Optimal Scheduling Policy of Queues with Impatient Customers *

Zheng-Xue Zhao† Shivendra S. Panwar†
Don Towsley‡

†*Polytechnic University, Brooklyn, New York 11201*
‡*University of Massachusetts, Amherst, Massachusetts 01003*

September 3, 1993

Abstract

The problem of call admission in telecommunications is equivalent to the problem of scheduling impatient customers in a queueing system. Given the distribution of the customer deadlines, a scheduling policy decides the customer service order and also which customer(s) to reject. We formulate this problem as a discrete time Markov decision process to maximize the average reward for serving a customer before its deadline. Under the assumptions that arrivals are described by a Bernoulli process, service times are arbitrarily distributed, and the deadline cumulative distribution is concave, we show that the last-in first-out policy that rejects customers whose waiting times exceed a threshold (LIFO-TO) is an optimal stationary policy. When buffer occupancy is the only information available for decision making, the optimal policy turns out to be a pushout LIFO policy (LIFO-PO) where the oldest customers are pushed out when the buffer size exceeds a threshold. This latter result is established under the additional assumption that service times are geometrically distributed. The extension of our results to the continuous time model is also discussed.

Key Words & Phrases: Impatient customers, Scheduling policy, Markov decision processes, Stochastic dominance, Call processing.

1 Introduction

The problem of call admission in telecommunications can considered equivalent to scheduling impatient customers in a queueing system. In an overloaded call admission system, if some

*This work was supported in part by the National Science Foundation under grant NCR-8909719, and by the New York State Science and Technology Foundation and its Center for Advanced Technology in Telecommunications (CATT) at Polytechnic University.

people start dialing before a delayed dial tone is heard, then the system will not receive all the digits dialed. However, the call is still processed and an unsuccessful call results. Another application of this kind of queueing system is the transmission of time-constrained messages over a communication channel. These messages have to reach their receivers within a certain time interval of their transmission, or they are useless to the receivers and considered lost. Thus the system performance is significantly affected by the behavior of impatient customers who should be served before their respective deadlines. Two possible scenarios are often encountered in this kind of system. The first is that the server of the queue is aware of each customer's deadline. The customers whose delay times exceed their deadlines are discarded without service. In the second scenario, the server is only aware of the deadline distribution of the customers. Therefore, some server work is useless because of the expiration of customers' deadlines.

This paper is devoted to searching the optimal policies when only the deadline cumulative distribution is known. Without knowing the deadline of every specific customer, the control action is to decide, at appropriate decision instants, which customer to serve and which customer(s) to reject. The rejection is necessary since customers whose deadlines have expired do not leave the queue automatically. Therefore, a customer could be either served in an order decided by a service discipline or discarded by a rejection scheme. From now on, we use the term *queueing policy* to represent the combination of the service discipline and rejection scheme in a queue. The following notation is used for some specific queueing policies in this paper:

(*i*) FIFO(or LIFO)-BL: first-in first-out (or last-in first-out) service discipline; a customer arriving to see a "full" buffer leaves immediately (blocked).

(*ii*) FIFO(or LIFO)-PO: first-in first-out (or last-in first-out) service discipline; a customer arriving to see a "full" buffer pushes out the "oldest" customer (the one with the longest waiting time) in the buffer and joins the queue.

(*iii*) FIFO(or LIFO)-TO: first-in first-out (or last-in first-out) service discipline; every arriving customer joins the buffer but will leave at a critical time after its arrival if it is still in the buffer at that time (time-out).

More precisely, the above notation is used for queueing policy classes. Those classes consist of the queueing policies with "full" buffer size (for BL and PO schemes) or critical time (for TO schemes) varying with the state of the queue.

Throughout the years, several papers have considered the problem of serving customers with unknown deadlines. Optimal service disciplines were investigated in [1, 2, 3] in which no rejection scheme is adopted. [4] discussed the optimal control problem for a non-preemptive M/M/1/k overloaded queue under FIFO-BL. It is proved that a fixed threshold type rejection decision is optimal for a BL scheme. Other work [1, 5] on this issue focuses on the performance evaluation of various queueing policies.

This paper is organized as follows. Section 2 contains a discrete time queueing model and the problem formulation. In Section 3, the discounted total reward problem is studied to derive an optimal stationary LIFO-TO policy. Next in Section 4, the optimality of this LIFO-TO policy is proved for the good throughput (goodput) problem. Finally, we determine the optimal policy under a reduced information structure in Section 5. The previous results for the discrete time model are also extended to continuous time.

2 The Model

We consider an infinite capacity queue with a slotted single server. Let $a_1 < a_2 < \cdots$ denote the arrival times of customers to this system where the i-th customer arrives in slot $a_i \in I\!N$, $1 \leq i$. We assume that arrivals are described by a Bernoulli process with p_a denoting the probability of an arrival in a slot. Associated with the customer i is a random relative deadline d_i, the time period in which it should begin service after a_i, $1 \leq i$. We assume that $\{d_i\}_{i=1}^{\infty}$ is an independent and identically distributed (i.i.d.) sequence random variables (r.v.'s) with cumulative distribution $F_d(k) = \Pr(d_i \leq k)$, $k = 1, 2, \cdots$. Unless noted otherwise, $F_d(k)$ is a concave function, i.e., $F_d(k+1) - F_d(k) \leq F_d(k) - F_d(k-1)$ for all

$k \geq 1$. Let $\{b_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. random variables that denote the customer service times in slots. In particular, the i-th customer that is scheduled into service is assigned b_i slots of service. Let $f_b(k) = \Pr[b_i = k]$, $k = 1, 2, \cdots$. Last, $\tilde{S} = (\{a_i\}_{i=1}^{\infty}, \{b_i\}_{i=1}^{\infty})$ is referred to as an input sample path.

We are interested in policies which decides whether or not a customer will receive service and, the order in which customers that receive service will be served. Customers that are chosen not to receive service are removed from the queue. We assume that these policies have available to them a complete history of the system including the amount of time that each customer, currently in the queue, has spent in the queue. Last, we restrict ourselves to non-preemptive and non-idling policies. Let Σ denote the class of such policies.

We are interested in the following performance metric, $N_\pi(t)$, the number of customers that begin service prior to their deadlines under $\pi \in \Sigma$ by the t-th slot, $t \geq 1$. Define the system goodput to be

$$G_\pi = \liminf_{t \to \infty} \frac{E[N_\pi(t)]}{t}, \tag{1}$$

Our objective is to determine the policy $\pi \in \Sigma$ that maximizes G_π.

Let W_i^π denote the waiting time of a served customer i under policy π, i.e. the time between its arrival and the beginning of its service. We adopt the convention that $W_i^\pi = \infty$ whenever i is rejected. We have

$$\begin{aligned} E[N_\pi(t)|\tilde{S}] &= \mathrm{E}\{\sum_{i \in \aleph_t(\tilde{S})} \Pr(W_i^\pi \leq d_i)\} \\ &= \mathrm{E}\{\sum_{i \in \aleph_t(\tilde{S})} [1 - F_d(W_i^\pi)]\}, \end{aligned} \tag{2}$$

where $\aleph_t(\tilde{S})$ is the set of customers that depart from the buffer by t under policy π, either due to a service completion or as a result of rejection conditioned on the arrival times and service times. The conditioning on $\tilde{S}$ can be removed to obtain $E[N_\pi(t)]$. Relation (2) allows us to use the same sample path to compare the performances of different queueing policies.

Since policies are non-preemptive, the decisions for customer service can be made only when the server is idle. Without loss of generality, we restrict ourselves to policies that reject customers only at service decision moments. At a decision instant, the system can be modelled by $s = (w_1, w_2, \cdots, w_n)$ where n denotes the number of customers in the queue and $w_1 \leq w_2 \leq \cdots \leq w_n$ denote the times they have spent in the queue. If $n = 0$, then $s = \emptyset$. The n customers are labeled as $c_1, c_2, \cdots, c_n$ where c_i has been in the queue for w_i units of time. Let $\mathcal{S}$ be the set of all feasible states of the queue, $\mathcal{S} = \{(w_1, w_2, \cdots, w_n) | w_1 < w_2 < \cdots < w_n; n \geq 0\}$.

3 Discounted Total Reward Problem

Instead of directly considering the average reward (goodput) metric as defined in (1), we first study the simpler problem of maximizing the *discounted total reward* $V_\pi^\alpha(s_0)$ defined below. Let the state of the system at time 0 be $s_0 \in \mathcal{S}$. At time instant $t_i^\pi \geq 0$, customer i is scheduled with reward $1 - F_d(W_i^\pi)$ under policy π. Define

$$V_\pi^\alpha(s_0) = \mathrm{E}\{\sum_{i=0}^{\infty} \alpha^{t_i^\pi}[1 - F_d(W_i^\pi)] | s_0\}, \tag{3}$$

with $0 < \alpha < 1$ as a discount factor. Let

$$V^\alpha(s_0) = \sup_\pi V_\pi^\alpha(s_0).$$

Policy π^* is said to be α-*optimal* if

$$V_{\pi^*}^\alpha(s_0) = V^\alpha(s_0) \qquad \forall s_0 \in \mathcal{S}.$$

We also introduce $V_\pi^\alpha(s, \tilde{S})$ to be the value function with initial state s_0 conditioned on input sample $\tilde{S}$.

Let A_s denote the set of possible actions that are permitted when the state of the system immediately preceding an action is s. If an action $a \in A_s$ is taken to serve customer c_l in system state s, it incurs a reward $R(a) = 1 - F_d(w_l)$. Otherwise, $R(a) = 0$. This reward is bounded in every state between zero and one. Therefore, (3) is also bounded between zero and $1/(1-\alpha)$ and is well defined. Furthermore from [8], an α-optimal policy can be found within the class of Markov stationary policies.

Now consider an extended state space, $\mathcal{E} = \mathcal{S} \times \{v|v = 0, 1, 2, \cdots\}$, defined at all the slot epochs. Here, v indicates the status of the server, $v = 0$ denoting an idle server and $v > 0$ denoting the number of slots of ongoing service. Thus, there is always a mapping in $\mathcal{S}$ space for any state $e \in \mathcal{E}$ with $v = 0$. With non-preemptive service and the restriction on rejections, the control action is pseudo in a state with $v \neq 0$ and incurs no reward ($R(a) = 0$). Let $P_{ee'}(a)$ denote the probability that the state at an action epoch is e' given that it was e at the previous action epoch and action a was taken. Then the sequence of states in space $\mathcal{E}$ forms a Markov chain under a stationary policy with transition probability $P_{ee'}(a)$. By using Bellman's equation for dynamic optimality [6, 7], we have

$$V^\alpha(e) = \max_a \{R(a) + \alpha \sum_{e'} P_{ee'}(a) V^\alpha(e')\} \qquad \forall e \in \mathcal{E}. \tag{4}$$

Turning our attention back to the state space $\mathcal{S}$ described in the last section, all control actions are active at the service decision moments. Under a stationary policy π, the control action a is given by a function $\pi : \mathcal{S} \to \{0, 1, 2, \cdots\} \times \{S|S \subset I\!N^+\}$. Here $\pi(s)$ is a pair (l, D) where either $l = 0$ if no customer is served, or $1 \leq l \leq n$, if a customer c_l is served. $D \subset \{1, 2, \cdots, n\}$ denotes the subset of customers in the buffer that are rejected, $l \notin D$. Obviously, $\pi(s) = (0, \emptyset)$ whenever $s = \emptyset$. In particular, $l = 0$ when $D = \{1, 2, \cdots, n\}$, i.e., when all the customers are rejected. Also, when $l \neq 0$ and $D = \{1, 2, \cdots, n\} - \{l\}$, no customer stays in the buffer after the action. For reward function, we have $R(a) = 1 - F_d(w_l)$ except $R(a) = 0$ when $l = 0$. Let $\tau_s(a)$ denote the time until the next decision epoch in state s' given that the system state at the preceding decision epoch was s and action a was taken. If $l = 0$, $\tau_s(a)$ is the time until the next arrival resulting in the state $s' = (0)$ and is geometrically distributed with parameter p_a. Otherwise if $l \neq 0$, $\tau_s(a)$ is a service time. In this case, s' consists of two sets of customer waiting times expressed as $s' = (s'_1, s'_2)$. The first, $s'_1 = (w'_1, w'_2, \cdots, w'_m)$, $w'_m > \cdots > w'_2 > w'_1 \geq 0$, contains the waiting times of m arrivals during the service period. If $\tau_s(a) = k$, then $k \geq w'_m + 1 \geq m > 0$. In case of no arrival in $\tau_s(a)$ ($m = 0$), $s'_1 = \emptyset$. The second, s'_2 contains the waiting times of those customers that were waiting in the system at the last decision moment. Their waiting times increase by $\tau_s(a) = k$, i.e., $s'_2 = \{w_i + k | w_i \in s, i \neq l, i \notin D\}$.

From the above discussion, equation (4) can be modified in $\mathcal{S}$ space as

$$V^\alpha(s) = \max_a \{R(a) + \sum_{s'} \sum_k \Pr[\tau_s(a) = k, s']\alpha^k V^\alpha(s')\} \qquad \forall s \in \mathcal{S}. \tag{5}$$

The range of k depends on the transition from s to s' after action a. Specifically, we have

$$\sum_k \Pr[\tau_s(a) = k, s']\alpha^k = \begin{cases} \sum_{k=1}^{\infty} p_a(1-p_a)^{k-1}\alpha^k, \ l = 0; \\ \sum_{k=w'_m+1}^{\infty} f_b(k) p_a^m (1-p_a)^{k-m}\alpha^k, \\ \qquad\qquad l \neq 0 \text{ and } D = \{1, 2, \cdots, n\} - \{l\}; \\ f_b(k) p_a^m (1-p_a)^{k-m}\alpha^k, \quad \text{otherwise,} \end{cases} \tag{6}$$

where $l = 0$ in the first expression gives $s' = (0)$ while the second leads to $s' = (s'_1, \emptyset)$. More generally in the third expression, k is fixed at the value given in s'_2.

3.1 General policy analysis

In this section, we derive structural properties of the optimal policy for our problem based on (3).

Lemma 3.1 *If the deadline cumulative distribution, $F_d(k)$, is concave in k over the non-negative integers, then there exists an optimal policy $\gamma \in \Sigma$ that serves those customers not rejected in LIFO order.*

Proof: Assume that there exists no optimal policy in the LIFO class. Let π be an optimal policy. Choose an input sample where π deviates from the LIFO service order at time k_1. Customers i and j are among the customers available for service and j is the one with the shortest waiting time in the buffer at that time ($j > i$). Under π, i gets served first. Customer j is either scheduled at time $k_2 \geq k_1 + b_i$ where b_i is the service time received by i, or rejected after k_1 under π. We construct a new policy π' as follows:

- At time k_1, π' serves j instead of i;

- π' serves i at time k_2 if π serves j at that time;

- π' rejects i if π rejects j;
- otherwise π' emulates π.

There are two cases.

Case 1 : If customer j is rejected by π after k_1, π' will reject i. We have

$$V_{\pi'}^{\alpha}(s,\tilde{S}) - V_{\pi}^{\alpha}(s,\tilde{S}) = \alpha^{k_1}[F_d(W_i) - F_d(W_j)] \geq 0 \tag{7}$$

since $W_i > W_j$ and $F_d(.)$ is a non-decreasing function. Hence, π' is at least as good as π in this case.

Case 2 : If π chooses to serve j at k_2, π' chooses to serve customer i instead. Then

$$\begin{aligned}
& V_{\pi'}^{\alpha}(s,\tilde{S}) - V_{\pi}^{\alpha}(s,\tilde{S}) \\
= \; & \alpha^{k_1}[1 - F_d(W_j)] + \alpha^{k_2}[1 - F_d(W_i + k_2 - k_1)] \\
& \quad -\alpha^{k_1}[1 - F_d(W_i)] - \alpha^{k_2}[1 - F_d(W_j + k_2 - k_1)] \\
= \; & \alpha^{k_1}\{[F_d(W_i) - F_d(W_j)] + \alpha^{k_2-k_1}[F_d(W_j + k_2 - k_1) - F_d(W_i + k_2 - k_1)]\} \\
> \; & \alpha^{k_2}\{[F_d(W_i) - F_d(W_j)] - [F_d(W_i + k_2 - k_1) - F_d(W_j + k_2 - k_1)]\} \\
\geq \; & 0.
\end{aligned}$$

The last inequality follows from the fact that $F_d(.)$ is a non-decreasing concave function. Therefore, π' is better than π in this case.

This yields a contradiction as the above two cases establishes that policy π' is better than π which was assumed to be optimal. Hence there exists an optimal policy that serves customers in LIFO order. ∎

Lemma 3.2 *For any deadline cumulative distribution $F_d(.)$, an α-optimal policy π^* never rejects a customer while another customer present in the buffer with a longer waiting time gets served later.*

Proof: Consider an arbitrary input sample path $\tilde{S}$. Let an optimal policy π reject customer j but not customer i where $W_j < W_i$. Further assume that i is eventually served under π. We can construct another policy π' which reverses the actions taken on j and i. The argument given in case 1 in the proof of Lemma 3.1 can be applied here to establish

$$V_{\pi'}^{\alpha}(s, \tilde{S}) - V_{\pi}^{\alpha}(s, \tilde{S}) \geq 0.$$

Removal of the conditioning on $\tilde{S}$ yields the desired result. ∎

As has been shown above, an optimal policy would reject all the customers with waiting time longer than W_j whenever customer j is rejected. Since a BL scheme rejects new arrivals, it cannot be optimal. If we use a PO scheme instead of a BL scheme so that the altered policy pushes out the "oldest" customer instead of blocking at every rejection moment, the queueing performance can be improved. Thus, we have the following corollary which is the same as in [11] under the FIFO service order.

Corollary 3.1 *There exists a policy using the PO rejection scheme which is better than the one using the BL scheme regardless of the customer service order. This is true for any customer deadline cumulative distribution.*

The following theorem is a consequence of the above two lemmas.

Theorem 3.1 *For customers with a concave deadline cumulative distribution, a stationary α-optimal policy exists in the class of LIFO-TO policies.*

3.2 The α-optimal policy

The rejection scheme of an optimal policy in the LIFO-TO class can be emulated by a *delayed rejection scheme* which makes use of Lemma 3.2. Let a rejected customer postpone leaving the buffer until either it reaches the server in LIFO order or another customer with a smaller waiting time is rejected. As a result, a delayed rejection always throws away the customer with the smallest waiting time and thus produces an empty queue. By serving the same set of customers all the time, the same performance is achieved under both the original rejection scheme and the delayed scheme.

If we restrict ourselves to the subset $\Gamma \subset \Sigma$ that schedules customers according to the LIFO rule, applies a delayed rejection scheme, and satisfies lemma 3.2, then the description of policy $\pi \in \Gamma$ is simplified. In particular, $\pi : \mathcal{S} \to \{0,1\}$ and is defined as follows. If $\pi(s) = 0$, then all customers are rejected and $s' = (0)$. If $\pi(s) = 1$, then the newest customer is served, no customer is rejected, and s' is the next state with m arrivals. Equations (5) and (6) can be modified as shown below to account for the preceding discussion for all state s.

$$\begin{aligned} & V^\alpha(s) \\ = \; & \max\{[1 - F_d(w_1)] + \sum_{s'}\sum_{k} f_b(k)p_a^m(1-p_a)^{k-m}\alpha^k V^\alpha(s'), \sum_{k=1}^{\infty} p_a(1-p_a)^{k-1}\alpha^k V^\alpha(0)\} \\ = \; & \begin{cases} \max\{[1 - F_d(w_1)] + \sum_{s'}\sum_{k=w'_m+1}^{\infty} f_b(k)p_a^m(1-p_a)^{k-m}\alpha^k V^\alpha(s'), V^\alpha(\emptyset)\}, & s = (w_1); \\ \max\{[1 - F_d(w_1)] + \sum_{s'} f_b(k)p_a^m(1-p_a)^{k-m}\alpha^k V^\alpha(s'), V^\alpha(\emptyset)\}, & \text{otherwise.} \end{cases} \end{aligned} \quad (8)$$

The optimal policy from Γ exhibits the following properties.

Lemma 3.3 *If $\pi \in \Gamma$ is an α-optimal policy, then $\pi(w_1) = 0$ implies $\pi(w_1 + x) = 0$ for all $x \geq 0$.*

Proof: Assume that π is an optimal policy and $\pi(s) = 0$ for $s = (w_1)$. This coupled with (8) implies that $[1 - F_d(w_1)] + \sum_{s'}\sum_{k=w'_m+1}^{\infty} f_b(k)p_a^m(1-p_a)^{k-m}\alpha^k V^\alpha(s') \leq V^\alpha(\emptyset)$. Consider state $s_x = (w_1 + x)$,

$$V^\alpha(s_x) = \max\{[1 - F_d(w_1 + x)] + \sum_{s'}\sum_{k=w'_m+1}^{\infty} f_b(k)p_a^m(1-p_a)^{k-m}\alpha^k V^\alpha(s'), V^\alpha(\emptyset)\}.$$

Since $[1 - F_d(w_1 + x)] \leq [1 - F_d(w_1)]$ and $\pi(s) = 0$, we have

$$[1 - F_d(w_1 + x)] + \sum_{s'}\sum_{k=w'_m+1}^{\infty} f_b(k)p_a^m(1-p_a)^{k-m}\alpha^k V^\alpha(s') \leq V^\alpha(\emptyset).$$

Hence $\pi(s_x) = 0$. ∎

Lemma 3.4 *If $\pi \in \Gamma$ is an α-optimal policy, then*

$$\pi(w_1, w_2, \cdots, w_n) = \pi(w_1), \qquad \forall\, (w_1, w_2, \cdots, w_n) \in \mathcal{S}.$$

Proof: There are two cases according to whether $\pi(w_1) = 1$ or 0. We begin with $\pi(w_1) = 1$. The proof is by contradiction. Assume that $\pi(w_1, \cdots, w_n) = 0$. This implies that $\pi(w_i, \cdots, w_n) = 0$, $i = 2, \cdots, n$ as a consequence of Lemma 3.2, i.e., all of the customers in the queue are removed. The system behaves as if all but the first customer are missing so that $\pi(w_1) = \pi(w_1, \cdots, w_n) = 0$ which contradicts our original statement.

The second case, $\pi(w_1) = 0$ is established using an induction argument on n, the number of customers in the system.

Basis Step: Consider $n = 2$. We have $V^\alpha(w_1, w_2) = V^\alpha(w_1) = V^\alpha(\emptyset)$. The first equality is a consequence of Lemma 3.3 and the fact that $w_1 < w_2 + x$, $x \geq 0$. The second equality is due to the definition of $\pi(w_1) = 0$. Hence, we conclude that $\pi(w_1, w_2) = 0$.

Inductive Step: Assume that the hypothesis is true for $n = 2, 3, \cdots, l$. We establish its validity for $n = l+1$. If $\pi(w_1) = 0$, then $\pi(\tilde{w}_1, \cdots, \tilde{w}_l) = 0$, if $w_1 \leq \tilde{w}_1$ and $(\tilde{w}_1, \cdots, \tilde{w}_l) \in \mathcal{S}$. This is a consequence of Lemma 3.3 and the assumption of hypothesis here for $n = l$. Therefore, $V^\alpha(w_1, \cdots, w_{l+1}) = V^\alpha(w_1) = V^\alpha(\emptyset)$. The first equality is because none of the older l customers will ever receive service as shown earlier and the second equality is a consequence of $\pi(w_1) = 0$. Hence we conclude that $\pi(w_1, \cdots, w_{l+1}) = 0$.

This completes the proof of the lemma. ∎

Lemma 3.4 claims that an α-optimal policy can make a decision based on w_1, the waiting time of the first customer in the queue. From Lemma 3.3, a critical time threshold exists for decision making. Under the LIFO service order, this threshold can extended by releasing the delayed scheme. That is $\pi(s) = 0$ when $w_1 > K$ for a threshold K. Consequently, the above two lemmas imply the following.

Theorem 3.2 *If the deadline cumulative distribution is concave, then the α-optimal policy is a LIFO-TO policy with a time threshold that is independent of the system state.*

4 Average Reward Problem – Goodput

Consider the problem of maximizing system goodput as defined in (1). The results from the previous section can be extended to the problem of maximizing goodput through the application of results in [9]. This extension requires that the model expressed in (4) satisfies the following three assumptions in space $\mathcal{E}$.

$(A1)$ For every state e and discount factor α, the quantity $V^{\alpha}(e)$ is finite.

$(A2)$ There exists a non-negative L such that $V^{\alpha}(e) - V^{\alpha}(\emptyset, 0) \geq -L$ for all e and α.

$(A3)$ There exists non-negative M_e, such that $V^{\alpha}(e) - V^{\alpha}(\emptyset, 0) \leq M_e$ for every e and α. For every e, there exists an action a such that $\sum_{e'} P_{ee'}(a) M_{e'} < \infty$.

Since only a pseudo action is taken for the states with $v \neq 0$, no reward is incurred until a latter state with an idle server $(v = 0)$ is reached. Therefore, the above assumptions can be verified if all the states $s \in \mathcal{S}$ satisfy them. Clearly $(A1)$ holds as we have already shown that $V^{\alpha}(s) \leq 1/(1-\alpha)$. Assumption $(A3)$ holds for similar reasons as well. Last, Assumption $(A2)$ follows from relation (8).

The Theorem in [9] can now be applied to yield the following result.

Theorem 4.1 *There exists an optimal stationary policy which maximizes the goodput defined in (1). If the customers' deadlines are concave distributed, this policy is the LIFO-TO policy with a fixed time threshold that is independent of the state of the system.*

5 Extensions

In this section, we first find the optimal rejection scheme when only the buffer occupancy is known to the server. We also consider the extension to a continuous time queueing system.

5.1 Reduced information structure

Now let us assume that just the buffer occupancy is given to the scheduler rather than customer waiting times. When the customer service time is *geometrically* distributed with a probability p_s of terminating in the next slot, we can extend the LIFO-TO results above to LIFO-PO policies.

Since we can compare the customer waiting times at any time instant by their arrival order, Lemma 3.1 and Lemma 3.2 still apply. The TO rejection scheme cannot be implemented because the customer waiting times are not given. Again an arbitrary policy can be emulated by a delayed rejection scheme. The rejected customers are tagged and leave the buffer only at the moment when a customer in the first waiting room is rejected. We denote the class of policies that schedules customers according to the LIFO rule, applies a delayed rejection scheme, satisfies Lemma 3.2, and uses only buffer occupancy information by Γ'.

Associated with customer i, $i = 1, 2, \cdots$ is a *push-up index* η_i. At the time that customer i arrives, $\eta_i \leftarrow 0$ if the server is idle. Otherwise, $\eta_i \leftarrow 1$. Index η_i is then updated whenever a new customer enters the system while customer i is present. If the number of younger customers including the new arrival exceeds η_i, then $\eta_i \leftarrow \eta_i + 1$; otherwise it remains unchanged. Let $W_i(\eta)$ denote the waiting time of customer i with push-up index η, from the end of its arriving slot to the time it reaches the head of queue for scheduling *in a system that does not allow rejections.* Consider a policy $\pi \in \Gamma'$ that rejects customers. Observe that if customer i receives service under π and its push-up index is η at the time of service, then its waiting time is statistically identical to $W_i(\eta)$, the waiting time in the no reject system. Furthermore, if i is rejected *as the youngest customer present in the queue*, then again its waiting time is statistically identical to $W_i(\eta)$. Last, due to the memoryless properties of the Bernoulli process, we have

$$\Pr[W_i(\eta) \leq k] = \Pr[W_j(\eta) \leq k] \qquad \forall k \geq 0, i \neq j. \tag{9}$$

This results from the fact that i and j do not simultaneously stay in the system with the same push-up index. Therefore, $W_i(\eta)$ and $W_j(\eta)$ are independent and identically distributed. We then are able to omit the subscript i and use $W(\eta)$ instead.

The state of the system at a decision epoch under a policy $\pi \in \Gamma'$ becomes now $s = (\eta_1, \eta_2, \cdots, \eta_n)$ where n denotes the number of customers in the queue and η_i denotes the push-up index of the i-th *youngest* customer, $\eta_1 < \cdots < \eta_n$. Let $\mathcal{S}'$ be the set of all feasible states of the queue, $\mathcal{S}' = \{(\eta_1, \eta_2, \cdots, \eta_n) | 0 \le \eta_1 < \eta_2 < \cdots < \eta_n; n \ge 0\}$. A policy $\pi \in \Gamma'$ is defined as follows, $\pi : \mathcal{S}' \to \{0, 1\}$. If $\pi(s) = 0$, then all customers are rejected. If $\pi(s) = 1$, then the youngest customer is served and no customer is rejected. With m new arrivals in state $s' = (s'_1, s'_2)$, $s'_1 = (\eta'_1, \cdots, \eta'_m)$, $s'_2 = \{\eta_i + \Delta\eta | i \ne 1, \Delta\eta \ge 1\}$ when $\pi(s) = 1$ and $s'_2 = \emptyset$ otherwise, equation (8) becomes

$$\begin{aligned} V^\alpha(s) &= \max\{E[1 - F_d(W(\eta)) | \eta = \eta_1] \\ &\quad + \sum_{s'} \sum_{k=\eta'+1}^{\infty} p_s(1-p_s)^{k-1} p_a^m (1-p_a)^{k-m} \alpha^k V^\alpha(s'), V^\alpha(\emptyset)\}, \quad \forall s, \end{aligned} \tag{10}$$

where $\eta' = \max(m, \eta'_m, \Delta\eta - 1)$

Using the standard notation from [10], we say that a random variable X is *stochastically larger* than another random variable Y, written $Y \le_{st} X$, if

$$E[f(Y)] \le E[f(X)], \qquad \forall \text{ increasing } f$$

We have the following stochastic property for $W(\eta)$.

Lemma 5.1 *With geometrically distributed customer service times, $W(\eta)$ is a stochastically increasing function of η, i.e.,*

$$W(0) \le_{st} W(1) \le_{st} \cdots \le_{st} W(n) \le_{st} \cdots.$$

Proof: The result follows from a standard coupling argument. First, $W(0) \le W(n)$, $n = 1, 2, \cdots$ since $W(0) = 0$. Consider a random customer, say i, whose push-up index is n, $n = 1, 2, \cdots$. There exists some customer, j, that arrives while i is in the queue ($i < j$) and has push-up index $n - 1$. Since j arrives after i, j will depart earlier. Hence $W_j(n-1) < W_i(n)$. Recalling from (9) that $W(n) =_{st} W_i(n)$ and $W(n-1) =_{st} W_j(n-1)$ completes the proof. ∎

Under the LIFO service order with the delayed rejection scheme, the following lemma is similar to Lemma 3.3 and Lemma 3.4 for $(\eta_1, \eta_2, \cdots, \eta_n)$.

Lemma 5.2 *If $\pi \in \Gamma'$ is an α-optimal policy, then* (*i*) $\pi(\eta_1) = 0$ *implies* $\pi(\eta_1 + h) = 0,\ h \geq 0$.

(*ii*) $\pi(\eta_1, \eta_2, \cdots, \eta_n) = \pi(\eta_1)$ *for all* $n \geq 1$ *and state* $(\eta_1, \eta_2, \cdots, \eta_n)$.

Proof (*i*): Assume that π is an optimal policy and $\pi(s) = 0$ where $s = (\eta_1)$. This coupled with (10) implies that $E[1-F_d(W(\eta))|\eta = \eta_1] + \sum_{s'} \sum_{k=\eta'+1}^{\infty} p_s(1-p_s)^{k-1} p_a^m (1-p_a)^{k-m} \alpha^k V^\alpha(s') \leq V^\alpha(\emptyset)$. Consider state $s_h = (\eta_1 + h)$,

$$V^\alpha(s_h) = \max\{E[1 - F_d(W(\eta))|\eta = \eta_1 + h] + \sum_{s'} \sum_{k=\eta'+1}^{\infty} p_s(1-p_s)^{k-1} p_a^m (1-p_a)^{k-m} \alpha^k V^\alpha(s'), V^\alpha(\emptyset)\}.$$

Now, as a consequence of Lemma 5.1, $E[1-F_d(W(\eta))|\eta = \eta_1 + h] \leq E[1-F_d(W(\eta))|\eta = \eta_1]$. This relation coupled with the fact that $\pi(s) = 0$ implies that $E[1-F_d(W(\eta))|\eta = \eta_1 + h] + \sum_{s'} \sum_k p_s(1-p_s)^{k-1} p_a^m (1-p_a)^{k-m} \alpha^k V^\alpha(s') \leq V^\alpha(\emptyset)$. Hence $\pi(s_h) = 0$.

Proof (*ii*): There are two cases according to whether $\pi(\eta_1) = 1$ or 0. We begin with $\pi(\eta_1) = 1$. The proof is by contradiction. Assume that $\pi(\eta_1, \cdots, \eta_n) = 0$. As a consequence, the last $n-1$ customers are rejected (c.f. Lemma 3.2). The system behaves as if all of the customers but the first one do not exist so that $\pi(\eta_1) = \pi(\eta_1, \cdots, \eta_n) = 0$ which contradicts our original statement.

The second case, when $\pi(\eta_1) = 0$, is established using an induction argument on n, the number of customers in the system.

Basis Step: Consider $n = 2$. The proof is by contradiction. Assume $\pi(\eta_1, \eta_2) = 1$, i.e., $V^\alpha(\eta_1, \eta_2) > V^\alpha(\emptyset)$. Now, a consequence of (*i*) above is that the second customer (with push-up index η_2) will always be rejected. Hence, the system behaves as if this second customer does not exist and $V^\alpha(\eta_1) = V^\alpha(\eta_1, \eta_2) > V^\alpha(\emptyset)$ contradicting the assertion that $\pi(\eta_1) = 0$.

Inductive Step: Assume that the hypothesis is true for $n = 2, 3, \cdots, l$. We establish it for $n = l + 1$ by contradiction. Assume that $\pi(\eta_1, \cdots, \eta_{l+1}) = 1$, i.e., $V^\alpha(\eta_1, \cdots, \eta_{l+1}) > V^\alpha(\emptyset)$. It follows from (*i*) above coupled with the induction hypothesis that all of the remaining l customers will be rejected. Hence the system behaves as if none of these customers are present so that $V^\alpha(\eta_1) = V^\alpha(\eta_1, \cdots, \eta_{l+1}) > V^\alpha(\emptyset)$ contradicting the original assertion that

$\pi(\eta_1) = 0$.

This completes the proof of the lemma ∎

The push-up indices are used to estimate the customer waiting times in the above lemma. Thus, a fixed critical time of a LIFO-TO optimal policy becomes a fixed push-out index. This in term implies a fixed buffer size threshold for pushing out customers. We have the following theorem which parallels Theorem 4.1.

Theorem 5.1 *Consider only the buffer occupancy is given for decision making. When the customers' service times are geometrically distributed, the optimal stationary policy is LIFO-PO with a fixed buffer size used as a rejection threshold.*

5.2 The continuous time queueing model

The continuous time queueing model can be derived as a limiting case of the discrete one. Now suppose that the slot length is decreased without limit and p_a, the probability of customer arrival in every slot, is decreased toward zero in such a way that the average number of arrivals in one time unit remains constant at value λ. Based on this, we claim that the customer arrival process becomes a Poisson process with an arrival rate λ while the total number of slots in one time unit approaches infinity. In addition, let customer service times and deadlines become continuous, independent and identically distributed positive random variables. Thus, a simple non-preemptive M/G/1 queue is considered here. When only the buffer occupancy is given for decision making, we are interested in the M/M/1 queueing model.

Our objective remains the same as described in (1) while the queueing state space has changed to an uncountable infinite one. Though Theorem 3.2 is still true for this model by using the same approach, an optimal queueing policy may not exist in the stationary class because of the uncountable state space. Therefore, we make the following conjectures:

(i) If the customers' deadlines are concave distributed, an optimal stationary policy for the M/G/1 queue is in the LIFO-TO policy class with a fixed critical time.

(*ii*) Consider the case when only the buffer occupancy is available for decision making. For the M/M/1 queue, the optimal stationary policy is the LIFO-PO with a fixed buffer size.

References

[1] B. T. Doshi and E. H. Lipper, "Comparisons of Service Disciplines in a Queueing System with Delay Dependent Customer Behavior", *Applied Probability-Computer Science: The Interface*, Vol. II, R. L. Disney and T. J. Ott, Eds. Cambridge, MA: Birkhauser, 1982, *pp.* 269-301.

[2] M. H. Kallmes, D. Towsley and C. G. Cassandras, "Optimality of the Last-In-First-Out (LIFO) Service discipline in Queueing Systems with Real-Time Constraints", *Proceeding of the 28th Conference on Decision and Control*, December 1989, *pp.* 1073-1074.

[3] J. G. Shanthikumar and U. Sumita, "Convex Ordering of Sojourn Times in Single-Server Queues: Extremal Properties of FIFO and LIFO Service Disciplines", *Journal of Applied Probability*, Vol. 24, 1987, *pp.* 737-748.

[4] J.P.C. Blanc, P.R. de Waal, P. Nain and D. Towsley, "Optimal control of admission to a multiserver queue with two arrival streams", *IEEE Transactions on Automatic Control*, Vol. 37, No. 6, June 1992, *pp.* 785-797.

[5] B. T. Doshi and H. Heffes, "Overload Performance of Several Processor Queueing Disciplines for the M/M/1 Queue", *IEEE Transactions on Communications*, Vol. Com-34, No. 6, June 1986, *pp.* 538-546.

[6] D. P. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models*, Prentice-Hall, New Jersey, 1987.

[7] R. Bellman, *Dynamic Programming*, Princeton University Press, New Jersey, 1957.

[8] S. M. Ross, *Introduction to Dynamic Programming*, Academic Press, New York, 1983.

[9] L. I. Sennott, "Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs," *Operation Research*, Vol. 37, *pp.* 626-633, 1989.

[10] S. M. Ross, *Stochastic Processes*, John Wiley & Sons, 1983.

[11] H. Yin and M. G. Hluchyj, "Implication of Dropping Packets from the Front of a Queue", *International Teletraffic Congress Specialists Seminar*, Morristown, 1991.