Optimal multiplexing of heterogeneous traffic with hard constraint

Philippe Nain, Keith W. Ross

OPTIMAL MULTIPLEXING OF HETEROGENEOUS TRAFFIC WITH HARD CONSTRAINT

by

Philippe NAIN[1], Keith W. ROSS[2]

[1]INRIA Centre de Sophia Antipolis, Route des Lucioles
06560 Valbonne, FRANCE

[2]Systems Engineering Department, University of Pennsylvania
Philadelphia, Pennsylvania 19104, U.S.A.

Abstract. Considered are optimal dynamic policies for multiplexing $K+1$ heterogeneous traffic types onto a single communication channel. The packet types arrive to the channel according to independent Poisson processes. The service requirements are exponential with type dependent means. The optimization criterion is to minimize a linear combination of the average delays for packet types 1 to K, while simultaneously subjecting the average delay of type-0 packets to a hard constraint. The optimal multiplexing policy is shown to be a randomized modification of the "μc rule". The optimization problem is thereby reduced to a problem of finding the optimal randomization factor; an algorithm, which can be implemented in real time, is given to do this for two particular cases.

I. INTRODUCTION

Consider the following multiplexing problem for an integrated communication network. Heterogeneous traffic types (e.g., interactive messages, file transfers, voice in packet form) arrive to a concentrator in the network, wait in a buffer and are given access to an outbound link according to a multiplexing policy. The designer is to choose a multiplexing policy that integrates the different data types equitably into the communication network. The data types may have different statistical properties and may place different demands on the system design : for interactive messages it may be desirable to have a constraint on the average time delay; for file transfers it is desirable to minimize the average time delay.

Motivated by the above application, we study in this paper optimal constrained multiplexing policies for a communication channel with packet types labeled $k = 0, 1, \ldots, K$. Specifically, the packet types arrive to an infinite capacity buffer according to independent Poisson processes with rate λ_k. Service times are assumed to be mutually independent and exponentially distributed with type dependent parameter μ_k, and independent from the arrival processes. The *decision points* are the *arrival* and *departure* epochs; therefore packets in service may be preempted at the arrival instants.

The multiplexing policy u specifies which packet type will be selected for service at each decision epoch; it is permitted to be *randomized* and to depend on the *past* and *current* line lengths [13]. For convenience, we will limit our attention to *work conserving* policies, i.e., to policies which always offer service to some packet when the buffer is nonempty.

Let c_k be a holding cost incurred for type-k packets, $k = 1, 2, \ldots, K$. Let $X_k(t)$ be the number of type-k packets in the system at time t and define $X(t) := (X_0(t), X_1(t), \ldots, X_K(t))$, for $t \geq 0$.

A policy u is said to be *feasible* if it meets the *hard constraint*, i.e., if u satisfies

$$V_n(u) := \limsup_t t^{-1} E_u[\int_0^t X_0(s)ds \,|\, X(0) = n] \leq \alpha, \tag{1.1}$$

for any initial state $n \in S := \mathbf{N}^{K+1}$, where $\mathbf{N}$ denotes the set of all nonnegative integers. A policy is said to be *optimal* if it minimizes, over the class of feasible policies, the long-run average cost

$$C_n(u) := \limsup_t t^{-1} E_u[\int_0^t \sum_{k=1}^K c_k X_k(s)ds \,|\, X(0) = n], \tag{1.2}$$

for any initial state $n \in S$. By Little's formula, the above optimization criterion is equivalent to minimizing a weighted sum of the average delays for packet types 1 through K, while simultaneously subjecting the average delay for type-0 packets to a hard constraint.

The optimization results of this paper can be extended to "nonexponential" queueing systems with or without preemption. Moreover, analogous results also hold for *discounted* and *finite horizon* criteria. However, these variations shall not be discussed since they only complicate the presentation without enhancing insight into the system behavior.

Klimov [8] studied a related problem in which packets, after having completed service, could change class and be routed back to the queue. The service requirements were general and only nonpreemptive policies were permitted. He considered the unconstrained optimization criterion of minimizing the average cost per unit of time and established the existence of an optimal nonrandomized static policy (see also [1], [2], [14]). However, for multiplexing integrated traffic onto a single channel (e.g., voice-data integration), it is natural to introduce hard constraints for certain traffic types (e.g., voice traffic, see [9]). The hard constraint leads to results and analyses which differ significantly from those of Klimov. The reader is also refered to [11] in which optimal constrained priority assignment was investigated for a *discrete time* queueing system.

In this paper we determine an optimal policy with a simple structure. To describe this policy, relabel the packet types 1 through K so that

$$\mu_1 c_1 \leq \mu_2 c_2 \leq \cdots \leq \mu_K c_K,$$

and let g_0 be the static policy that gives priority to packet types with the highest label, i.e., type-K packets receive the highest priority and type-0 packets receive the lowest priority. In particular, customers from classes 1 to K are served according to the so-called *μc rule* (e.g., see [1], [2], [7]). Consider also the policy g_k that is identical to g_0 except for giving priority to type-0 packets over types 1 through k for $k = 1, 2, \ldots, K$. In particular, g_K gives the highest priority to type-0 packets, and then gives priority to the other packet types according to the *μc rule*.

We show that if there is a policy that meets the constraint, then there is also an *optimal policy* that randomizes between two static policies g_{j-1} and g_j, for some fixed $j = 0, 1, \ldots, K$. The randomization mechanism involves repeated tosses of a coin with a bias factor that depends *neither* on the current state *nor* the past line lengths. Note that the special structure of the static policies g_{j-1} and g_j implies that the randomization is only necessary at the decision epochs for which there is contention between packet types 0 and j; otherwise, one simply applies the *μc rule*.

One feature of this optimization problem, which is not characteristic for many controlled queueing problems with multidimensional state space [15], is that we are able to completely specify the optimal policy for two particular cases. Indeed, once having established the aforementioned result, it only remains to determine the bias factor in order to completely specify the optimal policy. This is done for two cases : when the constraint value α is *tight* (to be made more precise); when $K = 1$. The optimal bias factor can then be calculated in real time for the case of time varying arrival and service rates.

In Section II, we collect some well known results from priority queues. The structure of the optimal policy is determined in Section III. In Section IV, the optimal bias factor is determined for the particular cases and numerical results are provided.

II. STATIONARY MULTIPLEXING POLICY

The line length process $\mathbf{X} := \{X(t),\ t \geq 0\}$ is a *Markov* process under any nonrandomized stationary policy. It is well known [5] that if

$$\rho := \sum_{k=0}^K \frac{\lambda_k}{\mu_k} < 1,$$

then X possesses a unique stationary distribution. However, under a randomized stationary policy, X is not a Markov process, but instead a *semi-Markov* process. This is because the random mix of two exponential sojourn times with different means is not exponential.

For a fixed initial state $n \in S$ and a given *work conserving* policy, consider the time it takes to empty the system of all packets. It is clear that the expectation of this quantity is independent of the particular work conserving policy employed. It therefore follows that if $\rho < 1$, then the semi-Markov process X has a unique stationary distribution for all work conserving stationary (possibly randomized) policies. We hereafter suppose that $\rho < 1$ and write $V(f)$ and $C(f)$ for $V_n(f)$ and $C_n(f)$ respectively, for any work conserving stationary policy f.

Recall the definition of the static policy g_k. With $\rho_k := \frac{\lambda_k}{\mu_k}$, for $k = 0, 1, \ldots, K$, we have the following well known result [7, p.39]

$$V(g_k) = \frac{\rho_0}{1 - \sum_{i=k+1}^{K} \rho_i} + \frac{\lambda_0(\frac{\rho_0}{\mu_0} + \sum_{i=k+1}^{K} \frac{\rho_i}{\mu_i})}{(1 - \sum_{i=k+1}^{K} \rho_i)(1 - \rho_0 - \sum_{i=k+1}^{K} \rho_i)}. \tag{2.1}$$

Define for $k = 1, 2, \ldots, K$, the *randomized μc rule*

$$f_{k,q} := [g_{k-1}, g_k, q]$$

which at each decision epoch, employs either the static policies g_{k-1} or g_k, with probability q and $1 - q$, respectively. Observe that

$$V(f_0) = V(g_k), \qquad V(f_1) = V(g_{k-1}). \tag{2.2}$$

We will need the following intuitively obvious result. A rigorous proof can be constructed following the proof of Theorem 5.4.3. of [12].

Lemma 2.1. *The average line length for type-0 packets $V(f_{k,q})$ is a continuous function of q over the interval [0,1], for each $k = 1, 2, \ldots, K$.*

These facts shall enable us to determine the structure of the optimal multiplexing policy in the following section.

III. THE STRUCTURE OF THE OPTIMAL POLICY

A constrained optimization problem can be reduced to one without constraints through the introduction of Lagrange multipliers. For each fixed multiplier $\omega \geq 0$, define the Lagrangian

$$J_n^{\omega}(u) := \limsup_t t^{-1} E_u \left[\int_0^t (\omega X_0(s) + \sum_{k=1}^{K} c_k X_k(s) ds) \,|\, X(0) = n \right],$$

and consider the *unconstrained problem* of minimizing $J_n^{\omega}(u)$ over the class of all policies.

It is well known (see e.g., [7, pp.199-200]) that, over the class of stationary policies, the *μc rule* is optimal. In other words if the multiplier ω satisfies

$$\frac{\mu_{k-1} c_{k-1}}{\mu_0} \leq \omega \leq \frac{\mu_k c_k}{\mu_0},$$

then the policy g_{k-1} is unconstrained optimal over the class of stationary policies, for $k = 1, 2, \ldots, K$.

In particular, if

$$\omega := \frac{\mu_k c_k}{\mu_0},$$

then the policies g_{k-1}, g_k and the randomized *μc rule* $[g_{k-1}, g_k, q]$ are all unconstrained optimal (over the class of stationary policies). Moreover, since the conditions of [3] are satisfied under the ergodicity condition $\rho < 1$, it follows that these policies are unconstrained optimal over the class of *all* policies.

We are now in position to determine the structure of the optimal constrained policy. Recall that the means $V(g_k)$ are explicitely given by (2.1)

Theorem 3.1. *If $V(g_K) > \alpha$, then there does not exist a feasible policy.*

If $V(g_K) \leq \alpha \leq V(g_0)$, then for some $q \in [0,1]$ the randomized μc rule $f_q := [g_{j-1}, g_j, q]$ is constrained optimal, where j is given by

$$j = \min\{i : V(g_i) \leq \alpha\}. \tag{3.1}$$

If $\alpha \geq V(g_0)$, then the static policy g_0 is constrained optimal.

Proof. The first and last assertions are obvious. Suppose $V(g_K) \leq \alpha \leq V(g_0)$ so that

$$V(g_j) \leq \alpha \leq V(g_{j-1}),$$

with j given by (3.1). From Lemma 2.1 and (2.2), it follows that there is a $q \in [0,1]$ such that

$$V(f_q) = \alpha. \tag{3.2}$$

Moreover, by the *μc rule*, the policy f_q minimizes $J_n^\gamma(u)$ for all $n \in S$, where

$$\gamma := \frac{\mu_j c_j}{\mu_0}.$$

Thus, for all initial states $n \in S$ and any feasible policy u, we have

$$\begin{aligned} C(f_q) + \gamma\alpha &= J^\gamma(f_q) \\ &\leq J_n^\gamma(u) \\ &\leq C_n(u) + \gamma V_n(u). \end{aligned} \tag{3.3}$$

Equation (3.3) in turn implies

$$C(f_q) - C_n(u) \leq \gamma(V_n(u) - \alpha),$$

$\forall n \in S$. Hence, $C(f_q) \leq C_n(u)$ for any feasible policy and any state $n \in S$, i.e., the policy f_q is *constrained optimal*. ∎

It is to be noted that the proof of the above theorem does not depend on the choice of the decision points. In particular, if the randomization is performed at the beginning of every busy period, then Theorem 3.1 holds, and in that case $V(f_q) = qV(g_{j-1}) + (1-q)V(g_j)$, where j is given by (3.1) (see also [7, pp. 186-187]). The optimal value of q, cf. (3.2), is then trivially determined. However, randomizing at the beginning of busy periods is not as natural as at the jump points, and moreover, it leads to a larger variance for the response times of type-0 packets.

In order to completely determine the optimal policy in the case where the decision points are the jump points, it remains to calculate the bias factor q that satisfied (3.2). This is the theme of Section IV.

IV. THE OPTIMAL BIAS FACTOR

For the remainder of this paper, we will assume that one of the two following conditions holds :

i) $\alpha \leq V(g_{K-1})$ (tight constraint value),

ii) $K = 1$.

In case i (and obviously in case ii) it follows that it is only necessary to study a system with packet types 0 and K. Because the function $q \longrightarrow V(f_q)$ is clearly *strictly increasing* and *continuous* in $[0,1]$ by Lemma 2.1 (here $f_q := [g_0, g_1, q]$ as in Theorem 3.1), the equation $V(f_q) = \alpha$ (cf. (3.2)) has a unique solution (the optimal bias factor) which can therefore be easily obtained numerically once $V(f_q)$ is known for all $q \in [0,1]$. For the remainder of this paper, we focus our attention on the determination of $V(f_q)$ for fixed $q \in (0,1)$. To this end we shall use the generating function method.

For $|x| \leq 1$, $|y| \leq 1$, we define the following two-dimensional generating function

$$F_q(x,y) := \lim_t E_{f_q}[\, x^{X_1(t)} y^{X_0(t)} \,|\, X(0) = n\,].$$

The above limit exists and is independent of the initial state $n \in S$ under the ergodicity condition $\rho < 1$ (see Section II). From now on it will be assumed that $\rho < 1$. Therefore,

$$V(f_q) = \frac{\partial}{\partial y} F_q(x,y)\Big|_{(x,y)=(1,1)}. \tag{4.1}$$

Consequently, it is seen from (4.1) that $F_q(x,y)$ must be determined in the vicinity of the point (1,1). To do this, we shall show that $F_q(x,y)$ satisfies a functional equation which can be solved analytically. To this end (see Remark 4.1), let $\mathbf{Q} := (q(m,n))_{(m,n)\in S^2}$ be the transition matrix for the Markov chain imbedded at the jump points of $\mathbf{X}$ when the policy f_q is applied for some fixed $q \in [0,1]$. Also let $\mathbf{v} := (v(n))_{n\in S}$ be the unique probability measure that satisfies

$$\mathbf{v} = \mathbf{v}\mathbf{Q}. \tag{4.2}$$

A classical result from the theory of semi-Markov processes [4, p.342] gives, for all $n \in S$,

$$\lim_t P_{f_q}(\, X(t) = n \,|\, X(0)\,) = \frac{v(n)\gamma(n)}{\mathbf{v}.\mathbf{\Gamma}}, \tag{4.3}$$

where $\mathbf{\Gamma} := (\gamma(n))_{n\in S}$ is the *conditional mean sojourn time* for the semi-Markov process $\mathbf{X}$, and $\mathbf{v}.\mathbf{\Gamma}$ is the inner product of $\mathbf{v}$ and $\mathbf{\Gamma}$.

Let us now introduce the transition matrix $\widehat{\mathbf{Q}} := (\widehat{q}(m,n))_{(m,n)\in S^2}$ for the *subordinated Markov chain* $(\widehat{X}_n)_{n\in S}$ (see [4, pp.259-260], [12, p.50]) defined by

$$\widehat{q}(m,n) := \begin{cases} \dfrac{q(m,n)}{\gamma(m)}, & \text{if } m \neq n; \\ 1 - \dfrac{1}{\gamma(m)}, & \text{otherwise.} \end{cases} \tag{4.4}$$

Let $\mathbf{\Pi} := (\pi(n))_{n\in S}$ be the unique invariant measure satisfying the equations

$$\begin{aligned} \mathbf{\Pi} &= \mathbf{\Pi}\widehat{\mathbf{Q}}, \\ \mathbf{\Pi}.\mathbf{1} &= 1. \end{aligned} \tag{4.5}$$

We then have the following :

Lemma 4.1.

$$\lim_t P_{f_q}(\, X(t) = n \,|\, X(0)\,) = \pi(n),\ \forall n \in S.$$

Proof. For $n \in S$

$$\begin{aligned} \sum_m v(m)\gamma(m)\widehat{q}(m,n) &= v(n)\gamma(n)(1 - \frac{1}{\gamma(n)}) + \sum_{m\neq n} v(m)q(m,n), \text{ from (4.4)}, \\ &= v(n)\gamma(n) - (v(n) - \sum_{m\neq n} v(m)q(m,n)), \\ &= v(n)\gamma(n), \text{ from (4.2). (Note that } q(n,n) = 0, \forall n \in S.) \end{aligned}$$

Consequently $\pi(m) = \dfrac{v(m)\gamma(m)}{\mathbf{v}.\mathbf{\Gamma}}$, $\forall m \in S$, from (4.5) and the uniqueness property. The proof is then concluded by considering (4.3). ∎

The above result shows that $F_q(x,y)$ is also the generating function of the *Markov chain* $(\widehat{X}_n)_{n\in S}$. Let us now consider the matrix $\widehat{\mathbf{Q}}$. With $e_0 := (1,0)$ and $e_1 := (0,1)$, a straightforward calculation gives for $m = (m_1, m_0) \in S^2$,cf.(4.4),

$$\widehat{q}(m, m+e_i) = \lambda_i;$$

$$\widehat{q}(m, m-e_i) = \begin{cases} 0, & m_i = 0; \\ \mu_i, & m_i > 0, m_{i-1} = 0; \\ \mu_i(i\xi + (1-i)(1-\xi)), & m_i > 0, m_{i-1} > 0; \end{cases} \qquad (i = 0,1);$$

$$\widehat{q}(m,m) = \begin{cases} 1-\lambda, & m_0 = m_1 = 0; \\ 1-\lambda-\mu_1, & m_0 > 0, m_1 = 0; \\ 1-\lambda-\mu_0, & m_0 = 0, m_1 > 0; \\ 1-\lambda-\mu_1\xi-\mu_0(1-\xi), & m_0 > 0, m_1 > 0, \end{cases}$$

where

$$\xi := q\frac{\lambda+\mu_0}{\lambda+q\mu_0+(1-q)\mu_1}; \qquad (q \in (0,1)); \tag{4.6}$$

$$\lambda := \lambda_1 + \lambda_0. \qquad \text{(By convention } m_{-1} = m_1.\text{)}$$

It then turns out that the transition matrix $\widehat{\mathbf{Q}}$ is exactly the same (up to an obvious change of notation) as the one of the process studied by Fayolle and Iasnogorodski [6, case "pq=$\mu_1\mu_2$"]. (In [6] the authors consider two parallel queues; for each queue the input process is Poisson and the service demands are exponentially distributed, with a rate which depends on the state (empty or not) of the other queue.)

Consequently the results obtained in [6] apply and it is seen that, for $\rho < 1$, $F_q(x,y)$ is the *unique* solution - analytic in $|x| < 1$, $|y| < 1$, continuous in $|x| \le 1$, $|y| \le 1$ - of the following functional equation :

$$K(x,y)F(x,y) = R(x,y)[\xi F(0,y) - (1-\xi)F(x,0)] + S(x,y)(1-\rho), \tag{4.7}$$

for $|x| \le 1$, $|y| \le 1$, where

$$K(x,y) := \lambda_1(1-x) + \lambda_0(1-y) + \xi\mu_1(1-\frac{1}{x}) + (1-\xi)\mu_0(1-\frac{1}{y});$$

$$R(x,y) := \mu_1(1-\frac{1}{x}) - \mu_0(1-\frac{1}{y});$$

$$S(x,y) := (1-\xi)\mu_1(1-\frac{1}{x}) + \xi\mu_0(1-\frac{1}{y}).$$

Differentiating twice (4.7) in the variable y and letting $(x,y) = (1,1)$ yields, cf. (4.1),

$$V(f_q) = \frac{\lambda_0 - \mu_0 \frac{\partial}{\partial y}F_q(0,y)\Big|_{y=1}}{\mu_0(1-\xi)-\lambda_0}, \qquad \text{for } \lambda_0 \neq \mu_0(1-\xi). \tag{4.8}$$

Starting now from the obvious relation

$$V(f_q) = \frac{\partial}{\partial x}F_q(x,x)\Big|_{x=1} - \frac{\partial}{\partial x}F_q(x,y)\Big|_{(x,y)=(1,1)},$$

we readily obtain a relation similar to (4.8),

$$V(f_q) = \frac{\lambda_0 + (\mu_1-\mu_0)\frac{\partial}{\partial y}F_q(0,y)\Big|_{y=1} + (\mu_1\xi + \mu_0(1-\xi)\frac{\partial}{\partial x}F_q(x,0)\Big|_{x=1}}{\mu_1\xi - \lambda_1}, \qquad \text{for } \lambda_0 = \mu_0(1-\xi). \tag{4.9}$$

(Note that the denominator in (4.9) cannot vanish under the condition $\rho < 1$.)

In [6] Fayolle and Iasnogorodski explicitly obtained the functions $F_q(x,0)$ and $F_q(0,y)$ respectively for $|x| < \sqrt{\frac{\mu_1\xi}{\lambda_1}}$ and $|y| < \sqrt{\frac{\mu_0(1-\xi)}{\lambda_0}}$. These results are given in Appendix with the change of notation previously mentioned. (The corresponding formulas (6.3)-(6.5) in [6] are slightly incorrect.)

Consequently for $\lambda_0 < \mu_0(1-\xi)$ the first derivative of $F_q(0,y)$ at point $y = 1$ can be derived from (A.2) (see Appendix), which gives $V(f_q)$ by (4.8).

For $\lambda_0 \geq \mu_0(1-\xi)$ we need to analytically continue the expression involving $F_q(0,y)$ given by (A.2) up to the unit circle, in order to obtain $\frac{\partial}{\partial y}F_q(0,y)\Big|_{y=1}$. (Note that for $\lambda_0 = \mu_0(1-\xi)$ the constant $\frac{\partial}{\partial x}F_q(x,0)\Big|_{x=1}$ also appearing in (4.9) can be directly computed from (A.1), since in that case we necessarily have $\lambda_1 < \mu_1\xi$ under the condition $\rho < 1$.)

This analytic continuation - which is not explicitly done in [6] - can be carried out as follows : using Theorem 3.1 and relation (3.2) in [6, pp.332- 333], we obtain for $\sqrt{\frac{\mu_0(1-\xi)}{\lambda_0}} \leq |y| \leq 1$,

$$F_q(0,y) = \frac{1}{\xi}[(1-\xi)F_q(k(y),0) - \frac{S(k(y),y)}{R(k(y),y)}(1-\rho)], \tag{4.10}$$

with

$$k(y) := \frac{\lambda_1 + \mu_1\xi + a(y) - \sqrt{(\lambda_1 + \mu_1\xi + a(y))^2 - 4\lambda_1\mu_1\xi}}{2\lambda_1};$$
$$a(y) := \lambda_0(1-y) + \mu_0(1-\xi)(1-\frac{1}{y}).$$

It is shown in [6, Lemma 2.2] that $|k(y)| < \sqrt{\frac{\mu_1\xi}{\lambda_1}}$ for $\sqrt{\frac{\mu_0(1-\xi)}{\lambda_0}} \leq |y| \leq 1$. Consequently for $\sqrt{\frac{\mu_0(1-\xi)}{\lambda_0}} \leq |y| \leq 1$, the function $F_q(k(y),0)$ in (4.10) can be computed with (A.1) (see Appendix). (The continuity of (4.10) and (A.2) (see Appendix) on the circle $|y| = \sqrt{\frac{\mu_0(1-\xi)}{\lambda_0}}$ can be proved using *Plemelj-Sokhotski formulas*, as done in [10, Lemma 5.1]. This not done here for sake of clarity.)

The above results can be summarized as follows :

- for $\lambda_0 < \mu_0(1-\xi)$, $V(f_q)$ is obtained from (4.8), (A.2);
- for $\lambda_0 \geq \mu_0(1-\xi)$, $V(f_q)$ is obtained from (4.8) (or from (4.9) if $\lambda_0 = \mu_0(1-\xi)$) and (4.10).

In Figure 1 we have plotted $V(f_q)$ versus the bias factor q, for two families of model parameters. The shape of these curves appears to be typical, since in each case the saddle point is exactly located at point q such that $\lambda_0 = (1-\xi)\mu_0$, which actually turns out to be a special point both from an analytical (see the present section) and queueing point of view. Table 1 gives the value of the optimal bias factor - called q_{opt} - for varying values of the constraint α.

Remark 4.1. The functional equation (4.7) can also be derived by considering the Markov process $\{(X(t),I(t)),\ t \geq 0\}$, where I(t) indicates the customer type being served at time t, if any. However this method yields lengthy computations which can be avoided in our case by noting the similarities of the transition matrix $\widehat{Q}$ with the corresponding matrix in [6].

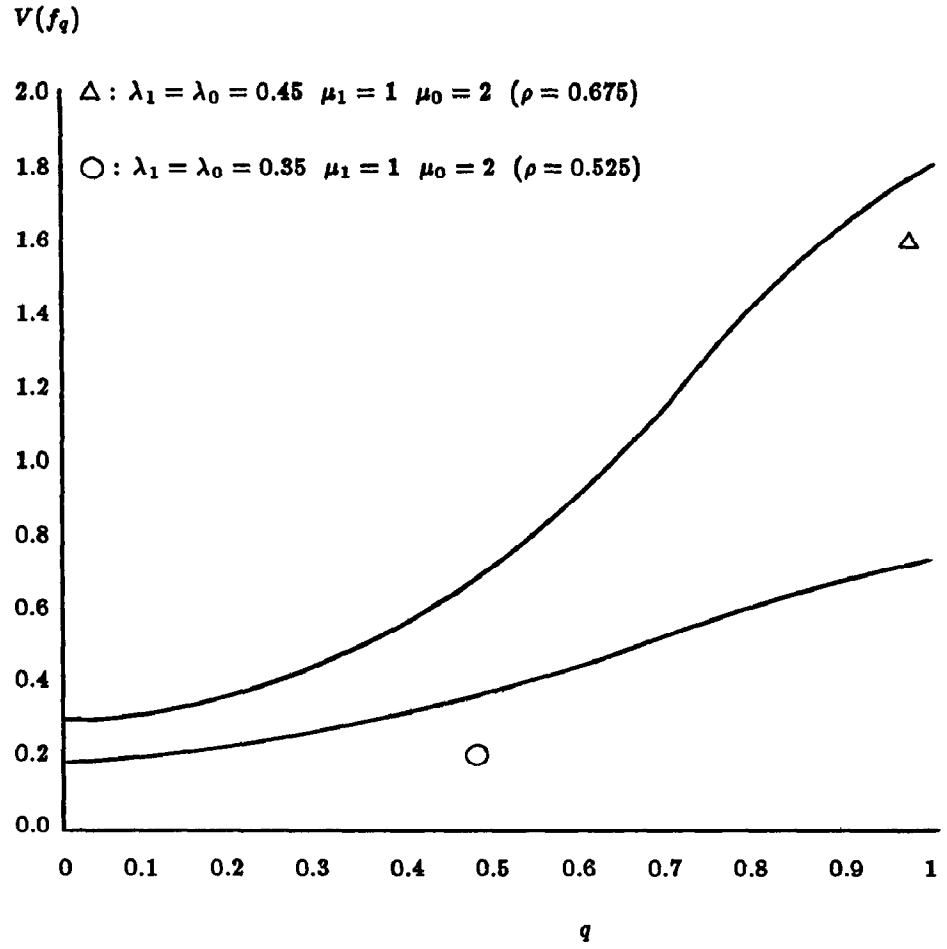

Figure 1
$V(f_q)$ versus q for $K = 1$

Case $\bigcirc$:	$\lambda_1 = 0.35$	$\lambda_0 = 0.35$	$\mu_1 = 1$	$\mu_0 = 2$	$(\rho = 0.525)$
		$V_{f_0} = 0.212$	$V_{f_1} = 0.765$		
	$p = 0.2$	$(\alpha = 0.6546)$	$\Longrightarrow$	$q_{opt} = 0.847$	
	$p = 0.4$	$(\alpha = 0.5440)$	$\Longrightarrow$	$q_{opt} = 0.707$	
	$p = 0.6$	$(\alpha = 0.4333)$	$\Longrightarrow$	$q_{opt} = 0.555$	
	$p = 0.8$	$(\alpha = 0.3227)$	$\Longrightarrow$	$q_{opt} = 0.351$	
Case Δ :	$\lambda_1 = 0.45$	$\lambda_0 = 0.45$	$\mu_1 = 1$	$\mu_0 = 2$	$(\rho = 0.675)$
		$V_{f_0} = 0.290$	$V_{f_1} = 1.825$		
	$p = 0.2$	$(\alpha = 1.5182)$	$\Longrightarrow$	$q_{opt} = 0.835$	
	$p = 0.4$	$(\alpha = 1.2112)$	$\Longrightarrow$	$q_{opt} = 0.709$	
	$p = 0.6$	$(\alpha = 0.9043)$	$\Longrightarrow$	$q_{opt} = 0.588$	
	$p = 0.8$	$(\alpha = 0.5973)$	$\Longrightarrow$	$q_{opt} = 0.417$	

$$(\alpha = (1-p)V(f_1) + pV(f_0)\ ,\ p \in [0,1])$$

Table 1
Computation of the optimal bias factor q_{opt} for varying values of the constraint α

APPENDIX

For $\rho < 1$, $|x| < \sqrt{\frac{\mu_1 \xi}{\lambda_1}}$, $|y| < \sqrt{\frac{\mu_0(1-\xi)}{\lambda_0}}$, $q \in (0,1)$,

$$F_q(x,0) = \frac{2\sqrt{\lambda_1 \mu_1 \xi}}{\pi} \int_0^{\pi} \frac{x \sin\theta \, \eta(\theta)\, d\theta}{\xi\mu_1 + \lambda_1 x^2 - 2x\sqrt{\lambda_1 \mu_1 \xi}\cos\theta} + 1 - \rho; \tag{A.1}$$

$$F_q(0,y) = \frac{2\sqrt{\lambda_0 \mu_0 (1-\xi)}}{\pi} \int_0^{\pi} \frac{y \sin\theta \, \nu(\theta)\, d\theta}{(1-\xi)\mu_0 + \lambda_0 y^2 - 2y\sqrt{\lambda_0 \mu_0 (1-\xi)}\cos\theta} + 1 - \rho, \tag{A.2}$$

where

$$\nu(\theta) := \frac{-\mu_1 \left(1 - \frac{1}{h(\theta)}\right) \sqrt{\frac{\mu_0(1-\xi)}{\lambda_0}} (1-\rho) \sin\theta}{\xi \left(\left[\mu_0 \sqrt{\frac{\mu_0(1-\xi)}{\lambda_0}} \cos\theta - \mu_1 (1 - \frac{1}{h(\theta)}) \right]^2 + \frac{\lambda_0 \mu_0}{1-\xi} \sin^2\theta \right)}; \tag{A.3}$$

$$h(\theta) := \frac{\lambda_1 + \mu_1 \xi + r(\theta) - \sqrt{[(\sqrt{\lambda_1} + \sqrt{\mu_1 \xi})^2 + r(\theta)][(\sqrt{\lambda_1} - \sqrt{\mu_1 \xi})^2 + r(\theta)]}}{2\lambda_1}; \tag{A.4}$$

$$r(\theta) := \lambda_0 + \mu_0(1-\xi) - 2\sqrt{\lambda_0 \mu_0 (1-\xi)}\cos\theta. \tag{A.5}$$

$\eta(\theta)$ is obtained from (A.3), (A.4), (A.5) by interchanging indicies 0 and 1 and by replacing ξ by $1 - \xi$; ξ is given in (4.6).

REFERENCES

[1] Baras, J. S., Dorsey A. J. and Makowski, A. M. (1985) Two Competing Queues with Linear Costs : the μc rule is often optimal. *Adv. App. Prob.* **17**, 186-209.

[2] Buyukkoc, C., Varaiya, P. and Walrand, J. (1985) The $c\mu$ Rule Revisited. *Adv. App. Prob.* **17**, 237-238.

[3] Borkar, V. S. (1983) Controlled Markov Chains and Stochastic Networks. *SIAM J. of Control* **21**, 652-666.

[4] Cinlar, E. (1975) *Introduction to Stochastic Processes*. Prentice-Hall, Englewood Cliffs, NJ.

[5] Cohen, J.W. (1983) *The Single Server Queue*. 2nd edition, North Holland Publ. Comp.

[6] Fayolle, G. and Iasnogorodski, R. (1979) Two Coupled Processors : the Reduction to a Riemann-Hilbert Problem. *Z. Wahrschein.* **47**, 325-351.

[7] Gelenbe, E. and Mitrani, I. (1980) *Analysis and Synthesis of Computer Systems*. Academic Press, London.

[8] Klimov, G. P. (1974) Time-Sharing Service System I. *Theory of Prob. and its Appl.* **19**, 532-551.

[9] Maglaris, B. and Schwartz, M. (1982) Optimal Fixed Frame Multiplexing in Integrated Line and Packet-Switched Communication Networks. *IEEE Trans. on Information Theory* IT-28, 263-273.

[10] Nain, P. (1986) On a Generalization of the Preemptive Resume Priority. *Adv. Appl. Prob.* **18**, No. 1.

[11] Nain, P. and Ross, K. W. (1985) Optimal Priority Assignment with Hard Constraint. *INRIA internal report* No. **459**. Submitted to IEEE Automatic Control.

[12] Ross, K. W. (1985) *Constrained Markov Decision Processes with Queueing Applications*. Thesis, CICE Program, University of Michigan.

[13] Ross, S. (1971) *Applied Probability Models with Optimization Applications*. Holden-Day, San Francisco.

[14] Rykov, V. V. and Lembert, E. Ye. (1967) Optimal Dynamic Priorities in Single-Line Queueing Systems. *Engrg. Cybernet.* **5**, 21-30

[15] Stidham Jr., S. (1985) Optimal Control of Admissions to a Queueing System. *IEEE Trans. Automat. Contr.*, Vol. AC-30, No. 8, 705-714.