SRC TR 88-95

TECHNICAL
RESEARCH
REPORT

Interchange Arguments In Stochastic Scheduling

by

P. Nain, P. Tsoucas, J. Walrand

SYSTEMS RESEARCH CENTER
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742

Interchange arguments in stochastic scheduling

Philippe Nain [1], *Pantelis Tsoucas* [2]
and
Jean Walrand

INRIA- Sophia Antipolis
Av. E. Hugues
06565 Valbonne Cedex
France

Systems Research Center
University of Maryland
College Park, MD 20742
USA

and

Department of Electrical Engineering and Computer Sciences
and Electronics Research Laboratory
University of California
Berkeley, CA 94720
USA

ABSTRACT

Interchange arguments are applied to establish the optimality of priority list policies in three problems. First, we prove that in a multi-class tandem of two $\cdot/M/1$ queues it is always optimal in the second node to serve according to the "$c\mu$" rule. The result holds more generally if the first node is replaced by a multi-class network consisting of $\cdot/M/1$ queues with Bernoulli routing. Next, for scheduling a single server in a multi-class node with feedback, a simplified proof of Klimov's result is given. From it follows the optimality of the index rule among idling policies for general service time distributions, and among pre-emptive policies when the service time distributions are exponential. Lastly, we consider the problem of minimizing the blocking in a communication link with lossy channels and exponential holding times.

November 28, 1988

1 Work performed while this author was visiting the Systems Research Center, University of Maryland, College Park, supported by the Office of Naval Research through grant number N00014-84-K-0614.

2 The work of this author was supported in part by the NSF Engineering Research Centers Program #NSFD-CDR-88-03012.

Interchange arguments in stochastic scheduling

Philippe Nain [1], *Pantelis Tsoucas* [2]
and
Jean Walrand

INRIA- Sophia Antipolis
Av. E. Hugues
06565 Valbonne Cedex
France

Systems Research Center
University of Maryland
College Park, MD 20742
USA

and

Department of Electrical Engineering and Computer Sciences
and Electronics Research Laboratory
University of California
Berkeley, CA 94720
USA

1. Introduction

This paper has two main aims. The first is to demonstrate the use of interchange arguments in proving optimality properties and the second is to obtain new results in stochastic scheduling. The main idea of our arguments appears in Varaiya et al [14] where it is used in order to obtain the optimality of index rules in multi-armed bandit problems. There, the objective is to maximize the expected total discounted reward. We use variations of this idea together with path-wise coupling techniques.

We first apply an interchange argument in Section 2 to partially characterize the optimal policy for scheduling two servers in a tandem of two nodes with M different classes of customers with exponential service times. The result, motivated by Ross and Yao [11], is that the optimal policy in the second node is a " c μ" rule. This is an easy extension of the results of Baras et al [2] and Buyukkoc et al [3]. The result can be extended to the case where the first node is a network consisting of $\cdot/M/1$ queues with Bernoulli routing.

Next, in Section 3 the problem of Klimov [7] is considered. A single server is to be scheduled in a network of $M/GI/1$ nodes. The objective is to minimize the expected long term average cost. It has been shown in [14] that this problem is equivalent to a multi-armed bandit problem. Our argument provides a simple proof of the result in [7], that the nonpre-emptive

---

1 Work performed while this author was visiting the Systems Research Center, University of Maryland, College Park, supported by the Office of Naval Research through grant number N00014-84-K-0614.

2 The work of this author was supported in part by the NSF Engineering Research Centers Program #NSFD-CDR-88-03012.

nonidling optimal policy is a priority rule. We also establish the optimality of that rule among idling policies. The priorities are determined and for the case where the service distributions are exponential we show that the same priority rule is optimal among pre-emptive policies. Remarkably, the optimal policy does not depend on the arrival rate. Our proof provides some insight into this fact.

Finally, in Section 4 we consider a problem of stochastic scheduling that does not fall in the framework of multi-armed bandit problems. Calls arrive at a communication link where N channels are available. There are probabilities of immediate loss associated with each channel and a successful call occupies the channel for an exponential amount of time. If the holding times are all independent and identically distributed, Anantharam et al [1] show that the time to reach the state where all channels are full is independent of the placement policy used. We provide a simple proof of this result and further prove that in the case where the holding times are not identically distributed, the time to reach the full state is stochastically maximized by assigning calls to the free channel with the shortest holding time.

2. Server scheduling in a multi-class network

Consider two $\cdot/M/1$ queues in tandem with M classes of jobs. Calls arrive at the first node at deterministic time instants $\{a_k\}_k$. The service rates at the first node are $\{\mu_i\}_{i=1}^M$ and there are associated holding costs denoted by $\{c_i\}_{i=1}^M$. In the second node the corresponding service rates are $\{\nu_i\}_{i=1}^M$ and the holding costs are $\{d_i\}_{i=1}^M$. Let $\mathbf{x}_t=(x_t^1,\ldots,x_t^M)$ (respectively $\mathbf{y}_t=(y_t^1,\ldots,y_t^M)$) be the vector of class populations in node 1 (respectively node 2). Assume that $d_1\nu_1 \geq d_2\nu_2 \geq \cdots \geq d_M\nu_M$. A pre-emptive server allocation policy is a function

$$\pi : (\mathbf{x}_t,\mathbf{y}_t) \rightarrow (\pi_t^1(\mathbf{x}_t,\mathbf{y}_t),\pi_t^2(\mathbf{x}_t,\mathbf{y}_t)) \in \{1,2,\ldots,M\}^2$$

The objective is to minimize over π the expected discounted cost incurred in the interval $[0,T]$ given by

$$J(\pi,T)=E\left[\int_0^T e^{-\beta t}\left(\sum_{i=1}^M c_i\, x_t^i+\sum_{i=1}^M d_i\, y_t^i\right)dt\right] \tag{2.1}$$

The following result shows that a "$d\nu$" policy is optimal for the second node. It is an extension of results in Baras et al [2] and Buyukkoc et al [3] who consider a single node. It also provides an extension to results in Ross and Yao [11] who consider multi-server scheduling in a network.

Theorem 2.1: In node 2 the optimal policy always serves job i, among the ones present in the queue, for which the quantity $d_i\,\nu_i$ is maximum.

Proof: The virtual service process of an exponential server with rate μ is a Poisson point process with parameter μ. A point of this process is a service completion if the queue is nonempty. Let $\{t_n^i\}$ (respectively $\{s_n^i\}$) be the points of the virtual service process for class $i\in\{1,\ldots,M\}$ in node 1 (respectively in node 2). We only need consider policies π switching at times

$$\{t_n\} = \{a_n\}\bigcup_{i=1}^M\{t_n^i\}\bigcup_{i=1}^M\{s_n^i\}.$$

For $T\geq 0$, condition on the number of points of the process $\{t_n\}$ in the interval $[0,T]$, say $0\leq t_1\leq t_2\leq \cdots \leq t_k\leq T\leq t_{k+1}\leq \cdots$. Optimality will be proved by induction on k. The result is trivially true for $k=0$ and assume that it holds for $k=1,\cdots,K$. We will prove that the result remains true for $k=K+1$.

By the optimality principle and the induction hypothesis $\pi_2(\cdot,\cdot)$ has to follow the "$d\nu$" rule at times $t_1,t_2,\ldots t_K$ and suppose that $\pi_0^2=i$ while $y_0^j>0$ with $i>j$. Then policy π cannot be optimal because it can be improved as follows. Denote by t_σ the first time when $\pi_{t_\sigma}^2=j$ and define policy $\bar{\pi}$ as follows.

$$\overline{\pi}_t^1 = \pi_t^1,\ t \geq 0,$$
$$\overline{\pi}_0^2 = j,$$
$$\overline{\pi}_t^2 = \pi_t^2,\ t = t_1, \ldots, t_{\sigma-1},$$
$$\overline{\pi}_{t_\sigma}^2 = i,$$
$$\overline{\pi}_t^2 = \pi_t^2,\ t = t_{\sigma+1}, t_{\sigma+2}, \cdots .$$

Then simple algebraic manipulation shows that

$$J(\pi, K+1) - J(\overline{\pi}, K+1) \geq 0,$$

if

$$p(j, t_k)d_j - p(i, t_k)d_i \geq 0,\ k = 1, \sigma,$$

where

$$p(l, t_k) = Pr\{t_k \in \{s_n^l\}\},\ l = i, j\ \ k = 1, \sigma.$$

It is easy to verify that because $\{a_n\}$ is a deterministic process and processes $\{t_n^i\}$, $\{s_n^i\}$ are Poisson,

$$\frac{p(j, t_k)}{p(i, t_k)} = \frac{\nu_j}{\nu_i}, k = 1, \sigma.$$

Therefore, $J(\pi, K+1) - J(\overline{\pi}, K+1) \geq 0$ since $d_j \nu_j \geq d_i \nu_i$ and π cannot be optimal. A similar argument shows that a policy that idles in node 2 at time 0 cannot be optimal. Note that policies π and $\overline{\pi}$ are not feasible because they are allowed to switch at all the points of process $\{t_n\}$, some of which are not observable. The above argument shows that the "$d\nu$" rule is optimal among all such policies. Yet, the "$d\nu$" rule is a feasible policy and is optimal among feasible policies as well.

□

Remarks 2.1

(a) In the above proof, policies π and $\overline{\pi}$ result in identical arrivals for node 2. The proof, except for the embedding, is identical to the one in Buyukkoc et al [7].

(b) The result stated in Theorem 2.1 remains true if node 1 is replaced by a network of $\cdot/M/1$ queues with Bernoulli routing (see Figure 2.1), and the cost function in (2.1) is modified in the obvious way.

3. Klimov's problem

3.1. The problem

The following situation was considered in Klimov [7]. There are N queues. The service times are independent and have the distribution function $G_i(t)$ in queue i $(1 \leq i \leq N)$. Customers arrive as an independent Poisson process with rate λ and are assigned to queue i with probability p_i. Write $\mathbf{p} = (p_1, \ldots, p_N)$. Upon service completion in queue i, a customer is sent to queue j with probability p_{ij}, and leaves the network with probability $p_{i0} = 1 - \sum_{j=1}^{N} p_{ij}$, independently of the state of the network. There is a single server that is allocated to one of the nodes at a time, in a nonpre-emptive way.

Assumptions

1. The matrix $P = [p_{ij}, 1 \leq i, j \leq N]$ is such that every customer eventually leaves, i.e., $P^n \to 0$ as $n \to \infty$. In particular, this implies that $(I - P)$ is invertible.

2. It is also assumed that $\int_0^\infty t \ dG_i(t) =: \beta_i < \infty$, for $1 \leq i \leq N$.

3. Finally, one assumes that $\lambda \mathbf{p}[I - R]^{-1}\beta < 1$, where $\beta = (\beta_1, \ldots, \beta_N)^T$ $((.)^T$ denotes transposition).

Denote by Z_t^i the number of customers at time $t \geq 0$ in queue $i \in \{1,...,N\}$ and let $\mathbf{Z}_t = (Z_t^1, \ldots, Z_t^N)$. Fix $c_i \geq 0$ for $1 \leq i \leq N$ and such that $\sum_{i=1}^N c_i = 1$. For a given server allocation policy π, one defines the average waiting cost per unit of time as

$$J(\pi) = \liminf_{T \to \infty} \frac{1}{T} E \int_0^T \sum_{i=1}^N c_i Z_t^i \ dt. \tag{3.1}$$

A policy is said to be *admissible* if it is non-idling, nonpre-emptive and nonanticipative. Non-idling means that the server is idle only when the system is empty. Nonanticipative means that the decision to allocate the server to queue i at time $t \geq 0$ is based on the evolution of the network up to time t. Under Assumptions 1,2 and 3 the system is ergodic under any non-idling policy (see Section 3.2).

A policy is optimal if it minimizes the cost (1.1) over all the admissible policies. The problem is to find an optimal policy.

Outline

We give a simple proof and provide two extensions of the result in [7]. Specifically, we show that a priority list policy that serves the non-empty node with the highest priority is optimal. Remarkably, the priorities do not depend on the parameters of the arrival process. In Section 3.2 we discuss the effects of idling in the simple case of two nodes with no feedback. Some auxiliary calculations are performed in Section 3.3 and are used subsequently in Section 3.4 to derive a priority index for each node. The optimality results extend to the case of pre-emptive policies for nodes where the service distributions are exponential.

3.2. The busy period

Decomposition

Convention: The set of nodes is partitioned into $\mathbf{n} = \{1,..,n\}$ and $\mathbf{n}^c = \{n+1,...,N\}$. Assume that while the nodes in $\mathbf{n}$ are not all empty the server serves according to the priority rule $1 > 2 > ... > n$.

Notation: For a matrix M and sets of natural numbers A and B, M_{AB} denotes the matrix $\{M_{ij}\}_{i \in A, j \in B}$. Similar notation for vectors has the obvious meaning.

Definition: Let $B^{(n)}$ be the time it takes to empty nodes $\mathbf{n}$, i.e.,

$$B^{(n)} = \inf\{t > 0 \mid \mathbf{Z}_{\mathbf{n}}(t) = 0\}$$

We represent the queueing process in the system during a busy period as a collection of trees (see Feller [4]). A job arriving in node i with service requirement S_j is represented as a node of type i with weight S_j. The children of each node are jobs arriving in the system while the customer is being served. Each job initially present in the system is the root of a tree and the length of a busy period is the sum of the weights of these trees. It is shown in [13] that under Assumptions 2 and 3 one has $E[B^{(n)}] < \infty$ for all n. It is easy to see by this construction that

Fact: The random variable $B^{(n)}$ does not depend on the order of service.

Furthermore, one obtains a decomposition for the mean of a busy period. Let $\mathbf{e}_i$ be the ith unit vector in $(-\infty,\infty)^N$ and for any $\mathbf{Z}(0)$ write $\mathbf{Z}(0) = m_1\mathbf{e}_1 + ... + m_N\mathbf{e}_N$. Then,

$$E[B^{(n)} \mid \mathbf{Z}(0)] = \sum_{i=1}^n m_i E[B^{(n)} \mid \mathbf{e}_i]. \tag{3.2}$$

3.3. Auxiliary calculations

Probabilities of transition

In this section we calculate the transition probabilities of a customer exiting the set of nodes $\mathbf{n}$, i.e., we are interested in the quantity

$r_{ij}^{(\mathbf{n})}$: Indicator function of the event that node $j \in \mathbf{n}^c$ is the first node not in $\mathbf{n}$ that a job visits starting in node i. The probabilities are given by

$$p_{ij}^{(\mathbf{n})} = E\,[r_{ij}^{(\mathbf{n})}]. \tag{3.3}$$

Lemma 3.1: The probabilities defined in (3.3) above are given, in matrix form by

$$P_{\mathbf{n}^c\mathbf{n}^c}^{(\mathbf{n})} = P_{\mathbf{n}^c\mathbf{n}^c} + P_{\mathbf{n}^c\mathbf{n}}(I - P_{\mathbf{nn}})^{-1} P_{\mathbf{n}^c\mathbf{n}^c} \tag{3.4}$$

Proof: For $i \in \mathbf{n}^c$ and $l \in \mathbf{n}$ we write the following first step equations.

$$p_{ij}^{(\mathbf{n})} = p_{ij} + \sum_{l \in \mathbf{n}} p_{il}\, p_{lj}^{(\mathbf{n})} \tag{3.5}$$

$$p_{lj}^{(\mathbf{n})} = \sum_{l' \in \mathbf{n}} p_{ll'}\, p_{l'j}^{(\mathbf{n})} + p_{lj} \tag{3.6}$$

The result is obtained by writing the above equations in matrix form and solving (3.5) for $P_{\mathbf{n}\mathbf{n}^c}^{(\mathbf{n})}$.

□

Expected sojourn times

We now calculate the expected sojourn time for each passage of a customer through the set of nodes $\mathbf{n}$. For this define

$S_j^{(\mathbf{n})}$: Total amount of service that a job receives until it exits $\mathbf{n}$ having started at node $j \in \mathbf{n}^c$.

We set

$$T_j^{(\mathbf{n})} = E\,[S_j^{(\mathbf{n})}] \tag{3.7}$$

Lemma 3.2: The expected sojourn times defined in (3.7) are given by

$$T_{\mathbf{n}^c}^{(\mathbf{n})} = P_{\mathbf{n}^c\mathbf{n}}(I - P_{\mathbf{nn}})^{-1} \boldsymbol{\beta}_{\mathbf{n}} \tag{3.8}$$

Proof: For $j \in \mathbf{n}^c$ and $l \in \mathbf{n}$ first step equation give

$$T_j^{(\mathbf{n})} = \sum_{l \in \mathbf{n}} p_{jl}\, T_l^{(\mathbf{n})}$$

$$T_l^{(\mathbf{n})} = \beta_l + \sum_{l' \in \mathbf{n}} p_{ll'}\, T_{l'}^{(\mathbf{n})}$$

The proof then proceeds as in Lemma 3.1 above.

□

We next turn our attention to a quantity that will be important in the computation of priority indices in Section 3.5. Define

$R_j^{(\mathbf{n})}$: Total amount of time needed to clear the set of nodes $\mathbf{n}$ of the arrivals resulting from serving a job in node $j \in \mathbf{n}^c$ and through its sojourn in $\mathbf{n}$.

Lemma 3.3: The expectation of $R_j^{(\mathbf{n})}$ is given by

$$E\,[R_j^{(\mathbf{n})}] = \lambda(\beta_j + T_j^{(\mathbf{n})}) \sum_{i=1}^{n} p_i\, E\,[B^{(\mathbf{n})} \mid \mathbf{e}_i\,] \tag{3.9}$$

Proof: By $A(t)$ denote the number of arrivals in the interval $[0,t]$. The result then follows easily from relationship (3.2) and by noting that

$$E[R_j^{(n)}] = E[E[B^{(n)} \mid \mathbf{A}(S_j + S_j^{(n)})]]$$
$$= \sum_{i=1}^{n} E[\mathbf{A}_i(S_j + S_j^{(n)})] E[B^{(n)} \mid \mathbf{e}_i]$$
$$= \lambda(\beta_j + T_j^{(n)}) \sum_{i=1}^{n} p_i E[B^{(n)} \mid \mathbf{e}_i]$$

The last step follows from the fact that the arrival process is Poisson.

□

3.4. Optimality

The nonpre-emptive case

In this section we prove that, as mentioned in the Section 3.1, the policy that minimizes the cost defined in (3.1) is a priority list which we determine. It is clear that an optimal policy also minimizes the expected cost in each busy period of the system given by

$$J(\pi,B) = E[\int_0^B \mathbf{c}\cdot\mathbf{Z}(t)dt]$$

We give expressions for the priority index of each node.

First, the nodes $\{1,2,...,N\}$ are renumbered as follows. Assign number 1 to the node that maximizes the quantity

$$\frac{c_i - \sum_k p_{ik} c_k}{\beta_i}, \quad i=1,...,N. \tag{3.10}$$

Recursively, for $1 \le n < N$, assign the number $n+1$ to the node $i \in \mathbf{n}^c$ that maximizes the quantity

$$\frac{c_i - \sum_{k \in \mathbf{n}^c} p_{ik}^{(n)} c_k}{\beta_i + T_i^{(n)}},$$

where $\mathbf{n}$ is the set of nodes $\{1,..,n\}$ in the new numbering. Denote by π the priority assignment list that corresponds to this ordering.

Theorem 3.1: Policy π is optimal among all nonpre-emptive, non-idling policies.

Proof: Let $J(\pi' \pi, B)(\mathbf{Z})$ be the cost incurred in a busy period starting from state $\mathbf{Z}$ and following policy π' in the first step and π thereafter. Then it suffices to prove that

$$J(\pi' \pi, B)(\mathbf{Z}) \ge J(\pi, B)(\mathbf{Z}), \text{ for all } \mathbf{Z} \in \{0,1,...\}^N. \tag{3.11}$$

It suffices to consider the case where $\pi'(\mathbf{Z}) = i \ne \pi(\mathbf{Z}) = j$ with $i > j$. This implies that $\mathbf{Z}=(0,...,0,Z^j, \ldots, Z^i, *,...,*)$ with $Z^i Z^j > 0$. For simplicity consider that $\mathbf{Z}_0 = \mathbf{Z}$. To establish (3.11) define ρ to be the first time that policy $\pi' \pi$ serves node j. By ϵ (respectively ς) denote the job that was served in node i at time 0 (respectively in node j at time ρ). In the context of Section 3.2, ρ is the time it takes to clear the system of the descendants of job ϵ that have priority higher than j. Let l be the node in $(\mathbf{j}-1)^c$ where job ϵ ends up after its sojourn in $\mathbf{j}-1$ and let $\mathbf{x}$ be the vector of the the rest of the descendants of ϵ after their sojourn in $\mathbf{j}-1$. Then define $\rho+\sigma$ to be the time it takes to serve job ς in node j and clear the system of the descendants of ς that have priority higher than j. Similarly, let ς end up in $k \in (\mathbf{j}-1)^c$ after its sojourn in $j-1$ and let **y** be the vector of the rest of the descendants of ς after their sojourn in **j**−1. For each sample path that is obtained by applying policy $\pi' \ \pi$ construct a sample path where π is followed until time σ, then any policy $\overline{\pi}$ such that $\overline{\pi}(\mathbf{Z}(\sigma)) = i$ is followed for one step, and π is resumed afterwards. Denote this policy by $\pi^{(\sigma)}\overline{\pi}\pi$. The arrival and service processes of jobs with priority higher than j are interchanged as in the construction of Section 3.2, i.e., the descendants of job ϵ with priority higher than j and the descendants of job ς with priority higher than j are the same in both realizations. The arrival and service processes of jobs with priority lower than j are the same in both realizations. One then obtains (see Figure 3.1),

$$J(\pi' \ \pi, B)(\mathbf{Z}) - J(\pi^{(\sigma)}\overline{\pi}\pi, B)(\mathbf{Z}) = \tag{3.12}$$

$$= c_j(\beta_i + T_i^{(j-1)} + E[R_i^{(j-1)}]) + (\sum_{l \in (\mathbf{j}-1)^c} p_{il}^{(j-1)} c_l + \mathbf{c}\cdot E[\mathbf{x}])(\beta_j + T_j^{(j-1)} + E[R_j^{(j-1)}])$$

$$-\{c_i(\beta_j + T_j^{(j-1)} + E[R_j^{(j-1)}]) + (\sum_{k \in (\mathbf{j}-1)^c} p_{jk}^{(j-1)} c_k + \mathbf{c}\cdot E[\mathbf{y}])(\beta_i + T_i^{(j-1)} + E[R_i^{(j-1)}])\}.$$

To simplify this expression we need to determine $E[\mathbf{x}]$ and $E[\mathbf{y}]$ as functions of the system parameters. For this, denote by $a_{ml}^{(j-1)}$ the expected number of jobs that enter node $m \in \mathbf{j}-1$ during a busy cycle that starts with a job in node $l \in \mathbf{j}-1$. From Section 3.2 one gets

$$\mathbf{c}\cdot E[\mathbf{x}] = \sum_{i' \in (\mathbf{j}-1)^c} c_{i'} \sum_{m \in \mathbf{j}-1} p_{mi'}^{(j,-1)} \sum_{l \in \mathbf{j}-1} a_{ml}^{(j-1)} \lambda p_l (\beta_i + T_i^{(j-1)})$$

From this, relation (3.9) and some rearrangement one gets that $J(\pi' \ \pi, B) - J(\pi^{(\sigma)}\overline{\pi}\pi, B) \geq 0$ if

$$\frac{c_j - \sum_k p_{jk}^{(j-1)} c_k}{\beta_j + T_j^{(j-1)}} \geq \frac{c_i - \sum_k p_{ik}^{(j-1)} c_k}{\beta_i + T_i^{(j-1)}}$$

To complete the proof one now argues that

$$J(\pi' \ \pi, B) \geq J(\pi^{(\sigma)m}\overline{\pi}\pi, B) \underset{m \to \infty}{\longrightarrow} J(\pi, B) \tag{3.13}$$

by bounded convergence. Policy $\pi^{(\sigma)m}\overline{\pi}\pi$ is defined recursively as $\pi^{(\sigma)}\pi^{(\sigma)(m-1)}\overline{\pi}\pi$ for $m = 1, 2, \ldots$.

3uppose now that policy $\overline{\pi}$ idles for some amount of time, at state **Z**. The above argument then shows that $\overline{\pi}$ can be improved and thus policy π is optimal among idling policies as well.

□

Remark 3.2: The argument used in the above proof is a variation of an argument in Varaiya et al [14]. We have followed closely the notation in Weiss [16], where a similar argument appears. Our argument also gives a simple proof of the results in Foss [5] who considers a generalized version of Klimov's problem and obtains a corresponding index rule.

The pre-emptive case

Assume now that the service time distributions are exponential at all the nodes and consider the coupling described above where π is now a pre-emptive policy following the same priority assignment list as in the nonpre-emptive case. One then sees that $J(\pi' \ \pi, B) - J(\pi^{(\sigma)}\overline{\pi}\pi, B)$ is the same as in (3.12). It follows that π is optimal among pre-emptive policies.

Remark 3.3: While this paper was under review, the paper of Lai and Ying [9] appeared. There, asymptotics of the "open bandit problem" are studied as the discount factor approaches 1. The above results are then derived. Our approach is simpler in that it does not rely on previous results on multi-armed bandits. At the same time, the results in [9] can be simply obtained by a direct argument similar to ours (see [8] and [16]).

4. A communication link model

In this section we demonstrate how interchange arguments can be employed in problems that do not fall in the multi-armed bandit framework.

The model

We consider the following model of a communication link. There are N channels to be used for the transmission of telephone calls. The calls arrive according to a deterministic sequence $\{a_k\}_k$. Each call is to be placed on one of the idle communication links, if one is available, and is lost otherwise. A call placed on link i is immediately lost with probability p_i and with probability $1-p_i$ it occupies the link for a period of time which is exponentially distributed with parameter μ_i. The system is described by the vector $\mathbf{Z}\in\{0,1\}^N$ where $Z^i=1$ if a call is present at link i and $Z^i=0$ otherwise. A placement policy is a function

$$\mathbf{u} : \mathbf{Z} \rightarrow \mathbf{u}(\mathbf{Z})\in\{1,...,N\}.$$

such that $\mathbf{Z}^{\mathbf{u}(\mathbf{Z})}=0$ if $\mathbf{Z}\neq(1, \ldots, 1)$. That is, in state $\mathbf{Z}$ policy $\mathbf{u}$ will place the next arrival on link $\mathbf{u}(\mathbf{Z})$. We have restricted ourselves to deterministic policies. Our arguments however, easily extend to randomized policies.

Outline

First, the case where $\mu_i=\mu$, $i=1,...,N$ is considered. We prove that from any state $\mathbf{Z}$ the time $T_{\mathbf{Z}}$ it takes to reach state $(1,...,1)\in\{0,1\}^N$ has a distribution that is independent of the policy $\mathbf{u}$. This result was obtained by Smith [12] and Anantharam et al [1] by explicit computation of the moment generating function of $T_{\mathbf{Z}}$.

Next, for unequal μ_i's, the problem of stochastically maximizing $T_{\mathbf{Z}}$ for any initial state $\mathbf{Z}$ is considered. We prove that the optimal policy always places calls on the free channel with the largest μ_i.

4.1. Invariance

In this subsection it will be assumed that $\mu_1=\mu_2=\cdots=\mu_N=\mu$.

Theorem 4.1: For any $\mathbf{Z}\in\{0,1\}^N$, the distribution of $T_{\mathbf{Z}}$ does not depend on the policy $\mathbf{u}$.

Proof: We will use a stochastic variation of the argument used in Section 3.5. Let $\mathbf{u}$ be a priority list assigning calls in the order $1,...,N$. As in the proof of Theorem 3.1 it will suffice to show that

$$T_{\mathbf{Z}}(\mathbf{u}'\,\mathbf{u}) \underset{st}{=} T_{\mathbf{Z}}(\mathbf{u}) \tag{4.1}$$

for any policy $\mathbf{u}'$. To establish (4.1) it suffices to consider the case where $\mathbf{Z}\neq(1, \ldots, 1)$ and $\mathbf{u}'(\mathbf{Z})=i\neq\mathbf{u}(\mathbf{Z})=j$ with $i>j$. This implies that $Z^l=1$, $l=1,...,j-1$, and $Z^i=Z^j=0$.

Arguing again as in the proof of Theorem 3.1, we will establish an analogue of relationship (3.13). To this end, denote by a_σ the first time that $\mathbf{u}'\,\mathbf{u}$ places a call on link j and assume for simplicity that $a_1=0$. Denote the virtual service processes of the links by $\{S_t^l\}_{l=1}^N$ and their points by $\{s_n^l\}_{l=1}^N$. Also, for $l=1,...,N$, set $r_n^l=1$ if the nth trial to engage link l is a success, and 0 otherwise.

For each sample path of $(\mathbf{Z}_t)$ resulting from $\mathbf{u}'\,\mathbf{u}$ we construct a sample path of a process $(\overline{\mathbf{Z}}_t)$ as follows. Consider a policy that places a call on link j at time a_1, follows policy u afterwards, and places a call on link i at time a_σ, if $a_\sigma<s_1^i$. Then, $\mathbf{Z}_{a_\sigma}=\overline{\mathbf{Z}}_{a_\sigma}$ and the paths of $(\mathbf{Z}_t)$ and $(\overline{\mathbf{Z}}_t)$ can be made to coincide from time a_σ onward. This would be the obvious argument in the case where $\mu=0$. It would suffice to let $\overline{r}_n^l = r_n^l$, $l=1,\cdots,N$, $n=1,2,...$.

On the other hand, if $s_1^i<a_\sigma$, the paths of $(\mathbf{Z}_t)$ and $(\overline{\mathbf{Z}}_t)$ can again be made to coincide from time s_1^i onward. This is achieved in the construction of $(\overline{\mathbf{Z}}_t)$ by letting $\{S_t^j\}$ be the virtual service process at link i, i.e., $\overline{s}_1^j = s_1^i$. Then, $\mathbf{Z}^i_{s_1^i} = \overline{\mathbf{Z}}^j_{s_1^i} = 0$. The two cases are illustrated in Figures 4.1(a) and 4.1(b) respectively.

Formally, in the construction of process $(\overline{\mathbf{Z}}_t)$ the arrival process remains $\{a_k\}$ and policy $\mathbf{u}$ is followed (recall $\mathbf{u}(\mathbf{Z}) = j$) until $\tau = a_\sigma \wedge s_1^i$ with

$$\overline{S}_t^l = S_t^l,\ l \neq i, j,\ t \geq 0,$$

$$\overline{r}_n^l = r_n^l,\ l = 1,\ldots,N,\ n = 1,2,\ldots$$

$$\overline{S}_t^j = S_t^i,\ t \leq \tau.$$

If $\tau = s_1^i = \overline{s}_1^j$ then continue with $\mathbf{u}$. Otherwise, i.e., if $\tau = a_\sigma$, follow $\overline{u}(\overline{\mathbf{Z}}_\tau) = i$ and then continue with $\mathbf{u}$. Denote this composite policy by $\tilde{\mathbf{u}}$. In either case set

$$\overline{S}_t^i = S_t^i,\ \overline{S}_t^j = S_t^j,\ t > \tau.$$

We have thus obtained that

$$T_{\mathbf{Z}}(\mathbf{u}'\ \mathbf{u}) \underset{st}{=} T_{\mathbf{Z}}(\mathbf{u}^{(\tau)}\tilde{\mathbf{u}}).$$

The proof now concludes as in Theorem 3.1.

□

Remark 4.1: The model considered here is similar to the well known repairman model (see e.g. Nash and Weber [10]). Our methods should apply to that model as well. In particular, one should be able to simply obtain the results in Hirayama [6] where a related optimization problem is studied.

4.2. Optimality

In this subsection we assume that $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_N$ and that policy $\mathbf{u}$ assigns calls according to the priorities $1,\ldots N$.

Theorem 4.2: For any $\mathbf{Z} \in \{0,1\}^N$, $T_{\mathbf{Z}}$ is stochastically maximized by policy $\mathbf{u}$.

Proof: Again, as in Theorem 3.1, we will show that

$$T_{\mathbf{Z}}(\mathbf{u}'\ \mathbf{u}) \underset{st}{\leq} T_{\mathbf{Z}}(\mathbf{u})$$

where $\mathbf{u}'(\mathbf{Z}) = i \neq \mathbf{u}(\mathbf{Z}) = j$ with $i > j$. With notation as in the proof of Theorem 4.1 one can check that using a similar construction for a process $(\overline{\mathbf{Z}}_t)$ we have

$$T_{\mathbf{Z}}(\mathbf{u}'\ \mathbf{u}) \leq T_{\mathbf{Z}}(\mathbf{u}^{(\tau)}\tilde{\mathbf{u}}),\text{a.s.}, \tag{4.2}$$

on some probability space, where the stopping time τ remains to be specified.

The construction of $(\overline{\mathbf{Z}}_t)$ only differs from the one in the proof of Theorem 4.1 in that $\overline{S}^j[0,\tau]$ is a superset of $S^i[0,\tau]$. This can be done since $\mu_j \geq \mu_i$. Also, note that in this case it suffices to define τ as

$$\tau = \inf\{t \mid \mathbf{Z}_t \geq \overline{\mathbf{Z}}_t\}$$

where the inequality is component-wise.

□

5. Acknowledgements

We would like to thank Professor A. Makowski, and Mr. I. Lambadaris for useful discussions.

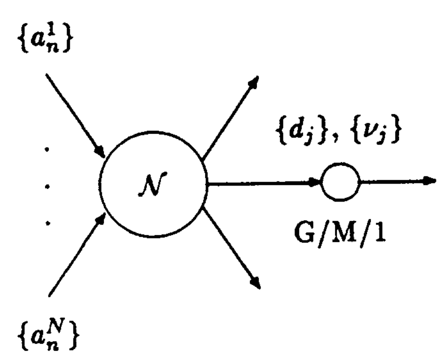
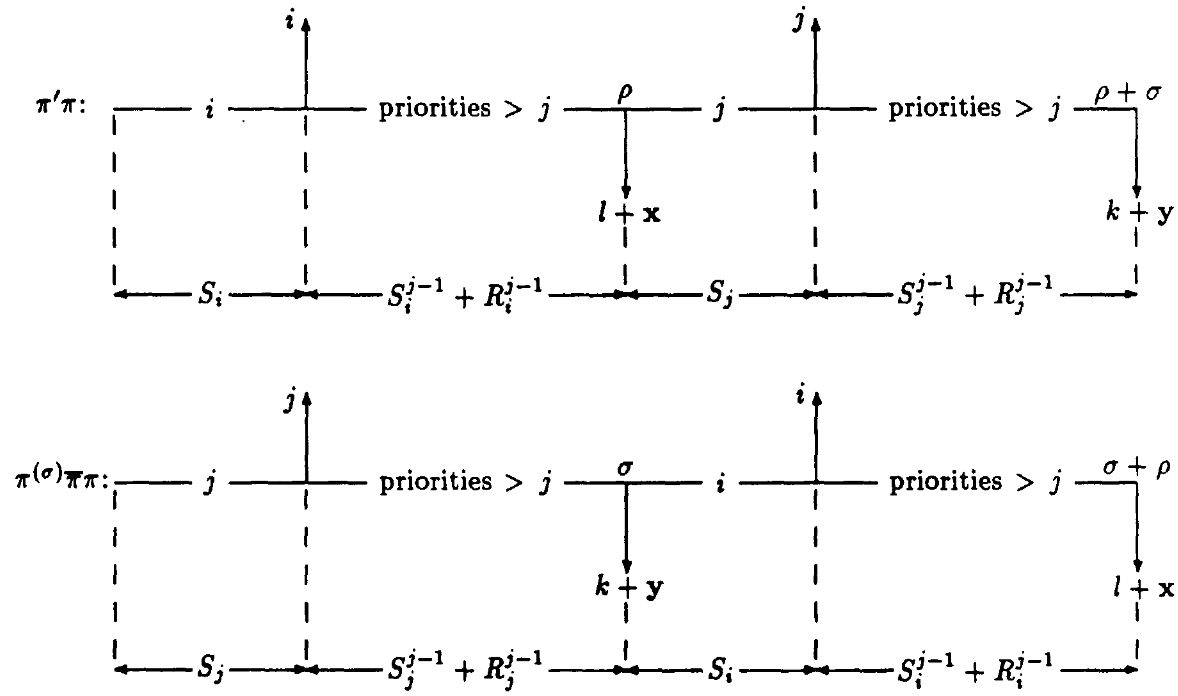


Figure 2.1



π′π:
i
priorities > j
ρ
j
priorities > j
ρ + σ
l + x
k + y
S_i
S_i^{j−1} + R_i^{j−1}
S_j
S_j^{j−1} + R_j^{j−1}
π^{(σ)}ππ:
j
priorities > j
σ
i
priorities > j
σ + ρ
k + y
l + x
S_j
S_j^{j−1} + R_j^{j−1}
S_i
S_i^{j−1} + R_i^{j−1}

Figure 3.1

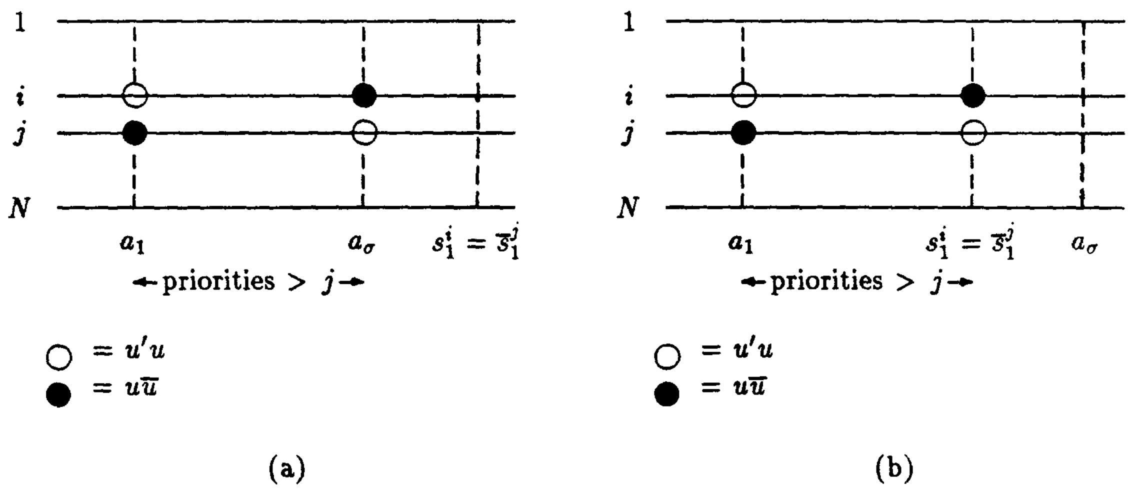

1
i
j
N
a_1
a_σ
$s_1^i = \overline{s}_1^j$
priorities > j
= u'u
= $u\overline{u}$
(a)
1
i
j
N
a_1
$s_1^i = \overline{s}_1^j$
a_σ
priorities > j
= u'u
= $u\overline{u}$
(b)

Figure 4.1