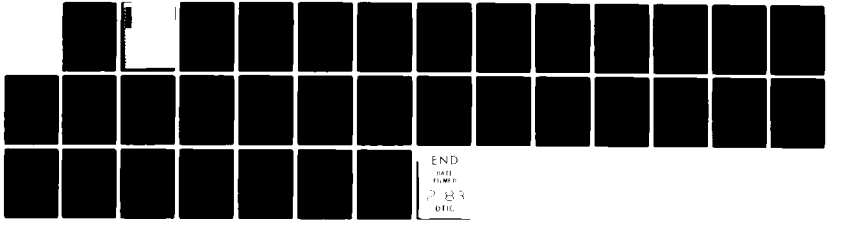
AD-A123 942 LINK CAPACITY CONTROL IN A COMPUTER COMMUNICATION NETWORK(U) ILLINOIS UNIV AT URBANA COORDINATED SCIENCE LAB M SCHLANSKER ET AL. JUL 80 R-889 N000014-79-C-0424 1/1

UNCLASSIFIED F/G 17/2 NL

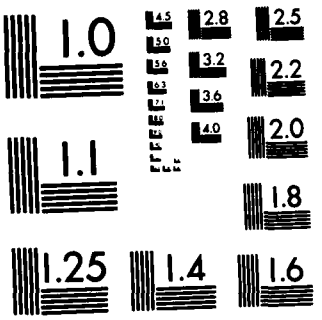


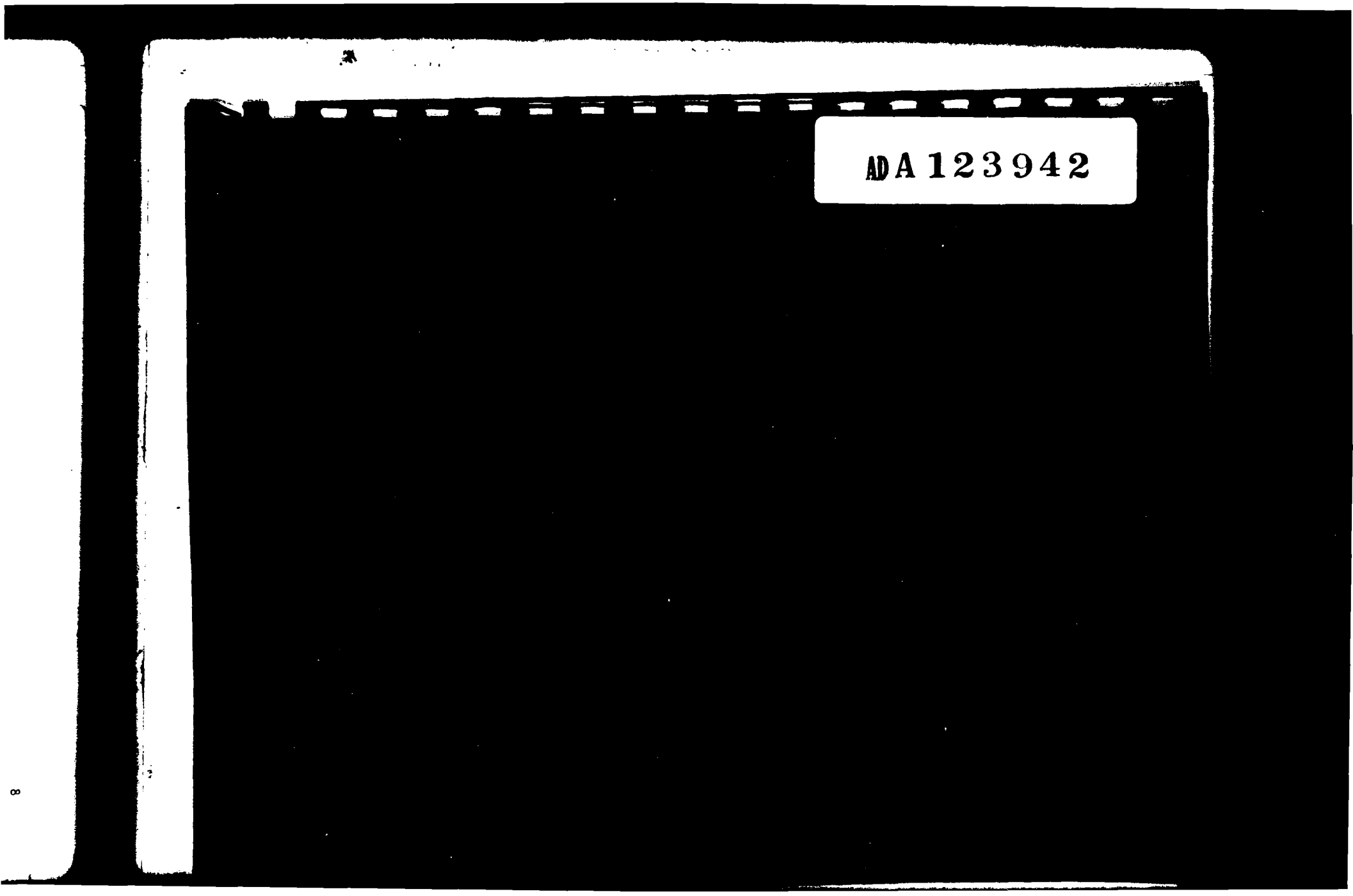
END

DATE FILMED

2 83

DTIC





UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A123942	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) LINK CAPACITY CONTROL IN A COMPUTER COMMUNICATION NETWORK		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER R-889; UILU-ENG 80-2221
7. AUTHOR(s) Michael Schlansker and Timothy Chou		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0424
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11 CONTROLLING OFFICE NAME AND ADDRESS Joint Services Electronics Program (U.S. Navy)		12. REPORT DATE July, 1980
		13. NUMBER OF PAGES 26
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

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

18. SUPPLEMENTARY NOTES

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

Circuit Switching, Computer Communication Network, Markov Decision Theory, Optimization, Packet Switching, Queuing Theory

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

This paper describes a line switching communication network where, as the network becomes congested, additional lines are opened to relieve congestion. We assume that a fixed charge is paid for each line opening or closing operation; a cost is paid per unit time for each line in use; a cost is paid per unit time for packet storage in a packet queue; and reward is earned for each packet transmission completed.

DD FORM 1 JAN 73 1473

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (Continued)

Under assumptions concerning Poisson arrival and service, a Markov model can be formulated which determines the average earning rate for the network under a specific policy. The policy is the control algorithm which specifies when lines should be opened or closed. The solution technique involves replacing an infinite subset of the state space by a finite set of states with equivalent behavior. A traditional technique called policy iteration can then be applied to the reduced finite model. The algorithm solves for the optimal policy, i.e. the policy yielding the highest earning rate.

The paper illustrates how optimal policies for the finite optimization problem are also exactly correct for the original infinite problem. While the work describes how a line switching network can be modelled and optimal control strategies determined; it also illustrates a modelling technique which can be applied to other optimization problems having an infinite state space.

UILU-ENG 80-2221

LINK CAPACITY CONTROL IN A COMPUTER COMMUNICATION NETWORK

by

Michael Schlansker and Timothy Chou

This work was supported in part by the Joint Services Electronics Program (U.S. Navy) under Contract N00014-79-C-0424.

Approved for public release. Distribution unlimited.

ABSTRACT

This paper describes a line switching communication network where, as the network becomes congested, additional lines are opened to relieve congestion. We assume that a fixed charge is paid for each line opening or closing operation; a cost is paid per unit time for each line in use; a cost is paid per unit time for packet storage in a packet queue; and reward is earned for each packet transmission completed.

Under assumptions concerning Poisson arrival and service, a Markov model can be formulated which determines the average earning rate for the network under a specific policy. The policy is the control algorithm which specifies when lines should be opened or closed. The solution technique involves replacing an infinite subset of the state space by a finite set of states with equivalent behavior. A traditional technique called policy iteration can then be applied to the reduced finite model. The algorithm solves for the optimal policy, i.e. the policy yielding the highest earning rate.

The paper illustrates how optimal policies for the finite optimization problem are also exactly correct for the original infinite problem. While the work describes how a line switching network can be modelled and optimal control strategies determined; it also illustrates a modelling technique which can be applied to other optimization problems having an infinite state space.

INDEX TERMS

Circuit Switching, Computer Communication Network, Markov Decision Theory, Optimization, Packet Switching, Queuing Theory

LINK CAPACITY CONTROL

IN A

COMPUTER COMMUNICATION NETWORK

by

Michael Schlansker

and

Timothy Chou

University of Illinois

Coordinated Science Laboratory

Urbana Illinois, 61801

LINK CAPACITY CONTROL

IN A

COMPUTER COMMUNICATION NETWORK

1. Introduction

Much of the work in the computer communication network area is concerned with networks operating with fixed capacity links between nodes. Typically the emphasis is on determining a fixed link capacity for each link within the network. Much of the results in the area are based on work by Kleinrock [1]. There are several capacity assignment algorithms and they are given in [2]. However, for networks with widely varying traffic loads, it may be beneficial to design a system where the link capacity is variable. This might be realistic for a network using multiple dial-up telephone lines where the link capacity can be increased by using another telephone line. In this paper, we will study the problem of dynamically controlling the link capacity between two nodes in a computer communication network. For a general network it is simple to apply the algorithm link by link throughout the network. The problem is treated using a Markov modeling technique where decision theory produces a optimal set of controls or policies yielding maximum average reward per unit time.

In section 2 we will define a Markov model describing the state of the link between the two nodes. This will be an infinite state space model which is difficult to solve using traditional analytic techniques. For that reason, in section 3 we will show a method of reducing the state space cardinality to that of a finite model. To do this we will note that at some point in time the packet queue is so heavily loaded

that any optimal policy will open all available lines. In section 4 we analyze this policy so that in section 5 we can show that this policy of opening all lines is indeed optimal when the packet queue is heavily loaded. Once the state space has been reduced, section 6 gives a brief description of the algorithm for determining the optimal policy for line management. Section 7 gives an example of the use of the algorithm and section 8 contains a summary and conclusions.

2. Model Description

Two nodes in a communication network communicate through an integral number of unidirectional lines each having fixed capacity, costing a fixed charge per unit time, and a fixed amount for each line opening or closing operation. In the two node network shown in figure 1, note that a queue of packets await transmission across the link from node N1 to node N2. The link between nodes represents an integral number of lines of equal bandwidth where, the number of lines is determined as a function of queue length at node N1. Thus information available at the source node is used to control the opening and closing of L unidirectional identical lines connecting node N1 to node N2.

The inter-departure times of packets served at each line are assumed exponentially distributed with parameter μ . The inter-arrival times of packets are also exponentially distributed but with parameter λ . Consider a collection of i open lines each serving packets at rate μ . The parallel collection of i open lines can be replaced by an equivalent single line with exponential inter-departure rate $i \cdot \mu$. This holds while there are at least as many packets as lines. When the

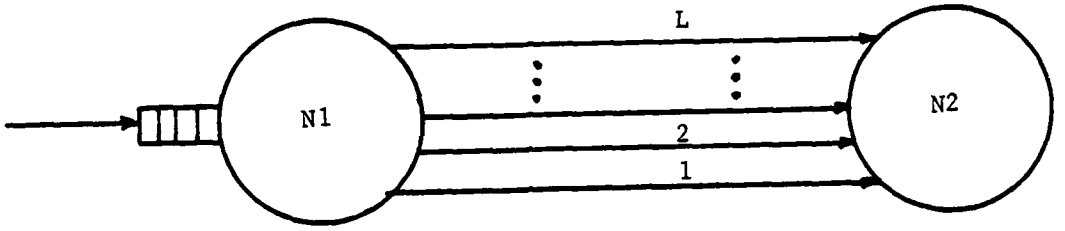


Figure 1

number of packets in the queue, j is less than the number of lines, then only j packets reside at a line server and the effective rate is $j \cdot \mu$. Thus, if i represents the number of open lines and j represents the packet queue length, the effective service rate $S_{i,j}$ is specified by: $S_{i,j}=\min\{i,j\} \cdot \mu$.

2.1 State and Policy

The line control problem described above can be modeled by a Markov process. Each state within the Markov process specifies the current number of lines in use, and the current queue length. Hold times and transition probabilities are determined; and a reward structure is defined; so that the Markov decision theory treated by Howard in [3] may be used to maximize the reward per unit time or gain of the process.

The <u>set of states</u>, $\{q_{i,j} | 1<=i<=L, 0<=j\}$ is defined so that at each state $q_{i,j}$ exactly i lines are open and j packets reside in the queue. Three possible policies exist at each state, a line may be opened, a line may be closed, or the line count may remain constant and the packet arrival/service process will operate. It is assumed that a line opening or closing operation will occur instantaneously. The policy function, $POL_{i,j} \subseteq \{\underline{o},\underline{c},\underline{r}\}$ represents the selected line control policy at $q_{i,j}$. A policy function will be chosen which will maximize the reward earned per unit time in steady state operation.

The choice of policy must be restricted at some states. We define the <u>range of policy</u> $RP_{i,j} \subseteq \{\underline{o},\underline{c},\underline{r}\}$ is the set of allowed policies at $q_{i,j}$:

TABLE 1 Range of Policy

state index	$RP_{i,j}$
i=1 , j>=0	{o,r}
1<i<L , j>=0	{o,c,r}
i=L , j>=0	{c,r}

Note that in line i=1, line closing is disallowed while in line i=L, line opening is disallowed. This limits both the maximum and the minimum number of lines which are active at any moment in time.

2.2 Holding Times

The average time spent within a state under a given policy is defined to be the hold time $HT_{i,j,\alpha}$. Under the assumptions of Poisson arrival and service, hold times can be computed from the process parameters as follows:

TABLE 2 Hold Time for $\alpha \in RP_{i,j}$

policy	$HT_{i,j,\alpha}$
α=r	$1/(S_{i,j}+\lambda)$
α=o or =c	0

for $1 \le i \le L$, $0 \le j$

2.3 Reward Structure

The reward structure describes the expected rewards minus the expected costs at each state within the system. The pertinent parameters are:

LTC	line cost per unit time
PSC	packet storage cost per unit time
REW	reward per packet for transmission
LOC	line opening cost
LCC	line closing cost

The reward $R_{i,j,\alpha}$ at state $q_{i,j}$ assuming the use of policy $\alpha \in RP_{i,j}$ is expressed in the following table:

TABLE 3 Reward Function for $\alpha \in RP_{i,j}$

policy	$R_{i,j,}$
$=\underline{r}$	$(REW \cdot S_{i,j} - i \cdot LTC - j \cdot PSC) \cdot HT_{i,j,\underline{r}}$
$=\underline{o}$	-LOC
$=\underline{c}$	-LCC

2.4 Transition Matrix

The transition probability from state $q_{i,j}$ to state $q_{m,n}$ under policy α is defined by the function $P_{i,j,m,n,\alpha}$:

TABLE 4 Transition Probability for $RP_{i,j}$

policy	state index	$P_{i,j;m,n}$
=r	m=i and n=j+1	$P=\lambda/(\lambda+S_{i,j})$
	m=i and n=min{0,j-1}	$P=S_{i,j}/(\lambda+S_{i,j})$
	otherwise	P=0
=o	m=i+1 and n=j	P=1
	otherwise	P=0
=c	m=i-1, and n=j	P=1
	otherwise	P=0

A state diagram for the line control Markov process is shown in figure 2.

3. Reducing State Space Cardinality

The transition probability matrix and reward structure describe the infinite Markov process for the line switching communication network. The model will be reduced to a finite Markov process by representing occupancy of an infinite sub-set of states within the process by entry into a single aggregate state with appropriate mean cost and hold time. Using this approach, we construct an imbedded finite Markov process which describes the interesting portion of the policy solution domain.

Let us partition the set of states into two regions called the inner region, $\{q_{i,j} | 1<=i<=L, j<K\}$ and the outer region, $\{q_{i,j} | 1<=i<=L, j>=K\}$. The parameter K is a positive integer constant chosen to designate the beginning of the regular portion of the optimal policy solution. In the outer region, we will show that the packet

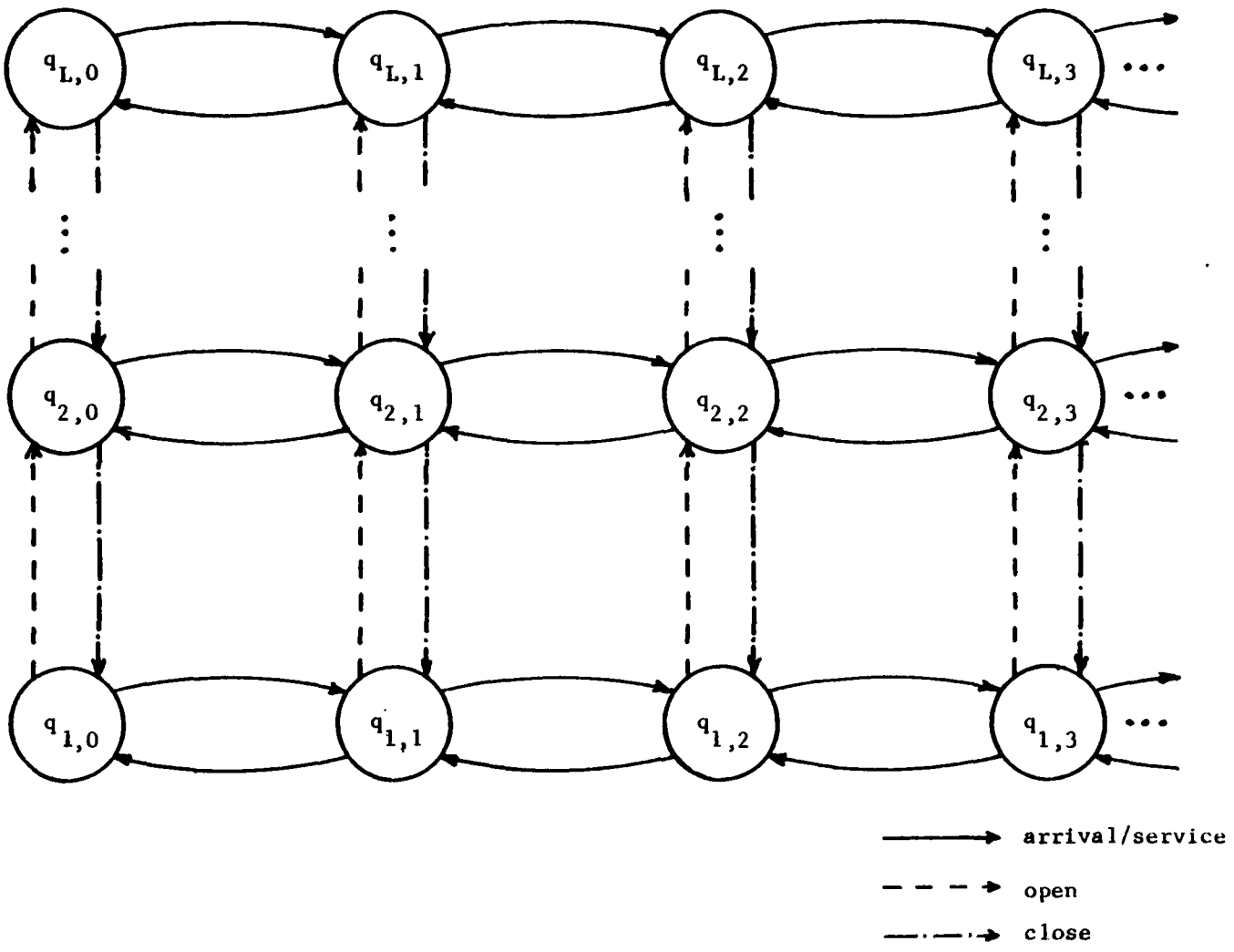

$q_{L,0}$
$q_{L,1}$
$q_{L,2}$
$q_{L,3}$
$q_{2,0}$
$q_{2,1}$
$q_{2,2}$
$q_{2,3}$
$q_{1,0}$
$q_{1,1}$
$q_{1,2}$
$q_{1,3}$
arrival/service
open
close

Figure 2

buffer is so heavily loaded that any optimal policy opens all lines to reduce future expected costs for maintaining packets in the queue. This is called the all lines open policy.

All Lines Open Policy

$POL_{L,j}=r$, $j>=K$.
$POL_{i,j}=o$, $1<=i<L$, $j>=K$

The outer region is shown in figure 3; note the use of the all lines open policy for all $j>=K$. In the outer region, where the queue is sufficiently deep ($j>=K$) the effective service rates, branch probabilities, and hold times are:

$$s_i=i\cdot\mu \quad p_i=\lambda/(\lambda+s_i) \quad ht_i=1/(\lambda+s_i) \quad \text{for} \quad 1<=i<=L$$

Lower case variables have been chosen here to indicate attributes of the outer region. The constant K must be chosen with $K>=L$ to insure that that there are enough packets in the queue to keep all L servers busy.

The uppermost row ($i=L$) of the outer region describes the basic queueing process which operates when the packet queue is heavily loaded and the all lines open policy is employed. The set of states $\{q_{L,j} \mid j>=K\}$ forms an infinite Markov chain where the only exit is the transition from $q_{L,K}$ to a member outside the set, $q_{L,K-1}$. Then, for all $j>=K$, p_L is the probability that a state $q_{L,j}$ makes the transition to $q_{L,j+1}$ while, $1-p_L$ is the probability of making the transition from $q_{L,j}$ to $q_{L,j-1}$. In all future discussions, it will be assumed that $L\cdot\lambda>u$ or equivalently, $p_L<1/2$. When $p_L>=1/2$, the Markov process is transient,

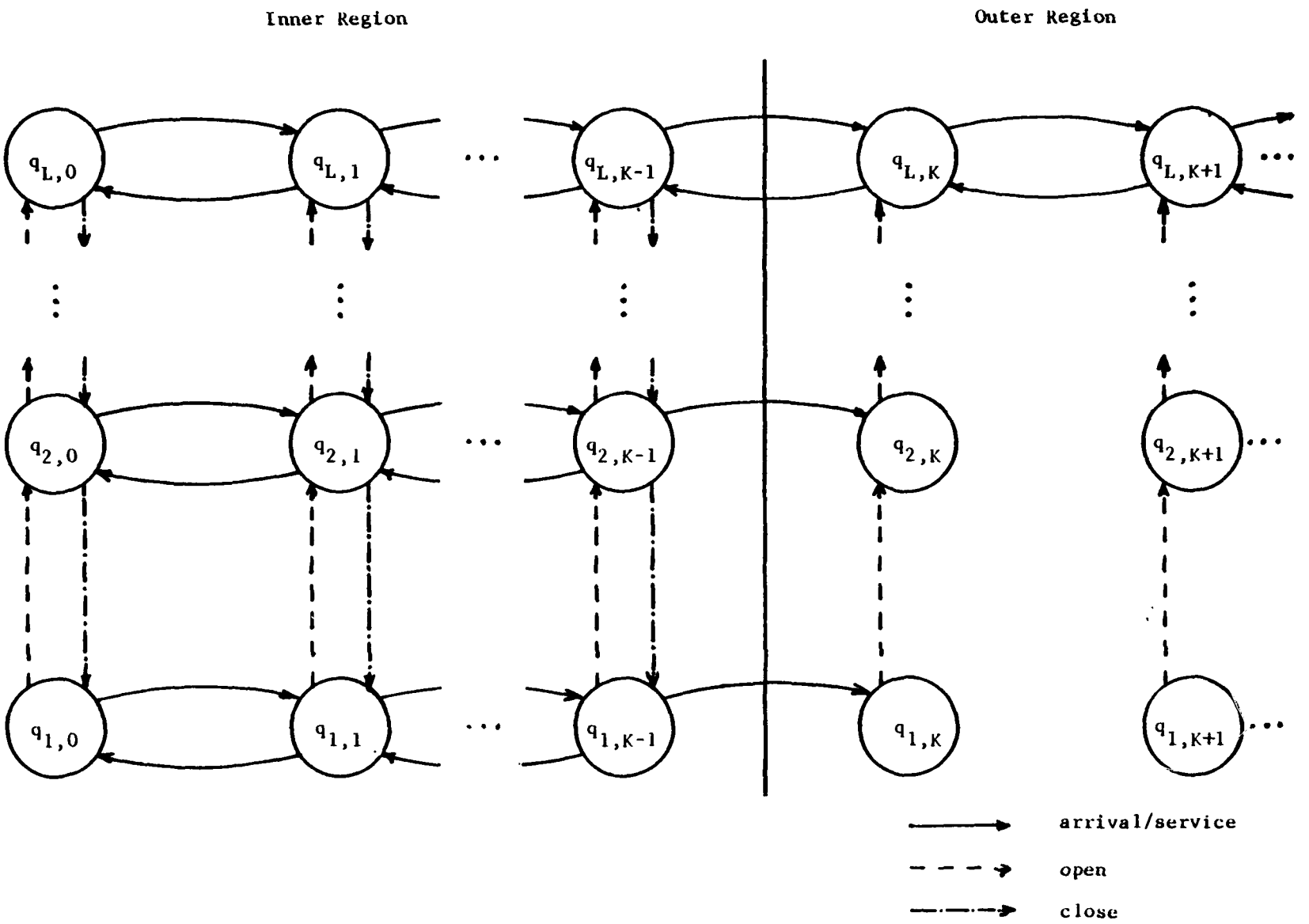

Inner Region
Outer Region
$q_{L,0}$
$q_{L,1}$
$q_{L,K-1}$
$q_{L,K}$
$q_{L,K+1}$
$q_{2,0}$
$q_{2,1}$
$q_{2,K-1}$
$q_{2,K}$
$q_{2,K+1}$
$q_{1,0}$
$q_{1,1}$
$q_{1,K-1}$
$q_{1,K}$
$q_{1,K+1}$
arrival/service
open
close

Figure 3

and the packet queue will grow without bound.

Define the earned reward function e(n), n>=0 to be the average cumulative reward earned after entry to state $q_{L,K+n}$, but before departure from the uppermost row of the outer region, $\{q_{L,j} | j>=K\}$ by branching to $q_{L,K-1}$. We can write the following balance equation by computing the earned reward e(n) at state $q_{L,K+n}$ in terms the earned reward from its neighbors:

$$e(n) = R_{L,K+n,\underline{r}} + (1-p_L)\cdot e_{n-1} + p_L \cdot e_{n+1} \quad \text{for } n>=0$$

$$e(-1) = 0 \qquad (1)$$

The initial condition for the balance equation, e(-1)=0 implies that no further reward is accumulated after the transition to $q_{L,K-1}$ which coincides with departure from the outer region.

The earned reward function can be readily evaluated by exploiting the symmetry of the uppermost row (i=L) of the outer region. Let H represent the expected time before the sub-chain, $Q_K=\{q_{L,j} | j>=K\}$ is exited assuming that the process starts in $q_{L,K}$. Let E represent the expected reward accumulated due to additional packet arrivals over the same interval. Then, the earned reward recurrence, e(n) expresses the fact that for any n>=0, the sub-chain $Q_{K+n}=\{q_{L,j+n} | j>=K\}$ has the same structure as the sub-chain Q_K except that an additional n entries must be maintained in the queue as long as the Markov process remains within Q_{K+n}. The hold time function, h(n) computes the time to exit Q_K from any state $q_{L,K+n}$, hence h(0)=H. These two recurrences are shown below:

$$e(n)=[a\cdot(K+n)+b]\cdot H + E + e_{n-1}$$

and (2)

$$h(n)=H + h_{n-1}\ ,\quad n>=0$$

Where, the initial conditions are: $e(-1)=0$ and $h(-1)=0$.

The parameters a and b are defined:

$a=-PSC$ and $b=(REW\cdot s_L - L\cdot LTC)$ so that,

$R_{L,j,r} = (a\cdot j + b)\cdot ht_L$ for $j>=K$.

Note that a and b above are the linear term in queue length j and constant term for the reward earning rate at states in the outer region. If the system is started in state $q_{L,K+n}$, the reward recurrence implies that K+n packets must be maintained for a time H equal to the time the system remains within the subset of states $\{q_{L,n+j} | j>=K\}$, E reflects additional costs from queue entries accumulated because of transitions to the right, and e_{n-1} represents costs accumulated after entering $q_{L,K+n-1}$. Symmetry and the Markov property dictate that H and E are well defined and independent of index n. The solution to these two recurrences is shown below:

$$e(n) = a\cdot[(n(n+1)/2)\cdot H] + (n+1)\cdot E + (a\cdot K+b)\cdot h(n)$$

and (3)

$$h(n) = (n+1)\cdot H \qquad \text{for all } n>=0$$

The constants H and E can be determined by substituting the solutions to the earned reward recurrence, (3) into the earned reward balance equation, (1) and equating coefficients of n:

$H = ht_L/(1-2\cdot p_L)$ and $E = ht_L\cdot p_L\cdot a/(1-2\cdot p_L)^2$.

Using these values for H and E, the solutions for e(n) and h(n) can be shown to satisfy both the balance equation of (1) and the recurrences of (2).

The solutions to the earned reward recurrence allow us to express the infinite line control Markov process by a finite one where in the uppermost row, an infinite set of states $\{q_{L,j} | j>=K\}$ is replaced by a single aggregate state $q'_{L,K}$ with appropriate mean cost $e(0)=E+(K\cdot a+b)\ H$ and mean hold time $h(0)=H$. The aggregate state $q'_{L,K}$ branches to $q_{L,K-1}$ with probability one thereby truncating the state space. The terminal states of lesser line index $q'_{i,K}$, $1<=i<L$ will be constrained to employ policy $\underline{o}$, a choice which will be shown to be optimal. This finite line control process is shown in figure 4. When K is chosen sufficiently large, the finite Markov process terminating at $q'_{i,K}$, $1<=i<=L$ describes an imbedded chain within the Markov chain formed by any optimal policy for the original infinite process under specific conditions which will be discussed later.

4. Relative Values under the All Lines Open Policy

In [3], Howard describes a procedure for the optimization of Markov decision processes. The policy optimization procedure involves an analysis step where, under a given policy, relative value equations are evaluated by solving a system of linear equations, and a policy improvement step where relative values from the previous policy are used to determine a new policy yielding higher average reward per unit time.

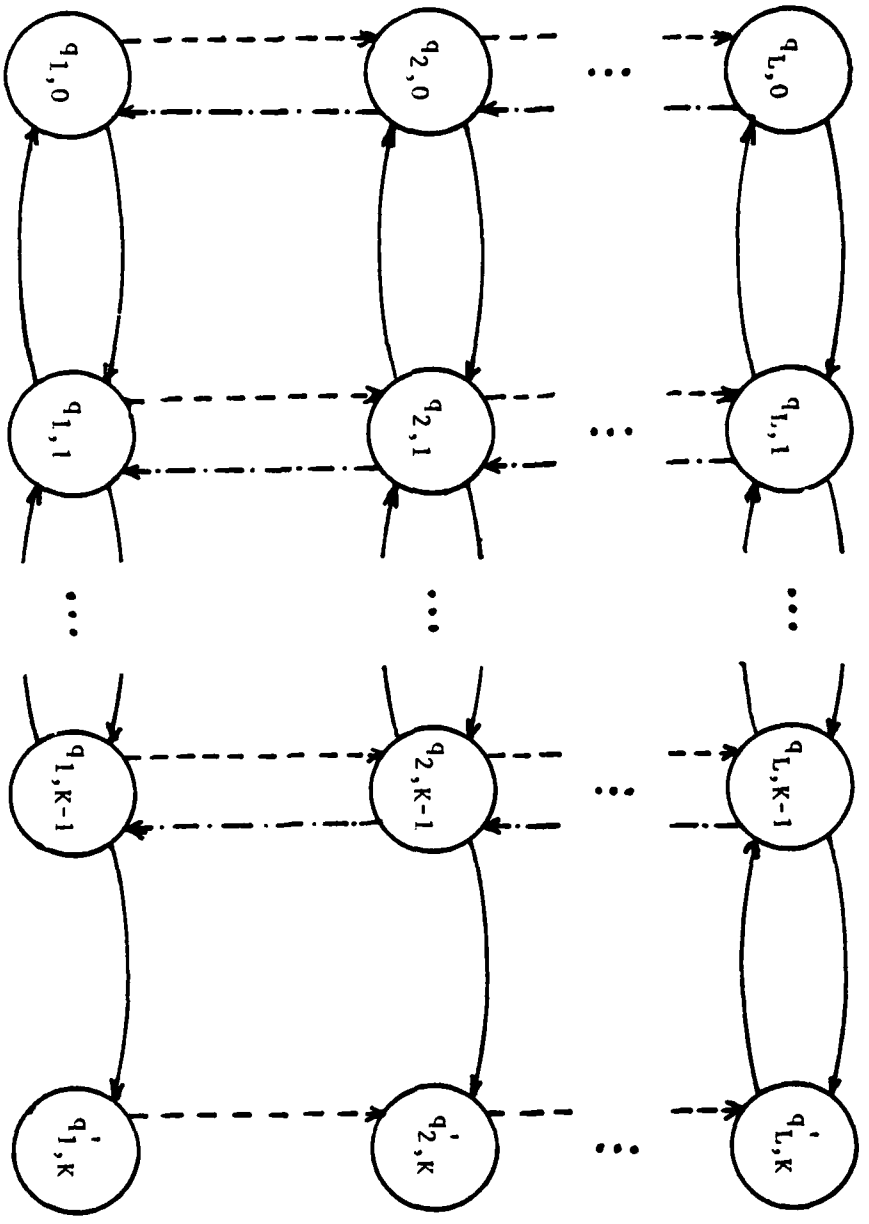

$q_{1,0}$
$q_{2,0}$
$q_{L,0}$
$q_{1,1}$
$q_{2,1}$
$q_{L,1}$
$q_{1,K-1}$
$q_{2,K-1}$
$q_{L,K-1}$
$q'_{1,K}$
$q'_{2,K}$
$q'_{L,K}$
arrival/service
open
close

Figure 4

When the average reward per unit time (the gain G) of the system is identical in two successive policy iteration steps, the iteration has converged and the associated policy is known to be optimal.

In this section, we will use the earned reward recurrence to construct the relative values in the outer region. The relative value equations for the line control Markov decision process appear as follows:

$$\text{define } \alpha \equiv POL_{i,j} \text{ then,}$$

$$V_{i,j} + G \cdot HT_{i,j,\alpha} = R_{i,j,\alpha} + \sum_{m,n} (P_{i,j,m,n,\alpha}) \cdot V_{m,n} \qquad (4)$$

$$\text{for} \quad 1 \leq i \leq L \;,\; 0 \leq j \;,\; 1 \leq m \leq L \;,\; 0 \leq n$$

In equation (4) above, $V_{i,j}$ is the relative value for state $q_{i,j}$, and G represents the process gain or average reward earned per unit time. Once the relative values have been determined under a given policy, a policy enhancement step is performed by maximizing the value oriented test quantity:

$$\underset{\alpha \in RP_{i,j}}{MAX} \left\{ R_{i,j,\alpha} - G \cdot HT_{i,j,\alpha} + \sum_{m,n} (P_{i,j,m,n,\alpha}) \cdot V_{m,n} \right\} \qquad (5)$$

$$\text{for} \quad 1 \leq i \leq L \;,\; 0 \leq j \;,\; 1 \leq m \leq L \;,\; 0 \leq n$$

This yields a policy with higher gain at each iteration until an optimal policy is reached.

The solutions of (3) for the earned reward and hold time recurrences may be used to compute relative values for all states within the uppermost row of the outer region, $\{q_{L,j} \mid j>=K\}$. A multistep relative value equation shown below will be used to compute the value of each of these states in terms of the value for state $q_{L,K-1}$ which will be given the arbitrary value V. This equation determines $q_{L,K+n}$'s value $V_{L,K+n}$ directly in terms of V and the average cost and time required to reach $q_{L,K-1}$ from any initial state $q_{L,K+n}$, $n>=0$.

$$V_{L,K+n} = -G \cdot h(n) + e(n) + V \qquad (6)$$

for $0<=n$, K fixed and sufficiently large

Here, h(n) represents the average duration of stay within the outer region assuming the system starts in $q_{L,K+n}$, e(n) accounts for the cost paid while within this region, and V represents the relative value of state $q_{L,K-1}$ which is reached after departing the outer region.

The relative values for states $q_{i,j}$ of lesser line index ($i<L$) under the assumption that $POL_{i,j}=\underline{o}$, $1<=i<L$, $j>=k$ can now be computed by repeatedly using equation (4) for lines L-1, L-2, ... ,1:

$$V_{i,K+n} = V_{L,K+n} - (L-i) \cdot LOC \qquad \text{for } 1<=i<=L \text{ , } 0<=n \qquad (7)$$
$$= -G \cdot h(n) + e(n) + V - (L-i) \cdot LOC$$

The relative values of (7) can be shown to satisfy the relative value equations of (4) where the all lines open policy is employed in the outer region. Equation (4) has been modified to employ the all lines open policy over the outer region yielding the equations below:

$$V_{L,j} = R_{L,j,\underline{r}} - G \cdot ht_L + (1-p_L) \cdot V_{L,j-1} + (p_L) \cdot V_{L,j+1}, \quad \text{for } j \geq K$$

and (8)

$$V_{i,j} = R_{i,j,\underline{o}} + V_{i+1,j} \quad \text{for } 1 \leq i < L,\ j \geq K.$$

Since the relative values satisfy (8), they must be the correct relative values of states within the outer region for the line control Markov process under the all lines open policy.

5. Optimal Policy for the Outer Region

The motivation for the work above is based on the assumption that the all lines open policy is optimal for states within the outer region. Now that solutions for the relative values have been established for states within the outer region under the all lines open policy, we shall determine whether these values indicate that the all lines open policy is indeed optimal. The conditions under which the optimality criterion is satisfied will now found by substituting the relative values from (7) into the test quantity of (5). The all lines open policy is optimal in the outer region exactly when (5) is maximized by selecting $POL_{L,j}=\underline{r}$ and $POL_{i,j}=\underline{o}$, $1 \leq i < L,\ j \geq K$.

Let us first consider the set of states within the uppermost row, $\{q_{L,j} | j \geq K\}$. It may be quickly shown that in this uppermost row of the outer region, the policy $POL_{L,j}=\underline{r}$ yields a test quantity which is at least as high as that of $POL_{L,j}=\underline{o}$ whenever LOC and LCC are non-negative. The policy $\underline{o}$ is not feasible and hence, not a candidate for maximizing (5). Thus, the policy $\underline{r}$ maximizes (5) in the uppermost row whenever line switching costs are non-negative.

For states of lesser line index, we will first show that the policy $POL_{i,j}=\underline{o}$ is strictly better than the policy $POL_{i,j}=\underline{r}$. Assume that the test value of (5) for $\alpha=\underline{o}$ is greater than that for $\alpha=\underline{r}$, $1<=i<L$, and the following inequalities result:

(9)

$$-R_{i,j,\underline{o}} - G\cdot HT_{i,j,\underline{o}} + V_{i+1,j} > R_{i,j,\underline{r}} - G\cdot HT_{i,j,\underline{r}} + (P_{i,j,i,j-1,\underline{r}})\cdot V_{i,j-1} + (P_{i,j,i,j+1,\underline{r}})\cdot V_{i,j+1}$$

$$\text{for} \quad 1<=i<L \quad , \quad j>=K$$

or

$$0 > -V_{i+1,j} + LOC + (a\cdot j+b)\cdot ht_i - G\cdot ht_i + (1-p_i)\cdot V_{i,j-1} + (p_i)\cdot V_{i,j+1}$$

After substituting the relative values for the outer region into (9), all second order terms in j cancel. If PSC is assumed to be positive and non-zero, the linear coefficient of j is negative and once the test inequality is satisfied, similar test quantities with higher index j must also satisfy this inequality. Selecting linear terms in j and simplifying, we have:

$$a\cdot p_i < a\cdot(H-h_i)/(2\cdot H) \quad , \quad 1<=i<L$$

$$\text{but, } h_L < h_i$$

$$\text{therefore, } \quad a\cdot p_i < a\cdot(H-h_L)/(2\cdot H) = a\cdot p_L$$

$$\text{or,}$$

$$\text{When } a=-PSC<0, \text{ we have:} \quad p_i > p_L \quad , \quad 1<=i<L$$

It can also be shown that $POL_{i,j}=\underline{o}$ yields higher test quantities than $POL_{i,j}=\underline{o}, j>=K$, $1<i<L$. Thus, if $p_i > p_L$, $1<=i<L$, then there exists an index K such that the optimal policy for all $q_{i,K+n}$, $1<=i<L$, $n>=0$ is

$=\underline{0}$. This follows because linear terms in j within equation (9) must eventually dominate constant terms for j sufficiently large. We conclude that whenever the effective service rate increases with number of lines, the all lines open policy will be optimal for some sufficiently large index K denoting the beginning of the outer region.

6. Policy Solution over the Inner Region

Once the infinite outer region of the state space has been replaced by the column of states, $q'_{i,K}$ the solution of the line control optimization problem becomes straightforward. Policy iteration techniques may be used to determine optimal policies for line management. The policy iteration cycle as described by Howard consists of first solving the system of equations of (4) along with a ground state definition equation (we chose $V_{1,0}=0$) for the relative values and the gain G, then a higher gain policy is determined using the previous relative values and the test quantity of (5). The iteration halts when the same policy is selected twice.

Some difficulties may be encountered because certain policies will yield a Markov process with multiple recurrent chains complicating the algebraic solution of the relative value equations which are potentially singular. This difficulty may be avoided either by using the policy iteration technique modified to treat polydesmic processes [3], or by carefully choosing an initial policy (e.g.: all lines open over the inner region) which defines exactly one recurrent chain within the uppermost row. Since, any state may experience an arbitrary large

number of arrivals in a small interval of time, all states have a non-zero probability of reaching some state $q_{i,j}$ with arbitrarily high queue length index j. Thus, if the all lines open policy is employed for all j>=K, all states can reach a single recurrent chain containing $\{q_{L,j} | j>=K\}$. This is a sufficient condition to insure that a policy forms a monodesmic Markov process. Not only does the initial policy form a monodesmic process, so do all successive policies searched from this initial policy in the iteration cycle.

The choice of K is sufficiently large whenever the optimal policy chosen for $q_{i,K-1}$, $1<=i<=L$ is identical to that chosen under the all lines open policy. Thus, whenever the finite policy iteration halts with an optimal policy where, $POL_{L,K-1}=\underline{r}$, and $POL_{i,K-1}=\underline{o}$, $1<=i<L$, this policy is optimal for the infinite process where of course the all lines open policy is employed in the outer region. This must be true because if (9) is satisfied for j=K-1, it will also be satisfied for all j>=K. When larger than necessary values for K are chosen, states which were within the outer region for smaller choices of K now lie within the inner region; however, the resulting policy and relative values for states within the smaller inner region will match those of corresponding states within the larger inner region. Hence, any choice of K which is too large will result in a correct solution.

7. Example

Figure 5 illustrates an application of this algorithm. The process parameters used in this example are shown in the table below:

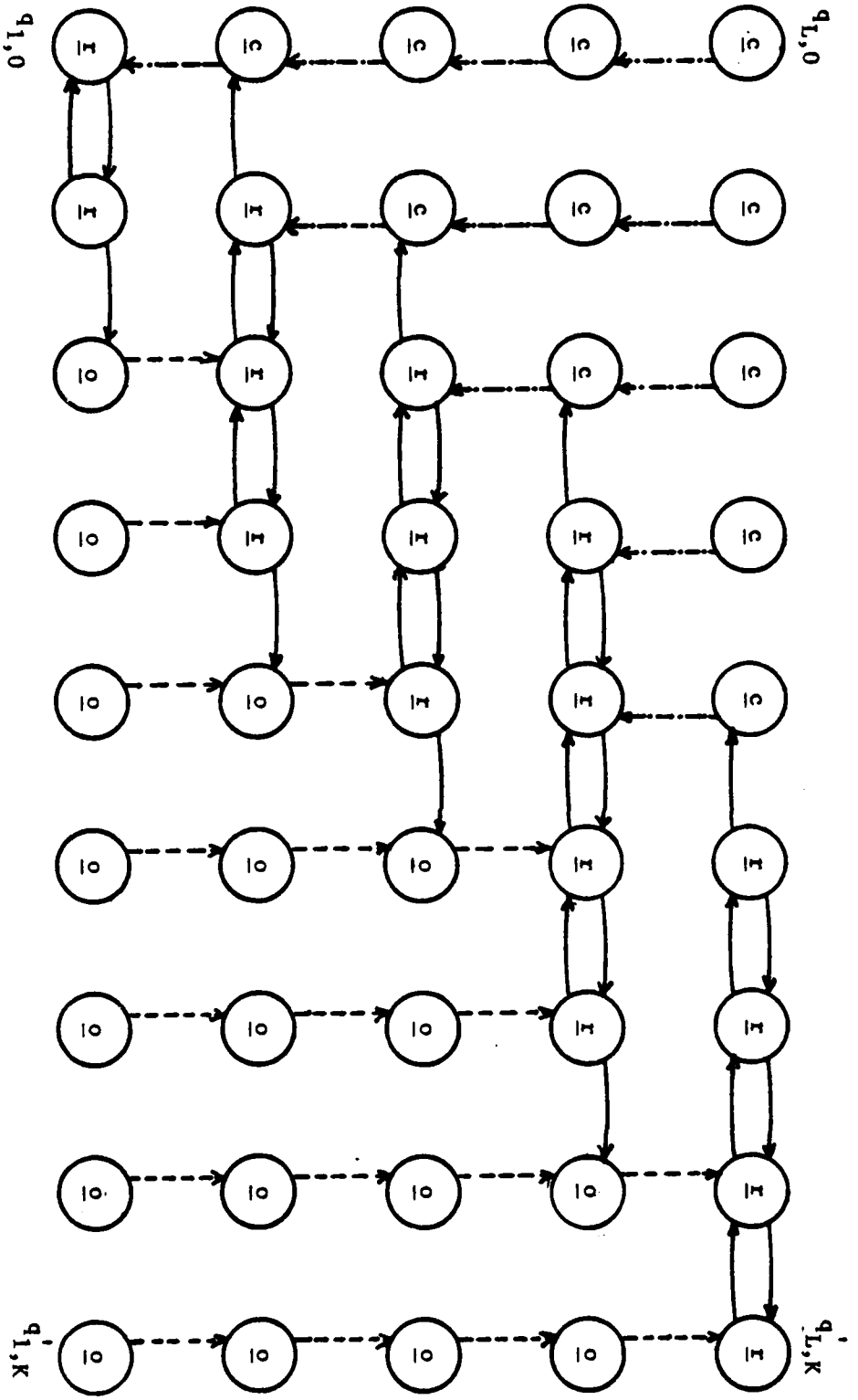


Figure 5

TABLE 5 Process parameters for example of figure 5

L	= 5	LTC	= 3.
K	= 9	PSC	= 3.
λ	= 2.	REW	= 1.
μ	= 1.	LOC	= 2.
		LCC	= 1.

The optimal policy resulting from use of the policy iteration algorithm shown within each state. Only transition arcs resulting from the optimal policy are shown. The parameter K=9 denotes the beginning of the outer region. Initially, K was chosen larger with identical results; however, K=9 was selected as the minimum value of K which still illustrates the optimality of the solution (all lines open for j=K-1). The policy iteration algorithm converged to the optimal policy from the initial all lines open policy in 7 iterations. Little is known of the rate of convergence of the policy iteration algorithm. While the number of possible policies is exponential in the number of states, all example problems converged very quickly to an optimal policy.

The relative values and gain resulting from the policy iteration algorithm are shown in figure 6. Note that for this example the gain is negative indicating that the network runs at a deficit. That is, rewards for packet transmission cannot pay for costs to mantain lines, switch lines and store packets; this results from a low choice of the value for REW. The relative values indicate the relative merit of residing within specific network states. The state $q_{1,0}$ was chosen as the ground state and therefore has value zero. All other relative values happen to be more negative than $q_{1,0}$ and are therefore more costly as initial states from which to resume message transmission.

	Queue Length								
Lines Open	0	1	2	3	4	5	6	7	8
5	-4.00	-8.22	-12.63	-17.14	-21.65	-26.17	-31.68	-38.20	-45.72
4	-3.00	-7.22	-11.63	-16.14	-20.65	-26.39	-33.08	-40.20	-47.72
3	-2.00	-6.22	-10.63	-15.25	-21.41	-28.39	-35.08	-42.20	-49.72
2	-1.00	-5.22	-10.06	-16.63	-23.41	-30.39	-37.08	-44.20	-51.72
1	0.00	-5.22	-12.06	-18.63	-25.41	-32.39	-39.08	-46.20	-53.72

Gain = -13.45

Figure 6

Note that the states relative value grows more negative with increasing queue depth as the costs for packet storage are certain to be higher.

8. Conclusion

The work above describes a model of line opening and closing operation within a computer communication network and illustrates the solution of optimal policies for line switching. The optimization is carried out under the assumption of exponential packet arrival and service. The assumption that packet service is exponential on each line is fairly realistic since, in a global sense it merely states that the packet service rate on each line is independent of the state (queue length and number of lines open). However, the assumption of an exponential arrival rate of messages is far more questionable since the intent of the control algorithm is to dynamically vary the number of lines in an environment of changing traffic load.

The algorithm which has been developed for an exponential arrival process could be directly applied to a non-exponential arrival process producing a reasonably good heuristic algorithm. However, there are a number of extensions to other more sophisticated heuristic approaches. If we assume that the the arrival process is approximately exponential, but the arrival rate is slowly varying in time, a superior approach would be to estimate the current arrival rate, and select a strategy compatible with the current arrival rate estimate.

The first part of this procedure is to determine the optimal policy over a wide range of arrival rates. This is done by statically resolving the problem for exponential arrivals over a wide range of

arrival rate parameter λ and determining specific ranges for λ where a particular policy is optimal or near optimal assuming an exponential process at the given rate. From this, the continum of the arrival rate parameter can be broken into ranges where a specific policy is preferred. This entire operation is performed statically at design time and is thus computationally feasible.

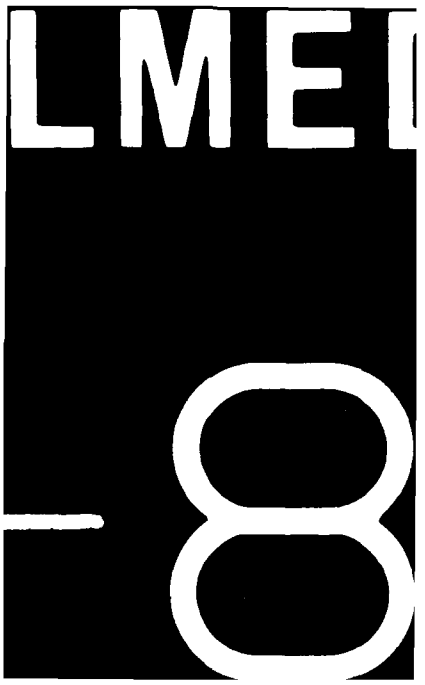
The second part of the procedure is to dynamically approximate the instantaneous arrival rate of the running network in order to select the most suitable policy from the tables produced above. The specific approach here depends strongly on the nature of the non-exponential arrival process, but a simple strategy will be described. If the source node breaks up time into fixed sized intervals and measures the number of arrivals within each interval, this sequence of interval measures can be used as a statistic from which one can derive an approximate arrival rate. For example, the rate estimator could, at each iteration, compute a current rate estimate as a weighted average of the old rate estimate and the current interval measure. This would geometrically decrease the significance of old interval measures and allow the construction of a simple rate estimator requiring very little computer time and memory space. The rate estimate could then be used to select the appropriate policy decision table over next time interval. More exotic schemes could be discussed but mean very little without a more careful characterization of the true nature of the arrival process.

REFERENCES

[1] L. Kleinrock, Communication Nets: Stochastic Message Flow and Delay, McGraw-Hill, New York, 1964. Reprinted, Dover publications, 1972.

[2] M. Schwartz, Computer-Communication Network Design and Analysis, Prentice-Hall, Inc., 1977.

[3] R. A. Howard, Dynamic Probabilistic Systems Vols.I &II Semi-Markov and Decision Processes, John Wiley and Sons, Inc.,.1971.


LME
8