Probability in the Engineering and Informational Sciences, **14**, 2000, 9–26. Printed in the U.S.A.

INDIVIDUAL EQUILIBRIUM DYNAMIC ROUTING IN A MULTIPLE SERVER RETRIAL QUEUE

ANTHONY C. BROOMS
Department of Engineering Science
University of Oxford
Oxford OX1 3PJ, United Kingdom
E-mail: anthony.brooms@eng.ox.ac.uk

Customers arrive sequentially to a service system where the arrival times form a Poisson process of rate λ. The system offers a choice between a private channel and a public set of channels. The transmission rate at each of the public channels is faster than that of the private one; however, if all of the public channels are occupied, then a customer who commits itself to using one of them attempts to connect after exponential periods of time with mean μ^{-1}. Once connection to a public channel has been made, service is completed after an exponential period of time, with mean ν^{-1}. Each customer chooses one of the two service options, basing its decision on the number of busy channels and reapplying customers, with the aim of minimizing its own expected sojourn time. The best action for an individual customer depends on the actions taken by subsequent arriving customers. We establish the existence of a unique symmetric Nash equilibrium policy and show that its structure is characterized by a set of threshold-type strategies; we discuss the relevance of this concept in the context of a dynamic learning scenario.

1. INTRODUCTION

Customers arrive sequentially to a system consisting of a multiple server retrial queue (see Figure 1) and a $\bullet/\mathrm{GI}/\infty$ queue; the former queue may be considered to be a set of public channels with customers reattempting to make a connection if blocked by a previous attempt to connect to one of these, denoted by Q_R, and the latter queue as a private channel, denoted by Q_P. Upon arrival, each customer chooses one of the following two options: either to connect to Q_R or to Q_P. Each customer is solely

© 2000 Cambridge University Press 0269-9648/00 .50

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

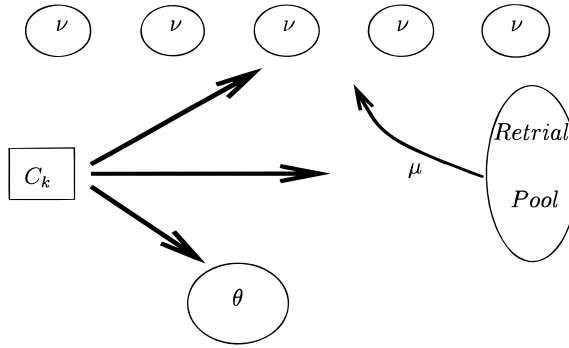


FIGURE 1. Multiple server retrial queue.

concerned with minimizing his own expected sojourn time. The number of retrialing customers and busy channels at Q_R may be observed by each arriving customer. The sojourn time of a particular customer in the retrial pool may be affected by the number of customers entering Q_R during its sojourn, which, in turn, is affected by the decisions of subsequent arriving customers.

The purpose of this paper is to present a series of stochastic order results which will allow us to infer certain structural results pertaining to the Nash equilibrium solution concept. We also briefly discuss the relevance of the Nash equilibrium in assessing the performance of the system under realistic operating conditions.

A similar *state-dependent* admission control problem to the one examined in this paper has been considered by Altman and Shimkin [4], in which their service system consists of a processor sharing queue and a $\bullet/\mathrm{GI}/\infty$ queue. They also present a dynamic learning scheme in which customers base their decisions on data collected by the system up to the point of decision; they demonstrate that a Nash equilibrium solution characterizes the long-term behavior of the system under this scheme. Generalizations of the model of [4] and the associated learning scheme were also considered by Ben-Shabar et al. [5], and Buche and Kushner [7], respectively.

Although similar to the overall paradigm of [4] and a few other models considered in the work of Brooms [6], our current work does include some important differences. For example, our model is two dimensional; as a consequence, we often have to condition on the number of public channels available to an arriving customer. We also find that properties of the stochastic processes relating to the number of busy channels and retrialing customers are established simultaneously, by exploiting the *forward induction* technique.

Other state-dependent game-theoretic queueing models have also been considered. Altman and Shimkin [2] consider the problem of socially optimal and individual equilibrium routing into a system of parallel queues. One of the models they consider is one in which the total service rate in all queues is constant, but the proportion of the service rate assigned to each queue varies with time. They also

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

consider a model in which there is a single server which always serves the longest queue. A *zero-sum* stochastic game analysis for the first model yields that the socially optimal strategy for the router is to *join the shortest queue*. Nash equilibrium strategies are shown to characterize the second model.

Altman [1] examines the role of an overall controller routing traffic into a two-queue system, where the service rate in each queue [given by $a_1(i)$, $i = 1,2$] is, in general, state and time dependent and unknown to the controller. The decision to be made by the controller for each new arrival is to decide whether to send it to queue 1 or 2 in order to minimize finite horizon and infinite horizon discounted cost criteria. The existence of optimal strategies of the router and of the server described by monotone switching curves are established.

A number of authors have considered customer retrial models, including Cohen [8], Keilson et al. [15], Falin [10], Hanschke [14], Falin and Fricker [11], Greenberg and Wolff [13], and Liang and Kulkarni [16]. A comprehensive survey of the subject can be found in the work of Falin and Templeton [12]. Tran-Gia and Mandjes [18] have examined the customer retrial phenomenon within a cell of a cellular mobile network, looking at both finite and infinite population models.

An outline of the rest of the paper is as follows. In the next section, we describe the model under consideration. In Section 3, we present results on the expected sojourn time in the multiple-server retrial system; these results are concerned with strict monotonicity and continuity which are used in establishing the main results of the section. We go on to show that an optimal strategy for each customer conditioned on the strategies of the other users is to adopt a threshold-type strategy.

The existence and uniqueness of a symmetric Nash equilibrium routing policy are established by invoking a related theorem from Altman and Shimkin [4]. A procedure for computing the symmetric Nash equilibrium is presented in Section 4. Finally, we discuss the relevance of the Nash equilibrium in modeling the performance of the system under an intuitively appealing dynamic learning rule.

2. THE MODEL

An arriving customer has to choose between a multiple-server retrial queue (denoted by Q_R) consisting of s servers or transmission channels along with a retrial pool (shared by other customers), or a private channel Q_P, where both options make up the service system.

Let A_k be the arrival time of customer C_k, where $0 = A_0 < A_1 < A_2 < \cdots$ and where $\{A_1, A_2 - A_1, A_3 - A_2, \ldots\}$ are i.i.d. exponential with mean $\lambda^{-1} < \infty$ (i.e., the arrivals form a Poisson process at rate λ).

A customer who connects to a channel in Q_R experiences an exponential transmission time with mean ν^{-1}; a channel may only be occupied by one customer at a time. If a customer decides to join Q_R, then it will try to connect to a public channel immediately; however, if a public channel is unavailable, the customer tries again to connect to one after an exponential period of time with mean μ^{-1}, and continues to do so until a connection is made.

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

The expected transmission time in Q_P is $\theta^{-1} < \infty$. Transmission times in Q_P are mutually independent, and independent of the behavior in Q_R.

Let the state of the system at time t be described by a pair $N(t) = (X(t), L(t))$, where $X(t)$ represents the number of retrialing or "milling" customers and $L(t)$ the number of occupied or "busy" channels.

We assume that there is a buffer of size M containing the milling customers, which may be either finite or infinite.

Customer C_k chooses which queue to join after observing $N(A_k)$, the system state just *prior* to its arrival.

A randomized decision rule $u_k(x,y) \in [0,1]$ defines the probability of joining Q_R if the state $N(A_k) = (x,y)$ is observed. Let $\mathbb{U}$ denote the collection of such decision rules. The vector of decision rules of all customers (i.e., a policy) is denoted by $\pi = (u_0, u_1, \ldots)$, with $\pi \in \mathbb{U}^\infty$.

Let $\{U_k\}_{k\geq 0}$ be a set of uniform random variables on the interval $[0,1]$ which are used in determining the decisions of arriving customers. So, for customer C_k arriving at time A_k, who joins Q_R with probability q (which depends on its strategy), U_k influences the decision in the following way: C_k joins Q_S if and only if $U_k \leq q$.

The sojourn time of a customer in the system is given by the time interval between arrival to and departure from the system. Denote the sojourn time of C_k by w_k. Also, given that C_k joins Q_R, denote its total milling time by m_k and its total sojourn time by v_k. Let $W_k(x,y,\pi)$ and $V_k(x,y,\pi)$ be the expected values of w_k and v_k, respectively, given that, upon arrival to the system, the customer finds x milling customers and y servers busy in Q_R and that customers are following the policy π. It follows that

$$W_k(x,y,\pi) = u_k(x,y)V_k(x,y,\pi) + (1 - u_k(x,y))\theta^{-1}.$$

Each customer is interested in minimizing its own expected sojourn time and is unconcerned with the quality of service experienced by any other customer. We consider this problem as a noncooperative game consisting of an infinite number of *players.*

3. ON THE EXISTENCE OF A UNIQUE SYMMETRIC NASH EQUILIBRIUM SOLUTION

For an arbitrary policy $\pi = (u_0, u_1, \ldots)$, let π^{-k} denote the policy that results by excluding the rule u_k of C_k, and let $[\pi^{-k}|u'_k]$ be the policy that results when u_k is replaced by u'_k. We say that *a decision rule u_k of C_k is optimal against the policy π^{-k} if*

$$W_k(x,y,[\pi^{-k}|u_k]) \leq W_k(x,y,[\pi^{-k}|u'_k]) \quad \forall x \in \mathbb{N}, y \in \{0,\ldots,s\}, u'_k \in \mathbb{U}.$$

A policy $\pi = (u_0, u_1, \ldots)$ is said to be a *Nash equilibrium policy* if, for every $k \geq 0$, u_k is optimal against π^{-k}.

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

3.1. Monotonicity Properties of the Expected Sojourn Time in Q_R When All Channels Are Busy

In this section, we show that if all the public channels are busy, then the expected sojourn time in Q_R of a new customer is strictly monotone increasing in the number of milling customers on its arrival. We first establish stochastic dominance relations on a certain pair of processes in order to establish the result of interest.

In the proofs to follow, we will assume that the buffer size $M = \infty$; however, the proofs for $M < \infty$ go through with only minor modification.

Assume C_0 joins Q_R. Then we define two processes as follows:

the x-process: with initial state $X(0) = x$ and $L(0) = s$

the $(x + 1)$-process: with initial state $\widetilde{X}(0) = x + 1$ and $\widetilde{L}(0) = s$ [quantities pertaining to the $(x + 1)$-process will have a tilde].

Under the above conditions, we claim the following lemma.

LEMMA 1: $L(t) \leq \widetilde{L}(t)$ *stochastically* $\forall t$.

PROOF OF LEMMA 1: The proof proceeds via a forward induction argument (see Ephremides et al. [9] for example). An event can either be an arrival, a potential reapplication for a channel, or a channel potentially becoming vacant.

Define the random variables $\{t_i\}$ to be the event times in the $(x + 1)$-process. Couple the processes by using the same set of arrival times $\{A_k\}_{k\geq 0}$ and random variables $\{U_k\}_{k\geq 0}$ for both processes.

To assist the induction, it will also prove useful to establish that $X(t) \leq \widetilde{X}(t)\ \forall t$. By the initial conditions, it follows that

$$\begin{aligned} X(0^+) &= x + 1 < x + 2 = \widetilde{X}(0^+), \\ L(0^+) &= s = \widetilde{L}(0^+). \end{aligned} \tag{1}$$

Assume that at time t_k^+, $k \geq 0$,

$$\begin{aligned} h &= X(t_k^+) \leq \widetilde{X}(t_k^+) = i, \\ n &= L(t_k^+) \leq \widetilde{L}(t_k^+) = o. \end{aligned} \tag{2}$$

Due to the lack of memory property of the exponential distribution, we may further couple the processes in the following way.

For the x-process, assign each of the h milling customers and n busy channels with the random variables $\{S_j\}_{j=1}^{h}$ and $\{D_j\}_{j=1}^{n}$, respectively. The $\{S_j\}$ and $\{D_j\}$ are from the exponential distributions with means μ^{-1} and ν^{-1}, respectively.

For the $(x+1)$-process, assign h milling customers and n busy channels with the random variables of the x-process. Assign the remaining milling customers and busy channels with random variables from the sets $\{S_j'\}_{j=1}^{i-h}$ and $\{D_j'\}_{j=1}^{o-n}$, respectively. The $\{S_j'\}$ and $\{D_j'\}$ are also from the exponential distribution with means μ^{-1} and ν^{-1}, respectively.

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

We say that milling customers or channels occupied by customers from each of the processes are "paired" if they have the same random variable; otherwise, a customer is "unpaired." At time t_{k+1}, one of the following cases must occur: (a) a customer arrives, (b) a channel potentially becomes available, or (c) a milling customer potentially makes an application for one of the channels. We consider the state of Q_R in the light of these three cases in order to show that

$$\begin{aligned} X(t_{k+1}^+) &\le \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &\le \widetilde{L}(t_{k+1}^+). \end{aligned} \tag{3}$$

Case (a): If $o < s$, then

$$\begin{aligned} X(t_{k+1}^+) &= h \le i = \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &= n+1 \le o+1 = \widetilde{L}(t_{k+1}^+). \end{aligned} \tag{4}$$

If $n < o = s$, then

$$\begin{aligned} X(t_{k+1}^+) &= h \le i \le \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &= n+1 \le s = o = \widetilde{L}(t_{k+1}^+). \end{aligned} \tag{5}$$

If $n = o = s$ and $h < i$, then

$$\begin{aligned} X(t_{k+1}^+) &\le h+1 \le i \le \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &= n = s = o = \widetilde{L}(t_{k+1}^+). \end{aligned} \tag{6}$$

However, if $n = o = s$ and $h = i$, then

$$\begin{aligned} X(t_{k+1}^+) &= \widetilde{X}(t_{k+1}^+) = h \text{ or } h+1, \\ L(t_{k+1}^+) &= n = s = o = \widetilde{L}(t_{k+1}^+), \end{aligned} \tag{7}$$

since the arriving customer will be faced with identical states in both processes.

Case (b): If $n = o = 0$, then the state remains unchanged at time t_{k+1}^+. Otherwise, suppose a paired customer vacates a channel. It follows that

$$\begin{aligned} X(t_{k+1}^+) &= h \le i = \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &= n-1 \le o-1 = \widetilde{L}(t_{k+1}^+). \end{aligned} \tag{8}$$

Suppose that an unpaired customer vacates a channel. Necessarily $n < o$ and

$$\begin{aligned} X(t_{k+1}^+) &= h \le i = \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &= n \le o-1 = \widetilde{L}(t_{k+1}^+). \end{aligned} \tag{9}$$

Case (c): If $h = i = 0$, then the state remains unchanged at time t_{k+1}^+. Otherwise, suppose a paired milling customer applies for a channel.

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

If $n = o = s$, then

$$\begin{aligned} X(t_{k+1}^+) &= h \le i = \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &= n = s = o = \widetilde{L}(t_{k+1}^+), \end{aligned} \tag{10}$$

since the customer is blocked under both processes.

If $o < s$, then

$$\begin{aligned} X(t_{k+1}^+) &= h - 1 \le i - 1 = \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &= n + 1 \le o + 1 = \widetilde{L}(t_{k+1}^+). \end{aligned} \tag{11}$$

If $n < o = s$, then

$$\begin{aligned} X(t_{k+1}^+) &= h - 1 < i = \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &= n + 1 \le s = o = \widetilde{L}(t_{k+1}^+), \end{aligned} \tag{12}$$

since the customer is blocked in the $(x + 1)$-process only.

Suppose an unpaired milling customer applies for a channel [which is necessarily from the $(x + 1)$-process]. Then, clearly, $h < i$. If $o = s$, then the customer is blocked and the state of the system at time t_{k+1}^+ remains unchanged. However, if $o < s$, then

$$\begin{aligned} X(t_{k+1}^+) &= h \le i - 1 = \widetilde{X}(t_{k+1}^+), \\ L(t_{k+1}^+) &= n < o + 1 = \widetilde{L}(t_{k+1}^+). \end{aligned} \tag{13}$$

■

We now use Lemma 1 to establish the following intuitively appealing result.

LEMMA 2: *For every policy* π, $k \ge 0$, *and* $x \in \mathbb{N}$, $V_k(x + 1, s, \pi) - V_k(x, s, \pi) > \delta_x$ *for* δ_x *positive*.

PROOF OF LEMMA 2: Consider customer C_0 without loss of generality. Assume that it has joined Q_R. Observe that $V_0(x, s, \pi) = E[v_0] = E[m_0] + \nu^{-1}$, and that $V_0(x + 1, s, \pi) = E[\tilde{m}_0] + \nu^{-1}$. Thus, it is enough to show that

$$E[m_0] < E[\tilde{m}_0]. \tag{14}$$

Let the potential applications of C_0 for a channel in the x-process be governed by the $\{R_i\}$ (which correspond to actual applications as long as C_0 is still in the retrial pool), where $\{R_1, R_2 - R_1, R_3 - R_2, \ldots\}$ are i.i.d. exponential with mean μ^{-1}. Also, define $N(s)$ to be the number of potential applications for service which arise in the time interval $(0, s]$. Define $\{\tilde{R}_i\}$ and $\widetilde{N}(s)$ in a similar way for the $(x + 1)$-process. Now, couple the two processes by setting $R_i = \tilde{R}_i \ \forall i$. Then,

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

$$\mathbb{P}\{m_0 > t\} = \sum_{n=0}^{\infty} \mathbb{P}\{m_0 > t | N(t) = n\}\mathbb{P}\{N(t) = n\}$$

$$= \sum_{n=0}^{\infty} \left(\int_{A_{n,t}} \mathbb{P}\{m_0 > t | N(t) = n, R_i = r_i, i = 1, \ldots, n\} \right.$$

$$\left. \times f_{\underline{R}}(\underline{r} | N(t) = n)\, dr_1\, dr_2 \ldots dr_n \right) \mathbb{P}\{N(t) = n\}, \qquad \textbf{(15)}$$

where $f_{\underline{R}}(\underline{r}|N(t) = n)$ is the probability density function of $\underline{R} = (R_1, R_2, \ldots, R_n)$ given $N(t) = n$, and $A_{n,t} := \{(r_1, r_2, \ldots, r_n) | 0 < r_1 < \cdots < r_n \leq t\}$. However, (15) is equivalent to

$$\sum_{n=0}^{\infty} \left(\int_{A_{n,t}} \mathbb{P}\{L(r_i) = s, i = 1, \ldots, n | N(t) = n, R_i = r_i, i = 1, \ldots, n\} \right.$$

$$\left. \times f_{\underline{R}}(\underline{r} | N(t) = n)\, dr_1\, dr_2 \ldots dr_n \right) \mathbb{P}\{N(t) = n\}. \qquad \textbf{(16)}$$

However,

$$\mathbb{P}\{L(r_i) = s, i = 1, \ldots, n | N(t) = n, R_i = r_i, i = 1, \ldots, n\}$$

$$\leq \mathbb{P}\{\tilde{L}(r_i) = s, i = 1, \ldots, n | N(t) = n, R_i = r_i, i = 1, \ldots, n\}, \qquad \textbf{(17)}$$

since Lemma 1 holds on the restricted set $\{N(t) = n, R_i = r_i, i = 1, \ldots, n\}$. Observe that due to the coupling,

$$\{N(t) = n\} \equiv \{\tilde{N}(t) = n\},$$

$$\{R_i = r_i, i = 1, \ldots, n\} \equiv \{\tilde{R}_i = r_i, i = 1, \ldots, n\}.$$

It therefore follows that (17) is equivalent to

$$\mathbb{P}\{\tilde{L}(r_i) = s, i = 1, \ldots, n | \tilde{N}(t) = n, \tilde{R}_i = r_i, i = 1, \ldots, n\}$$

and that $f_{\underline{R}}(\underline{r}|N(t) = n) \equiv f_{\underline{\tilde{R}}}(\underline{r}|\tilde{N}(t) = n)$. Hence, (16) is bounded above by $\mathbb{P}\{\tilde{m}_0 > t\}$, which implies that $E[m_0] \leq E[\tilde{m}_0]$.

To establish strict monotonicity, consider the event B_x defined solely in terms of the x-process:

i. A (paired) customer vacates a channel,
ii. C_0 applies for a channel,

where (i) and (ii) are the first events to occur after the arrival of C_0, and the coupling of Lemma 1 is implemented, except that we require that C_0 *not* be paired with itself across the two processes just prior to (ii) (which is always possible to do).

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

When (ii) occurs, C_0 connects to a public channel in the x-process, but is still in the retrial pool in the $(x+1)$-process. B_x is clearly an event which occurs with positive probability. Therefore,

$$E[\tilde{v}_0 - v_0] = E[\tilde{m}_0 - m_0] \geq E[\tilde{m}_0 - m_0 | B_x]\mathbb{P}\{B_x\} > 0,$$

since $m_0 < \tilde{m}_0$ on B_x and $E[\tilde{m}_0 - m_0 | B_x^c] \geq 0$. To justify the latter observation, we note that since m_0 is stochastically less than or equal to $\tilde{m}_0$, there exists a coupling for which $m_0 \leq \tilde{m}_0$ almost surely; taking expectations under this coupling over the set B_x^c leads to the desired result. ■

Next, we show that if all customers adopt precisely the same decision rule and, moreover, that it is of *threshold type*, then the expected sojourn time is strictly increasing in that threshold. Randomized decision rules are acceptable in this setup and this concept is made more precise in the following definition.

DEFINITION: *An $[L,q]$ threshold-type strategy $u \in \mathbb{U}$ is given by*

$$u(x,y) = \begin{cases} 1 & \text{if } y < s \\ 1 & \text{if } x < L,\ y = s \\ q & \text{if } x = L,\ y = s \\ 0 & \text{if } x > L,\ y = s \end{cases} \tag{18}$$

for $0 \leq q \leq 1$, where $u(x,y)$ defines the probability of entering the system given that the state is (x,y) just prior to its arrival. So, a customer enters Q_R if either $y < s$, or $y = s$ and $x < L$. A customer enters with probability q if $y = s$ and $x = L$; otherwise, the customer joins Q_P.

An $[L,q]$ threshold-type strategy may be represented more compactly by $[g]$, where $g = L + q$.

LEMMA 3: *$\forall k \geq 0$ and $x \in \mathbb{N}$, $V_k(x,s,[g]^\infty)$ is strictly increasing in $g \in [0,M]$.*

PROOF OF LEMMA 3: Let g and $\tilde{g}$ be two thresholds with $g < \tilde{g}$. We wish to show that

$$V_0(x,s,[g]^\infty) < V_0(x,s,[\tilde{g}]^\infty) \qquad \forall x \in \mathbb{N}.$$

We consider two processes:

the g-process: with $\pi = [g]^\infty$,

the $\tilde{g}$-process: with $\pi = [\tilde{g}]^\infty$,

where both processes start from initial state $N(0) = (x,s)$ and C_0 joins Q_R. Establishing that $L(t) \leq \tilde{L}(t)$ stochastically $\forall t$ is done in an identical manner to the proof of Lemma 1, with the following modification to case (a): If $n = o = s$ and $h = i$, then either (7) holds or

$$\begin{aligned} X(t_{k+1}^+) &= h, \qquad \tilde{X}(t_{k+1}^+) = h+1, \\ L(t_{k+1}^+) &= n = s = o = \tilde{L}(t_{k+1}^+), \end{aligned} \tag{19}$$

since the difference in the thresholds may lead to a customer being admitted in the $\tilde{g}$-process but rejected in the g-process.

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

The rest of the proof is almost the same as Lemma 2. To establish strict monotonicity, consider the following event D in which

i. $L - x$ arrivals occur before the first public channel is vacated with $U_{L-x} \in (q, 1 - (1 - \tilde{q})I\{L = \tilde{L}\}]$,
ii. a public channel is vacated,
iii. C_0 applies for a public channel in the g-process

in that order, where (i)–(iii) are the first events to occur after the arrival of C_0, and the coupling of Lemma 1 is implemented while ensuring that C_0 is *not* paired with itself in the two processes just prior to (iii) (which is always possible).

As before, by conditioning on the event D, it may be shown that $E[\tilde{v}_0 - v_0] > 0$ as required. ■

3.2. Continuity of the Expected Sojourn Time

We will make use of Lemmas 2 and 3 to establish the following continuity result.

LEMMA 4: *For every $k \geq 0$ and $x \in \mathbb{N}$, $V_k(x, s, [g]^\infty)$ is continuous in $g \in [0, M]$.*

PROOF OF LEMMA 4: Consider customer C_0 without loss of generality, and the processes as defined in the previous lemma. Let $g < \tilde{g}$ be such that $g = L + q$ and $\tilde{g} = L + \tilde{q}$, with $0 \leq q < \tilde{q} \leq 1$. It is enough to show that $\forall \varepsilon > 0$, there exists a $\delta > 0$ such that, for $\tilde{g} - g < \delta$,

$$V_0(x, s, [\tilde{g}]^\infty) - V_0(x, s, [g]^\infty) < \varepsilon, \tag{20}$$

since $V_0(x, s, [g]^\infty)$ is increasing in g.

Consider the event

$$F = \{\exists k \geq 1 | A_k \leq v_0, N(A_k) = (L, s), q < U_k \leq \tilde{q}\}.$$

Define the stopping time

$$k_0 := \inf\{k \geq 1 | A_k \leq v_0, N(A_k) = (L, s), q < U_k \leq \tilde{q}\},$$

where $\inf \varnothing := \infty$. In fact, if $k_0 = \infty$, then $v_0 = \tilde{v}_0$. We note that the event F is equivalent to $\{k_0 < \infty\}$. Thus,

$$E[\tilde{v}_0 - v_0] = \sum_{k=1}^{\infty} E[\tilde{v}_0 - v_0 | k_0 = k] \mathbb{P}\{k_0 = k\}.$$

Now,

$$E[\tilde{v}_0 - v_0 | k_0 = k] = E[\tilde{v}_0 - A_k | k_0 = k] - E[v_0 - A_k | k_0 = k],$$

which is equivalent to

$$V_0(L, s, [\tilde{g}]^\infty) - V_0(L - 1, s, [g]^\infty).$$

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

Denote this quantity by $\delta(\tilde{g}, g)$. By Lemmas 2 and 3, this quantity is strictly positive. In addition,

$$\begin{aligned}\mathbb{P}\{k_0 = k\} &\le \mathbb{P}\{A_k \le v_0,\, q \le U_k \le \tilde{q}\} \\ &= (\tilde{q} - q)\mathbb{P}\{A_k \le v_0\}.\end{aligned}$$

By noting that $\tilde{g} - g = \tilde{q} - q$, we see that

$$\begin{aligned}E[\tilde{v}_0 - v_0] &\le \delta(\tilde{g}, g)(\tilde{q} - q)\sum_{k=1}^{\infty}\mathbb{P}\{A_k \le v_0\} \\ &= (\tilde{g} - g)\delta(\tilde{g}, g)N(g), \qquad \textbf{(21)}\end{aligned}$$

where $N(g) = E[\sum_{k=1}^{\infty} 1\{A_k \le v_0\}]$ is the expected number of arrivals during the sojourn of C_0 in Q_R. It is easily seen that (21) may be bounded above by

$$(\tilde{g} - g)\delta(L+1, L)N(L+1),$$

as required. ■

3.3. Main Results

The next two results are concerned with the structure of Nash equilibria.

LEMMA 5: *Let π be an arbitrary policy and let U_k^* be the set of strategies of C_k which are optimal against π^{-k}.*

i. *Every $u_k^* \in U_k^*$ is a threshold-type strategy (with finite threshold) and is given by*

$$u_k^*(x,y) = \begin{cases} 1 & \text{if } y < s \\ 1 & \text{if } V_k(x,s,\pi) < \theta^{-1} \\ q_x & \text{if } V_k(x,s,\pi) = \theta^{-1} \\ 0 & \text{if } V_k(x,s,\pi) > \theta^{-1} \end{cases} \qquad \textbf{(22)}$$

$0 \le q_x \le 1$. (Recall that θ^{-1} is the expected service time in Q_P.)

ii. *Consequently, the set U_k^* is given as follows. Let L^* be the smallest nonnegative integer such that $V_k(L^*, s, \pi) \ge \theta^{-1}$. If $L^* \ge M$, then U_k^* consists of the single threshold-type strategy $[M,0]$. If $L^* < M$ and $V_k(L^*, s, \pi) = \theta^{-1}$, then $U_k^* = \{[L,q] | 0 \le q \le 1\}$. Otherwise, U_k^* consists of the single threshold-type strategy $[L^*,0]$.*

PROOF OF LEMMA 5:

i. Suppose that C_k observes y busy channels, where $y < s$. Then, clearly, the *optimal* strategy is to join Q_R, since the customer will be able to connect to a public channel with mean transmission time $\nu^{-1} < \theta^{-1}$. Now suppose that C_k observes x milling customers and $y = s$ (corresponding to all public channels busy) in Q_R at arrival. C_k should now choose to join either Q_R, where the

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

expected sojourn time is $V_k(x,s,\pi)$, or Q_P, where the expected service time is θ^{-1}. Clearly, an optimal strategy chooses the lower one as in Eq. (22). But by Lemma 2, $V_k(x,s,\pi)$ is strictly increasing in x, ensuring that Eq. (22) is indeed a threshold-type rule. $V_k(x,s,\pi) \to \infty$ as $x \to \infty$ by Lemma 7 (see Appendix); hence, there exists a finite x such that $V_k(x,s,\pi) > \theta^{-1}$, which means that the threshold is finite even for $M = \infty$, as required to establish (i).

ii. Part (ii) follows by using (i) and using the fact that $V_k(x,s,\pi)$ is strictly increasing in x. ■

THEOREM 1: *If* $\pi = (u_0,u_1,\ldots,\ldots)$ *is an equilibrium policy, then each* u_k *is a threshold-type strategy.*

PROOF OF THEOREM 1: Suppose $\pi = (u_0,u_1,\ldots,\ldots)$ is an equilibrium policy. Then, by definition, each u_k is optimal against π^{-k}. Therefore, it follows that u_k is a threshold-type strategy by Lemma 5. ■

We can also use previous results to establish the following.

THEOREM 2: *There exists a unique symmetric Nash equilibrium policy* $\pi^* = (g^*,g^*,\ldots,\ldots)$.

PROOF OF THEOREM 2: By replacing $\bar{V}_0(\cdot,[g]^\infty)$ by $V_0(\cdot,s,[g]^\infty)$, and by noting that we are dealing with *threshold-type* rather than *threshold* policies, the proof is almost precisely that of Theorem 1(ii) of Altman and Shimkin [4]. ■

4. COMPUTATIONAL RESULTS

4.1. Characterization of the Nash Equilibrium

The next lemma tells us how to compute the symmetric Nash equilibrium. We consider customer C_0 for concreteness.

LEMMA 6: *The equilibrium threshold* $[L^*,q^*]$ *may be found via the following procedure.*

(a) *Let* L^* *be defined as* $\min\{L \geq 0 | V_0(L,s,[L,1]^\infty) > \theta^{-1}\}$.

(b1) *If* $V_0(L^*,s,[L^*,0]^\infty) \geq \theta^{-1}$, *then the equilibrium threshold is* $[L^*,0]$.

(b2) *If* $V_0(L^*,s,[L^*,0]^\infty) < \theta^{-1}$, *then the equilibrium threshold is* $[L^*,q^*]$, *where* $0 < q^* < 1$ *is the unique solution of*

$$V_0(L^*,s,[L^*,q]^\infty) = \theta^{-1}.$$

PROOF OF LEMMA 6: The proof is precisely that of Lemma 4 of Altman and Shimkin [4] with $V_0(L^*,[L^*,q^*]^\infty)$ replaced by $V_0(L^*,s,[L^*,q^*]^\infty)$. ■

4.2. Calculating the Expected Sojourn Time

In this section, we show that by conditioning on the next event to take place after the arrival of a customer, we can compute $V(x,y,[g]^\infty)$ for any given policy $[g]^\infty$,

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

where $g = L + q$. In particular, we are going to make use of dummy events, in a way similar to Lippman [17].

Let $Z(x,y)$ be the expected remaining milling time of a customer in the retrial pool, given that there are x milling customers including C_0, and y busy channels. Then, for $[L,q]$ fixed, $\alpha := \lambda + \mu(L+1) + s\nu$, $Z(x,y)$ is the solution to the following set of $(S+1)(L+1)$ linear equations in $(S+1)(L+1)$ unknowns:

$$Z(x,0) = \frac{1}{\alpha}[1 + \lambda Z(x,1) + (x-1)\mu Z(x-1,1) + (L+1-x)\mu Z(x,0) + s\nu Z(x,0)], \quad 0 < x < L, \tag{23}$$

$$Z(L,0) = \frac{1}{\alpha}[1 + \lambda Z(L,1) + (L-1)\mu Z(L-1,1) + \mu Z(L,0) + s\nu Z(L,0)], \tag{24}$$

$$Z(L+1,0) = \frac{1}{\alpha}[1 + \lambda Z(L+1,1) + L\mu Z(L,1) + s\nu Z(L+1,0)], \tag{25}$$

$$Z(x,y) = \frac{1}{\alpha}[1 + \lambda Z(x,y+1) + (x-1)\mu Z(x-1,y+1) + (L+1-x)\mu Z(x,y) + y\nu Z(x,y-1) + (s-y)\nu Z(x,y)], \tag{26}$$

$$0 < x < L,\ 0 < y < s, \tag{27}$$

$$Z(L,y) = \frac{1}{\alpha}[1 + \lambda Z(L,y+1) + (L-1)\mu Z(L-1,y+1) + \mu Z(L,y) + y\nu Z(L,y-1) + (s-y)\nu Z(L,y)], \tag{28}$$

$$0 < y < s, \tag{29}$$

$$Z(L+1,y) = \frac{1}{\alpha}[1 + \lambda Z(L+1,y+1) + L\mu Z(L,y+1) + y\nu Z(L+1,y-1) + (s-y)\nu Z(L+1,y)], \quad 0 < y < s, \tag{30}$$

$$Z(x,s) = \frac{1}{\alpha}[1 + \lambda Z(x+1,s) + \mu(L+1)Z(x,s) + s\nu Z(x,s-1)], \quad 0 < x < L, \tag{31}$$

$$Z(L,s) = \frac{1}{\alpha}[1 + \lambda q Z(L+1,s) + \lambda(1-q)Z(L,s) + \mu(L+1)Z(L,s) + s\nu Z(L,s-1)], \tag{32}$$

$$Z(L+1,s) = \frac{1}{\alpha}[1 + \lambda Z(L+1,s) + \mu(L+1)Z(L+1,s) + s\nu Z(L+1,s-1)]. \tag{33}$$

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

Each of Eqs. (23)–(33) can be viewed as the expected time until the next event occurs, plus the expected remaining time from that point onward.

The expected sojourn time (for C_0 say) is then given as follows:

$$V_0(x, y, [L, q]^\infty) = \nu^{-1}, \qquad y < s, \tag{34}$$

and

$$V_0(x, s, [L, q]^\infty) = Z(x + 1, s) + \nu^{-1}. \tag{35}$$

5. A DYNAMIC LEARNING SCHEME

Consider a dynamic learning scheme for our system as defined earlier in which customers base their joining decisions on data collected by the system up to the point of decision. Under this scheme, each customer on arrival connects to a public channel if one is available; otherwise, it follows the decision rule

$$\text{join } Q_R \text{ with probability } S_\epsilon(\theta^{-1} - \nu^{-1} - \hat{V}_t(X_t, s)), \tag{36}$$

where ϵ is a small positive parameter, S_ϵ is an increasing function, with $S_\epsilon(x) = 0$ for $x \leq -\epsilon$ and $S_\epsilon(x) = 1$ for $x \geq \epsilon$; also, $\hat{V}_t(x, s)$ is the empirical average milling time of the customers which have been assigned to a channel by time t, but which entered when there were x milling customers and s busy channels.

We assume that the buffer is finite and that a fraction of the arrivals are *uncontrolled*; that is to say, a proportion $p \in (0,1)$ of the arrival stream attempt to join Q_R irrespective of the routing policy. These assumptions ensure that learning never stops in any particular state of the system and are an aid to establishing certain positive recurrence properties of the system (see [3]). The Nash equilibrium being considered is the one which takes into account these additional assumptions (noting that our previous results hold with only minor adjustments).

Let T_n be the nth time at which a customer exits the retrial pool. Denote the state of the *learning process* at time T_n to be $\xi_n \in \mathbb{R}^{2M+2}$ which is defined in the following way. Let the first component of the vector be the state of Q_R at time T_n [i.e., $N(T_n) = (x_n, l_n)$ say]. Let the next x_n components consist of the number of milling customers which each customer present in the retrial pool found on arrival to the system; assign the value 0 to components $x_n + 3, \ldots, M + 2$ of the vector. Similarly, let entries $M + 3$ to $M + x_n + 2$ consist of the time spent in the retrial pool so far by the customers still present there; again, assign the value 0 to components $M + x_n + 3, \ldots, 2M + 2$ of the vector.

With this choice of learning state, coupled with the analysis found in Section 5 of Altman and Shimkin [3], the following theorem now follows.

THEOREM 3: *Let $\hat{u}_t$ denote the decision rule implied by* (36). *Then, $\hat{u}_t$ converges* (*almost surely*) *to a decision rule u^ϵ, which is a threshold-type rule for ϵ sufficiently small. Furthermore, u^ϵ converges to the decision rule implied by* $[g^*]$ *as* $\epsilon \to 0$.

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

6. CONCLUSIONS

We considered the structure of the unique symmetric Nash equilibrium routing policy for customers deciding whether to enter a multiple channel retrial system or a private channel.

A series of stochastic dominance results between various stochastic processes of interest were established via stochastic coupling and, in particular, by exploiting the technique of forward induction. It is expected that forward induction could be exploited to analyze even more complicated models.

A solution procedure for calculating the symmetric Nash equilibrium was presented, and we went on to show that it is capable of predicting long-term behavior under an intuitively appealing dynamic learning rule.

An interesting problem for future research would be to establish the existence and characterize the structure of a socially optimal policy.

Acknowledgments

The author is grateful to Dr. E. J. Collins at the University of Bristol, United Kingdom, for his valuable comments on the material covered in this paper.

The author was supported by EPSRC Postgraduate Studentship Award No. 95006181 while at the School of Mathematics, University of Bristol, United Kingdom.

References

1. Altman, E. (1994). A Markov game approach for optimal routing into a queueing network. Technical Report, INRIA, Sophia-Antipolis Cedex, France.
2. Altman, E. & Shimkin, N. (1993). Worst-case and Nash routing policies in parallel queues with uncertain service allocations. Technical Report, INRIA, Sophia-Antipolis Cedex, France.
3. Altman, E. & Shimkin, N. (1996). Individual equilibrium and learning in a processor sharing system. Technical Report, INRIA, Sophia-Antipolis Cedex, France.
4. Altman, E. & Shimkin, N. (1998). Individual equilibrium and learning in processor sharing systems. *Operations Research* 46(6): 776–784.
5. Ben-Shabar, I., Orda, A., & Shimkin, N. (1998). Dynamic service sharing with heterogeneous preferences. Technical Report, Electrical Engineering Department, Technion, Israel.
6. Brooms, A.C. (1998). Decentralized Nash equilibrium routing policies in various queueing systems. [Unpublished] Ph.D. thesis, School of Mathematics, University of Bristol, United Kingdom.
7. Buche, R. & Kushner, H.J. (1998). Stochastic approximation and user adaptation in a competitive resource sharing system. In *Proceedings of the 37th IEEE conference on Decision and Control.*
8. Cohen, J.W. (1957). Basic problems of telephone traffic theory and the influence of repeated calls. *Philips Telecommunication Review* 18: 49–100.
9. Ephremides, A., Varaiya, P., & Walrand, J. (1980). A simple dynamic routing problem. *IEEE Transactions on Automatic Control* AC-25: 690–693.
10. Falin, G. (1986). On the waiting-time process in a single-line queue with repeated calls. *Journal of Applied Probability* 23: 185–192.
11. Falin, G. & Fricker, C. (1991). On the virtual waiting time in an M/G/1 retrial queue. *Journal of Applied Probability* 28: 446–460.
12. Falin, G. & Templeton, J. (1997). *Retrial queues.* London: Chapman & Hall.
13. Greenberg, S.G. & Wolff, R.W. (1987). An upper bound on the performance of queues with returning customers. *Journal of Applied Probability* 24: 466–475.
14. Hanschke, T. (1987). Explicit formulas for the characteristics of the M/M/2/2 queue with repeated attempts. *Journal of Applied Probability* 24: 486–494.

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

15. Keilson, J., Cozzolino, J., & Young, H. (1968). A service system with unfilled requests repeated. *Operations Research* 16: 1126–1137.
16. Liang, H. & Kulkarni, V.G. (1993). Stability condition for a single-server retrial queue. *Advances in Applied Probability* 25: 690–701.
17. Lippman, S. (1975). Applying a new device in the optimization of exponential queuing systems. *Operations Research* 23: 687–709.
18. Tran-Gia, P. & Mandjes, M. (1997). Modeling of customer retrial phenomenon in cellular mobile networks. *Journal on Selected Areas in Communications* 15: 1406–1414.

APPENDIX

Crucial to establishing that the optimal decision rule of customer C_k, say, against some policy π^{-k} of the other customers has a *finite* threshold for $y = s$, is to show that $V_k(x,s,\pi)$ tends to infinity as $x \to \infty$. We address this issue in the following lemma.

LEMMA 7: *Suppose* $M = \infty$. *Then*,

$$V_k(x,s,\pi) \to \infty \qquad \text{as } x \to \infty.$$

PROOF OF LEMMA 7: Consider customer C_0 without loss of generality. Define $H(x,s)$ to be the expected remaining sojourn time of a customer which is currently in the retrial pool with $x-1$ other retrialing customers, with all s channels being occupied, and arriving customers following the policy π.

First observe that

$$H(x,s) = \frac{1}{\lambda + s\nu} + \frac{\lambda}{\lambda + s\nu}\left[\pi_k(x,s)H(x+1,s) + (1-\pi_k(x,s))H(x,s)\right] + \frac{s\nu}{\lambda + s\nu}H(x,s-1) \tag{37}$$

for some k. However,

$$H(x+1,s) = V_0(x,s,\pi) \geq V_0(x-1,s,\pi) = H(x,s);$$

hence, the term in square brackets is bounded below by

$$\pi_k(x,s)H(x,s) + (1-\pi_k(x,s))H(x,s) = H(x,s).$$

Therefore, a lower bound for (37) is

$$\frac{1}{\lambda + s\nu} + \frac{\lambda}{\lambda + s\nu}H(x,s) + \frac{s\nu}{\lambda + s\nu}H(x,s-1). \tag{38}$$

It follows therefore that

$$H(x,s) \geq \frac{1}{s\nu} + H(x,s-1). \tag{39}$$

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

Also,

$$H(x,s-1) \geq \frac{1}{\lambda+(s-1)\nu+x\mu} + \frac{(x-1)\mu}{\lambda+x\mu+(s-1)\nu} H(x-1,s)$$
$$\geq \frac{1}{\lambda+s\nu+x\mu} + \frac{(x-1)\mu}{\lambda+s\nu+x\mu} H(x-1,s). \quad \textbf{(40)}$$

Combining relations (39) and (40), we have that

$$H(x,s) \geq \frac{1}{s\nu} + \frac{1}{\lambda+s\nu+x\mu} + \frac{(x-1)\mu}{\lambda+s\nu+x\mu} H(x-1,s),$$

which is bounded below by

$$\frac{1}{s\nu} + \frac{(x-1)\mu}{\lambda+s\nu+x\mu} H(x-1,s).$$

By iterating further, we can show that a lower bound for the above quantity is

$$\frac{1}{s\nu} + \frac{1}{s\nu}\left[\frac{(x-1)\mu}{\lambda+s\nu+x\mu} + \frac{(x-1)(x-2)\mu^2}{(\lambda+s\nu+x\mu)^2} + \frac{(x-1)(x-2)(x-3)\mu^3}{(\lambda+s\nu+x\mu)^3} \right.$$
$$\left. + \cdots + \frac{(x-1)(x-2)\cdots 3\cdot 2\mu^{x-2}}{(\lambda+s\nu+x\mu)^{x-2}} \right]$$
$$+ \frac{(x-1)(x-2)(x-3)\cdots 2\cdot 1\mu^{x-1}}{(\lambda+s\nu+x\mu)^{x-1}} H(1,s). \quad \textbf{(41)}$$

Observing that the first and last terms of (41) are nonnegative, we have that

$$H(x,s) \geq \frac{1}{s\nu} \sum_{j=1}^{x-2} I_j(x),$$

where

$$I_j(x) = \prod_{i=1}^{j} \frac{(x-i)\mu}{\lambda+s\nu+x\mu}.$$

Observe that

$$I_j(x) \to 1 \qquad \text{as } x \to \infty. \quad \textbf{(42)}$$

Also, for fixed x, $I_j(x)$ is decreasing in $j \in \mathbb{N}$ since

$$I_j(x) = \left[\frac{(x-j)\mu}{\lambda+s\nu+x\mu} \right] I_{j-1}(x),$$

noting that the term in square brackets is less than 1.

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press

Now select any $C > 0$ and $\delta \in (0,1)$. Choose any $m \in \mathbb{N}$ such that $m > C\delta^{-1}$. By (42), there exists an N'_m such that

$$x \geq N'_m \Rightarrow I_m(x) > \delta.$$

Set $N_m := \max\{m, N'_m\}$. Then, for $x \geq N_m + 2$, we have that

$$\sum_{j=1}^{x-2} I_j(x) \geq \sum_{j=1}^{m} I_j(x) > m\delta > C,$$

as required to show that $H(x,s) \to \infty$ as $x \to \infty$. ■

https://doi.org/10.1017/S0269964800141026 Published online by Cambridge University Press