGORITHMS

##### A. An Outline

Consider the task of finding the optimal c.c. access control policy for a particular MSMR system specified by a service to resource map  $A$  and a resource vector  $b$ , with stationary distribution given by (1). If no characterization of the optimal c.c. policy is known, we would use the following algorithm:

**Algorithm A:**

- 1) Find all coordinate convex subsets of  $Z$ .
- 2) Find the average revenue generated by each subset.
- 3) Choose the subset that generates the highest average revenue.

The number of coordinate convex subsets, however, becomes prohibitively large when the number of services is greater than one and the number of states is still relatively small.

In the previous section, we have found that the optimal c.c. access control policy for the continuous MSMR model must be convex, while the optimal c.c. policy for the discrete MSMR model need not be. Furthermore, this convexity is lost due to granularity in resource usage. For systems that accommodate moderate numbers of each service type, we therefore propose restricting our search for optimal policies to *convex* c.c. policies, i.e. of the form (5). We could use the following algorithm:

**Algorithm B:**

- 1) Approximate the discrete product form distribution (1) by a continuous system (4), with:

$$f_i(x_i) = \frac{\rho_i^{x_i}}{\text{Gamma}(x_i + 1)} \quad (6)$$

- 2) Find the optimal c.c. policy for this system by repeatedly removing type "I" sets that produce below average revenue from the exterior of the state space:

*outer loop:*

*calculate average revenue*

*find all extreme points of the state space*

*introduce a new constraint tangent to the state space at the extreme point with the lowest revenue*

*inner loop:*

*calculate average revenue on the intersection of the constraint with the state space*

*lower the constraint by control\_step*

*until average revenue on constraint exceeds*

*average revenue of state space*

*until no constraint with below average revenue can be found.*

We investigate the complexity of these algorithms in the next subsection.

##### B. Complexity

Define the following quantities:

$z$  = number of states in the state space  $Z$ .

$q$  = maximum of  $x_i$  such that  $x \in Z$ , for square state spaces.

$s$  = number of c.c. subsets of  $Z$ .

$sc$  = number of convex c.c. subsets of  $Z$  considered in algorithm B.

$CA$  = complexity of algorithm A.

$CB$  = complexity of algorithm B.

The complexity of algorithm A is proportional to the number of candidate c.c. policies:

$$CA \sim s \text{ (complexity of evaluating each subset)}$$

The number of such subsets,  $s$ , is dependent upon the structure of the state space. In general<sup>1</sup>, we can only say that  $z \leq s \leq 2^z - 1 + 1$ . Consider, however, a system with 2 service types and a square state space. For this system, it can be shown that<sup>2</sup>:

$$s = 2^{O(q)} = 2^{O(\sqrt{z})}$$

The complexity of evaluating each subset depends on the amount of recursion and reuse of information used in finding the first moments of the Markov chain. At most, this is  $O(zn)$ . The complexity of algorithm A is therefore dominated by  $s$ .

For algorithm B:

$$CB \sim sc \text{ (complexity of evaluating each subset)}$$

The number of policies examined,  $sc$ , is reduced from  $s$  for two reasons. First, only *convex* c.c. subsets are legal policies. Second, in a continuous product form space, the difference in average revenue produced by two simultaneous infinitesimal changes to the state space is equal to the sum of the differences if these two changes were taken separately. Boundary changes (removable sets) can thus be considered individually, rather than as pairs.

For the same 2 service type system considered above, there are  $(q+1)^2$  convex policies satisfying (3), and algorithm B con-

1. A one dimensional system would have  $s=z$ . The  $z-1$  dimensional system  $Z = \{x \mid \sum_{i=1}^{z-1} x_i \leq 1\}$  would have  $s=2^{z-1}+1$ .

2. For this system,  $z=q^2$ . The complexity is found by simple combinatorial arguments.

siders at most  $2*(q+1)/control\_step$  removable sets. If we set  $control\_step = 1$ :

$$sc = O(q) = O(\sqrt{z})$$

Furthermore, the complexity of evaluating each subset can be reduced by using numerical integration on the continuous state space, rather than sums on the discrete state space. If we choose an integration step size equal to the standard deviation of  $x_i = O(q^{1/2})$ , then the complexity of evaluating each subset is further reduced.<sup>1</sup> This component, however, is still likely to be exponential in  $n$ .<sup>2</sup>

In our versions of these algorithms, which were not coded for efficiency, we found that algorithm A had a running time that was exponential in the number of states in the system, and progressed from infinitesimal to impractical<sup>3</sup> at about 30 states (and 3 service types). We found that algorithm B had a running time that was exponential in the number of service types, and progressed from infinitesimal to impractical at about 7 service types (and millions of states).

### V. EXAMPLES

In this section, we discuss access control in a system with 2 services X and Y and 2 resources A and B. Assume service X occupies 2As and 1B and service Y occupies 1A and 2Bs. Assume there are 60 As and 60 Bs in the system. The state space for this system is shown in Fig. 3.

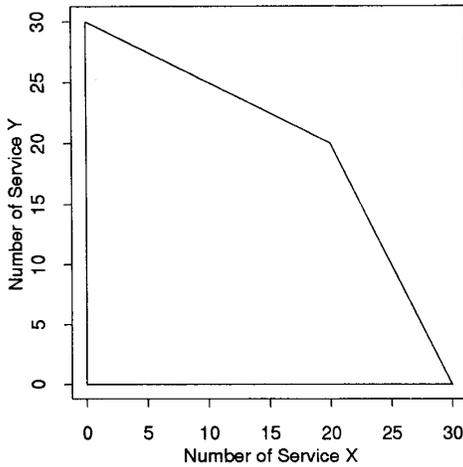


Fig. 3. The state space.

The example was chosen to include a moderate number of states, so that moments could be calculated by either the dis-

1. If the complexity of evaluating each subset is proportional to the number of elements in the subset, then this reduction is from  $O(q^n)$  to  $O(q^{n/2})$ . This suggests that numerical integration, or equivalently discrete state grouping, increases the dimension of practical problems by a factor of 2.

2. Use of recursion techniques proposed for MSSR systems may significantly reduce this complexity if they can be extended to MSMR systems.

3. Since these complexities are exponential, running times quickly increased from sub-second to days as the appropriate parameter increased past the noted value.

crete or continuous algorithms, but so that optimal control policies could only be calculated in reasonable time by the continuous algorithms.

#### A. Blocking Probabilities and Average Revenue

The majority of the work in the numerical integration routine is in calculating the density at each integration point. Therefore, rather than calling the integration routine separately to find each moment, we chose to have the integration routine keep one running sum for the normalization constant, plus one running sum for each first moment. One call thus returns all of the first moments. This suffices to find throughputs, blocking probabilities, and average revenue<sup>4</sup>:

$$L_i = \mu_i EX_i$$

$$P(F_i) = 1 - EX_i / \rho_i$$

$$R = \sum_i r_i EX_i$$

If we chose to also calculate second moments, then we could have found sensitivities using these results and others from [1]:

$$\frac{\partial L_i}{\partial \lambda_j} = \begin{cases} \frac{\mu_i}{\lambda_j} \text{cov}(x_i, x_j), & \text{if } i \neq j \\ \frac{\mu_i}{\lambda_i} \text{var}(x_i), & \text{if } i = j \end{cases}$$

To help illustrate the use of first moments and to demonstrate the need for control, consider our 2-service 2-resource example with  $r=(1,3)$  and with a load of  $\rho_Y=20$ . What happens as we vary  $\rho_X$ ?

At low  $\rho_X$ , there is very little blocking of either service. As the load on service X starts to increase,  $EX$  increases with it, and  $EY$  decreases slightly since services X and Y are substitutes (Fig. 4). As the probability of blocking of service Y becomes significant, the trade-off becomes more severe and  $EY$  drops precipitously.

Since  $r_Y > r_X$ , we should be concerned about this drop. Indeed, in the uncontrolled system, average revenue (Fig. 5) increases only until  $\rho_X \approx 23$  and then drops off with increasing  $\rho_X$ .

Allowing control of the system corrects this problem. Since  $r_Y > r_X$ , the optimal control for the continuous model is simply a restriction on the number of service X in the system. With  $\rho_Y=20$ , as  $\rho_X$  increases, this restriction becomes more severe, eventually allowing only states with  $X \leq 20$  (Fig. 4). As we would demand of the controlled system, revenue no longer dips at high  $\rho_X$  (Fig. 4).

#### B. Control Policies

Consider the same 2 service type system with  $r=(1,3)$ ,  $\rho_X=30$ , and  $\rho_Y=20$ . How sensitive is the system to the amount

4. It should be noted that use of recursion techniques proposed for MSSR systems may significantly reduce this complexity if they can be extended to MSMR systems.

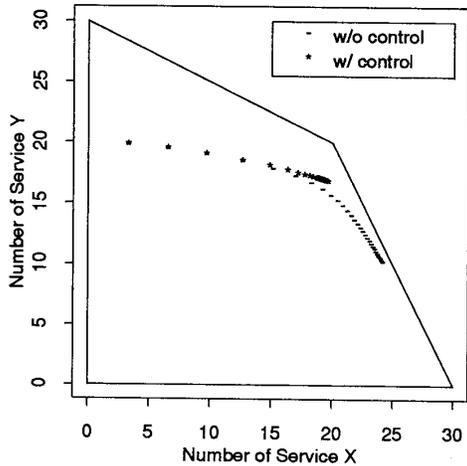


Fig. 4. Increasing load on X: state space

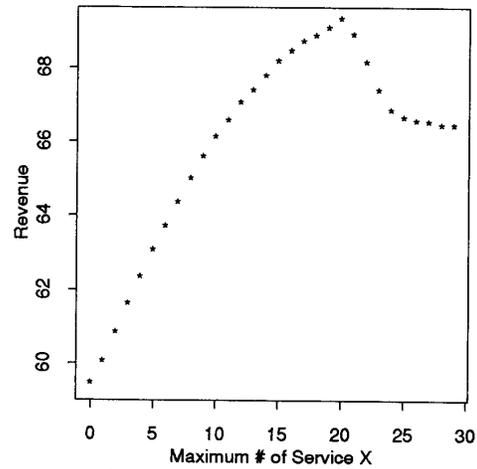


Fig. 6. Varying control of service X: low revenue ratio

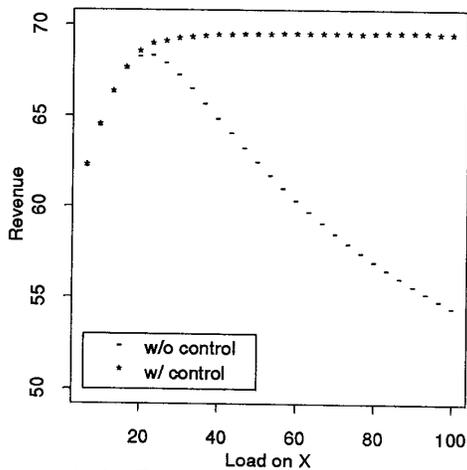


Fig. 5. Increasing load on X: revenue

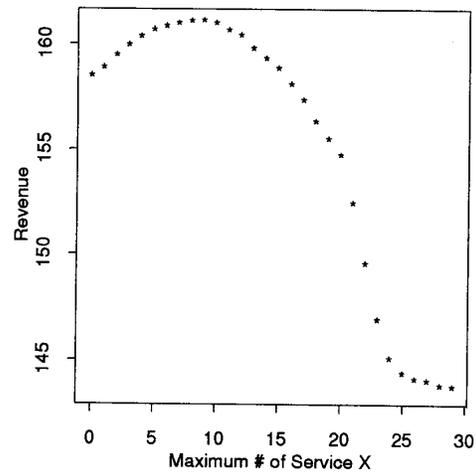


Fig. 7. Varying control of service X: high revenue ratio

of control exercised? As we vary the maximum number of service  $X$  allowed, revenue moves as shown in Fig. 6. The vertical difference at  $\rho_X=30$  on Fig. 5 shows up here as the difference between the peak value of  $R$  and the value at maximum  $X=30$ . Here the peak is prominent and the control algorithm can find it easily in 3 iterations of the outer control loop.

The optimal control point becomes much more sensitive at slightly higher  $r_Y/r_X$ . If we set  $r=(1,8)$ , the peak is rounder (Fig. 7). The optimal control is stronger, but is also tougher to find.

## VI. PARTING THOUGHTS

In this paper, we have approximated the MSMR product form stationary distribution by a similar distribution on a continuous space. We have shown that the resulting optimal coordinate convex control policy is convex, and presented a characterization of it. We have presented an algorithm to find a near-optimal control policy for the original MSMR problem that is less complex than existing methods.

This analysis, however, is limited by the requirement that resource usage is simultaneous. A service may not need two types of resources simultaneously, but first one and then the other. This latter type of sharing can not be addressed by the models in this paper, but instead require a model cognizant of changes in time.

## ACKNOWLEDGMENTS

The author would like to thank P.P. Varaiya for guidance, and the reviewers for suggestions to make this paper more readable.

## APPENDIX PROOFS

The proof falls into three parts. First, we show if the optimal policy is convex, then it must exclude all states excluded from  $Z^*$ . Second, we show if the optimal policy is convex, then it must include all states included in  $Z^*$ . Finally, we show the optimal policy must be convex.

*Lemma 1:* If  $X$  has a distribution on  $Ax \leq b$  given by (4), and the optimal policy is of the form of (3), then  $G_I \subseteq G_I^* \forall I$ , where  $G_I^*$  is given by (5).

*Proof:* The proof proceeds by showing that removed states form removable sets of type "I" cross-sections generating below average revenue.

Consider  $G_I^* = BM_I(R^*) \uparrow I$  for some  $I \subset N$ . The set of states in "I" type cross sections that have been removed is:

$$V_I = \{x \in Z \mid x \uparrow I \notin G_I^*\}$$

By construction,  $V_I$  is removable from  $Z$ . Furthermore, the expected rate of revenue on  $V_I$  is:

$$E_{V_I} r(X) = \int_{V_I} r(E_{C_Z(x \uparrow I)} X) \pi(C_Z(x \uparrow I)) d(x \uparrow I)$$

and  $r(E_{C_Z(x \uparrow I)} X) < R^*$  on  $V_I$  since the cross section contains none of the above average type "I" cross sections, namely  $V_I \cap M_I(R^*) = \emptyset$ .

Thus  $Z^*$  should not contain  $V_I$ , i.e.  $Z^* \not\supseteq V_I \forall I$ , or namely  $G_I \subseteq G_I^* \forall I$ .  $\square$

*Lemma 2:* If  $X$  has a distribution on  $Ax \leq b$  given by (4) satisfying (DR), and the optimal policy is of the form of (3), then  $G_I \supseteq G_I^* \forall I$ , where  $G_I^*$  is given by (5). Furthermore,  $G_I$  is convex.

*Proof:* The proof proceeds by showing that the concavity property (DR) implies that any additional removable sets of type "I" would generate an above average rate of revenue.

Suppose that there exists an  $I$  such that  $G_I \subset G_I^*$ . Define

$$A = \text{boundary}(G_I) - \text{boundary}(G_I^*)$$

and

$$A' = \text{boundary}(G_I^*) - \text{boundary}(G_I).$$

Now  $R[x \uparrow I \in A] \geq R^*$ ; any average revenue less than  $R^*$  would justify removal of the set. Similarly,  $R[x \uparrow I \in A'] \geq R^*$ , from the construction of  $G_I^*$ . Now since  $r(\cdot)$  is linear, i.e.  $R[\text{set}] = r(E_{\text{set}} X)$ , using property (DR) we can conclude that

$$R[\text{convex hull of } (x \uparrow I \in A) \& (x \uparrow I \in A')] \geq R^*.$$

But since this convex hull is annexable, this would imply that  $G_I$  was not optimal.

Similarly, we can conclude that  $G_I$  must be convex, since a similar construction could be made by setting  $G_I^*$  to be the convex hull of  $G_I$ .  $\square$

*Theorem 3.1:* If  $X$  has a distribution on  $Ax \leq b$  given by (4) satisfying (DR), then the optimal c.c. subset  $Z^* \subseteq Z$  is convex and defined by (5).

*Proof:* The proof proceeds by showing:

1) Any set  $V$  removable from  $Z^*$  must contain both above average and below average states, and thus must contain average states.

2) The set of average states of  $Z^*$  is contained in the union of the  $M_I$ .

3) Any such  $V$  must contain sets in the  $M_I$  that generate a below average rate of revenue.

4) (DR) concavity implies that no such sets exist.

This proof of the convexity of  $Z^*$  relies upon lemmas 1 and 2. We suspect that there is probably a simpler, shorter proof that does not rely upon these but instead shows directly that any optimal  $Z^*$  must be convex.

Consider  $Z^*$  as defined in (5). If  $Z^*$  is not optimal, then there exists a set  $V$  removable from  $Z^*$  with  $E_{V_I} r(X) < R^*$ .

By construction, the  $Z^*$  proposed in (5) has already stripped off all removable sets on the exterior of  $Z$  with below average rates of revenue. Therefore, if  $V$  is removable and if  $E_{V_I} r(X) < R^*$ , then  $V$  must contain both sets on the exterior of  $Z^*$  generating above average rates of revenue and sets further in  $Z^*$  generating below average rates of revenue. We will use the concavity property (DR) to show that the above average states in  $V$  outweigh the below average states in  $V$  and result in  $E_{V_I} r(X) \geq R^*$ .

Consider the set of states generating an average rate of revenue:

$$T = (r(x) = R^*) \cap Z^*$$

By continuity, if  $V$  contains both below average and above average states,  $V$  must contain average states, namely  $V \cap T \neq \emptyset$ . We will show that  $T$  is contained in  $M_I$  regions and that no such  $V$  can exist in  $M_I$  regions.

Assume for now that  $T \subseteq \cup_{I \subset N} M_I$ . This will be shown later. If this is so, then  $V \cap T \neq \emptyset$  and  $V$  removable implies that  $V$  must contain sets in the  $M_I$  that generate a below average rate of revenue i.e.

$$\exists V_I \subseteq M_I \ni E_{V_I} r(X) < R^*$$

Let  $V_I$  be the largest such set, i.e.  $ss(M_I - V_I) \cap M_I = \emptyset$

Now on  $M_I$ ,  $r(E_{C_Z(x \uparrow I)} X) > R^*$ . But  $V_I$  can be represented as an integral over  $x \uparrow I$ , and

$$r(E_{V_I(x \uparrow I)} X) \geq r(E_{C(x \uparrow I)} X) > R^*$$

at every point in this integral, since  $V_I(x \uparrow I)$  is the upper part of  $C(x \uparrow I)$ , so  $E_{V_I} r(X) > R^*$ , namely no such  $V_I$  can exist, i.e.

$$\forall V_I \subseteq M_I \ni E_{V_I} r(X) < R^* \& ss(M_I - V_I) \cap M_I = \emptyset$$

It only remains to show that  $T \subseteq \cup_{I \subset N} M_I$  to establish that there exists no set  $V$  removable from  $Z^*$  with  $E_{V_I} r(X) < R^*$  and prove the theorem.

To help visualize these regions, consider Fig. 8, which depicts a two service, one resource system. In this picture, the  $M_I$  do not cover  $T$ .

We will show that  $T$  is contained in  $\cup M_I$  by partitioning  $T$  into  $T_I$ , each of which is contained in the corresponding  $M_I$ . Consider the boundary of  $T$ . All points  $x$  on the boundary either belong to  $E_i$  for some  $i$ , corresponding to points that are on the boundary of  $Z$ , or  $x \uparrow I \in \text{boundary of } G_I^*$  for some  $I$ , corresponding to points that are on the boundary of  $Z^*$  but not on the boundary of  $Z$ . Therefore we can partition the boundary of  $T$  into

$$\{\bar{T}_I\} \ni x \in \bar{T}_I \Rightarrow x \text{ is on the boundary of } T \&$$

$$\left[ \begin{array}{l} x \uparrow I \in G_I^*, |I| > 1 \\ x \uparrow I \in G_{\{i\}}^* \text{ or } x \in E_i, I = \{i\} \end{array} \right]$$

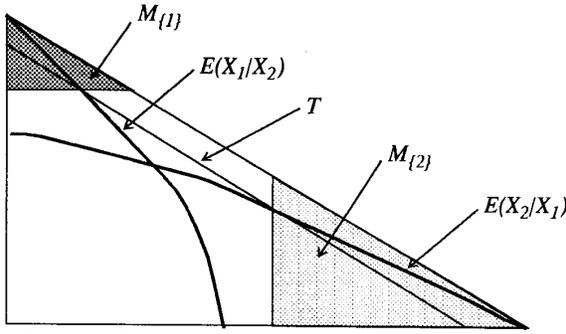


Fig. 8. A proposed system contradicting convexity

Now note that  $EX \in T$  since  $r(EX) = Er(X) = R$ . From  $EX$  to the boundary of  $T$ , we can partition  $T$  in a linear fashion into

$$\{T_I\} \ni (x \in T) \& x = \alpha EX + (1 - \alpha) y$$

for some  $0 \leq \alpha \leq 1$  &  $y \in \bar{T}_I$

where the multiplication is taken component-wise.

In Fig. 9, for our example, with a different  $T$  shown from Fig. 8 for purposes of illustration

$$T_{\{1\}} = \text{line segment from } \bar{T}_{\{1\}} \text{ to } EX$$

$$\& T_{\{2\}} = \text{line segment from } \bar{T}_{\{2\}} \text{ to } EX$$

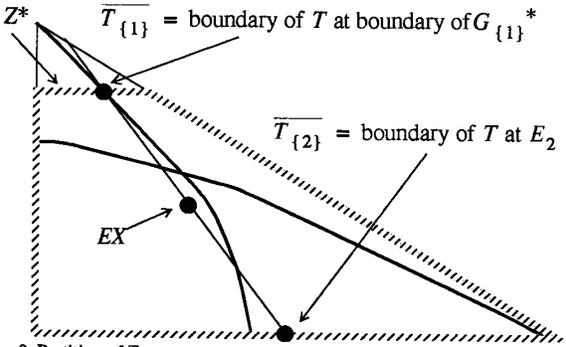


Fig. 9. Partition of  $T$

We will show that  $T \subseteq \bigcup_{I \in N} M_I$  by showing that  $T_I \subset M_I \forall I$ .

First note that  $EX \in \text{chull}(E_{C_Z^*}(x \uparrow I) X) \forall I$ , since

$$EX = \int E_{C_Z^*}(x \uparrow I) X \pi(C_Z^*(x \uparrow I)) d(x \uparrow I)$$

and  $E_{C_Z^*}(x \uparrow I) X$  is concave with respect to  $x \uparrow I$  from (DR).

Thus  $EX \in M_I \forall I$ , since  $EX$  must be both in  $T$  and in the convex hull of  $E_{C_Z^*}(x \uparrow I) X$ .

For our example,  $EX \in M_{\{2\}}$  because  $EX$  must be both on  $T$  and in the convex hull of  $E(X_2/X_1)$  and hence on the line segment  $AB$  in Fig. 10.

Furthermore the "I" boundary of  $T, \bar{T}_I \in M_I$ , from its definition.

We have shown that  $EX \in M_I$  and that  $\bar{T}_I \in M_I$ . From construction of  $T_I$  and from concavity (DR), it follows that all of  $T$  from  $EX$  to  $\bar{T}_I$  must be in  $M_I$ , namely,  $T_I \subset M_I \forall I$ .

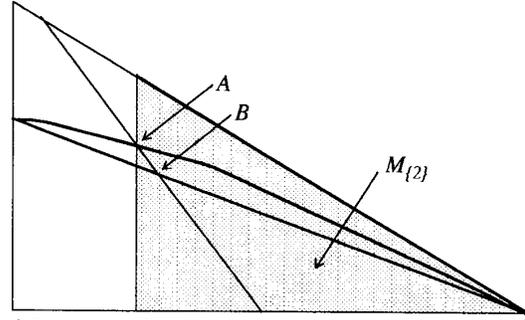


Fig. 10. Showing that  $EX \in M_{\{2\}}$

Thus  $T = (\bigcup T_I) \subset (\bigcup M_I)$ , and the proof is complete.  $\square$

REFERENCES

- [1] S. Jordan and P.P. Varaiya. "Throughput in multiple service, multiple resource communication networks." *IEEE Trans. Commun.*, vol. 39, pp. 1216-1222, Aug. 1991.
- [2] S. Jordan and P.P. Varaiya. "Control of multiple service, multiple resource communication networks." *IEEE Trans. Commun.*, vol. 42, Nov. 1994.
- [3] J. Aein. "A multi-user-class, blocked-calls-cleared, demand access model." *IEEE Trans. Commun.*, vol. 26, pp. 378-385, Mar. 1978.
- [4] J. Kaufman. "Blocking in a shared resource environment." *IEEE Trans. Commun.*, vol. 29, pp. 1474-1481, Oct. 1981.
- [5] B. Kraimeche and M. Schwartz. "Circuit access control strategies in integrated digital networks." in *Proc. of the Conf. on Comp. Commun. (IEEE Infocom)*, pp. 230-235, San Francisco, CA, Apr. 1984. IEEE.
- [6] K. Ross and D. Tsang. "Teletraffic engineering for product-form circuit-switched networks." *Adv. in App. Prob.*, vol. 22, pp. 657-675, 1990.
- [7] D. Tsang and K. Ross. "Algorithms to determine exact blocking probabilities for multirate tree networks." *IEEE Trans. Commun.*, vol. 38, pp. 1266-1271, Aug. 1990.
- [8] Z. Dziong and J. Roberts. "Congestion probabilities in a circuit-switched integrated services network." *Perf. Eval.*, vol. 7, pp. 267-284, 1987.
- [9] D. Mitra. "Asymptotic analysis and computational methods for a class of simple, circuit-switched networks with blocking." *Adv. in Appl. Prob.*, vol. 19, pp. 219-239, 1987.
- [10] W. Whitt. "Blocking when service is required from several facilities simultaneously." *AT&T Tech. Journal*, vol. 64, pp. 1807-1856, Oct. 1985.
- [11] F. Kelly. "Blocking probabilities in large circuit-switched networks." *Advances in Applied Probability*, vol. 18, pp. 473-505, 1986.
- [12] K. Ross and D. Tsang. "The stochastic knapsack problem." *IEEE Trans. Commun.*, vol. 37, pp. 740-747, July 1989.
- [13] S. Zachary. "Control of stochastic loss networks, with applications." *Journal of the Royal Statistical Society B*, vol. 50, no. 1, pp. 61-73, 1988.
- [14] K. Ross and D. Yao. "Optimal dynamic scheduling in jackson networks." *IEEE Trans. Automat. Contr.*, vol. 34, pp. 47-53, Jan. 1989.
- [15] P. Key. "Optimal control and trunk reservation in loss networks." *Prob. in the Eng. and Info. Sci.*, vol. 4, pp. 203-242, 1990.
- [16] G. Foschini and B. Gopinath. "Sharing memory optimally." *IEEE Trans. Commun.*, vol. 31, pp. 352-360, Mar. 1983.
- [17] S. Jordan. "Concavity properties in multi-server loss systems." in preparation.

Scott Jordan (S'86-M'90) received the B.S./A.B., M.S., and Ph.D. degrees from the University of California, Berkeley, in 1985, 1987, and 1990, respectively.

He is currently an Assistant Professor at Northwestern University. His teaching and research interests are the modeling and analysis of behavior, control, and pricing of computer/telecommunication networks.