Dynamic Price Optimization for an M/M/k/N Queue with Several Customer Types

Eugene A. Feinberg* and Fenghsu Yang†

1. INTRODUCTION

We consider an M/M/k/N queue with \( m \) customer types, where \( m \geq 1 \). Customers of type \( j \), \( j = 1, 2, \ldots, m \), arrive at the system according to independent Poisson processes with rates \( \lambda_j \), where \( 0 < \lambda_j < \infty \). There are \( k \) identical servers in the system, where \( 1 \leq k \leq N \). The service times are independent, do not depend on the customer types, and are exponentially distributed with rate \( \mu \), where \( 0 < \mu < \infty \).

When there are \( n \) customers in the system, the total service rate of the system is \( \mu_n \), where \( \mu_n = \mu n \) for \( n = 0, 1, \ldots, k-1 \), and \( \mu_n = k \mu \) for \( n = k, k+1, \ldots, N \). Moreover, there is no preemption for customers. The queue follows the first-in-first-out (FIFO) rule.

We assume that different types of customers have different rewards. Let \( r_j \), \( j = 1, 2, \ldots, m \), be the reward of type \( j \) customers and \( \infty > r_1 > r_2 > \ldots > r_m > 0 \). Moreover, we assume that the system manager is aware of these rewards in the market and no customer in the market is willing to pay the price \( p_n \), \( 0 < p_n < \infty \), at the entrance. If an arriving customer sees \( n \) customers in the system, the system manager posts a price \( p_n \), \( 0 < p_n < \infty \), and enters the system. Meanwhile, the system manager incurs a non-negative holding cost, \( h_n \), as a lump sum to the customer who pays \( p_n \) and enters the system. If an arriving customer sees \( n \) customers in the system and its reward is less than \( p_n \), the customer will leave the system immediately and the system will not collect any price or incur any holding cost. We assume that \( 0 \leq h_0 = h_1 = \ldots = h_{k-1} \leq h_k \leq \ldots \leq h_{N-1} \) and \( h_0 < r_1 \). If there are \( N \) customers in the system, the system manager posts a price \( p_N = r_0 > r_1 \) and all arrivals are rejected. The system manager may reject all arrivals before the system is full by posting the price \( r_0 \). The objective is to maximize the revenue of the system under average reward criterion. We call a policy that maximizes the revenue an optimal policy. In addition to optimal policies, we also consider three more selective policies: canonical, bias optimal, and Blackwell optimal policies.

Two different problems, optimal pricing and optimal admission control, have been studied for queuing systems. In an optimal pricing problem, the system posts prices to customers and the objective is to find a pricing rule, under which the system earns the highest revenue. In the other hand, in an optimal admission control problem, the system decides whether an arriving customer is admitted or not by its customer type, where different customer types have different rewards. If a customer is admitted, the system collects the reward from the customer. The objective is to find an admission rule for each customer type, under which the system earns the highest revenue.

Miller [7] studied optimal admission control problem for an M/M/k/loss queue with \( m \) customer types and without holding costs. He ordered the rewards \( \infty > r_1 > r_2 > \ldots > r_m > 0 \) and formulated his problem as a continuous-time Markov Decision Process (CTMDP). In his study, he introduced the concept of a trunk reservation policy (TRP) and proved the existence of an optimal TRP for his problem. Below we modify the definition of a TRP for this problem:

**Definition 1.1.** We say that a policy \( f \) is a TRP if the price \( f_n \), \( n = 0, 1, \ldots, N \), posted by the system manager when there are \( n \) customers in the system under this policy is a non-decreasing function of \( n \).

Haviv and Puterman [4] studied optimal admission control for an M/M/k/N queue with one customer type and convex increasing holding costs. They proved that there exist at most two optimal control levels for the customers and, if there are two different optimal control levels, the difference between them is one. They also showed that the bias optimal policy is unique, is the policy that selects the larger optimal control level for the customers, and is also Blackwell optimal. Lewis et al. [5] proved that a similar result holds for an M/M/k/N queue with several customer types and without holding costs. Feinberg and Yang [1] studied optimal admission control for an M/M/k/N queue with several types of customer and convex holding costs. They showed that the bias optimal policy for their problem is unique, is Blackwell optimal, and is the optimal TRP with the largest optimal control level for each customer type.

Low [6] studied dynamic pricing for an M/M/k/N queue with several prices \( p_1, p_2, \ldots, p_k \). Each price \( p_i \) is associated with an arrival rate \( \Lambda_i \). He let \( 0 < p_1 < p_2 < \ldots < p_k \) and \( 0 < \Lambda_1 < \Lambda_2 < \ldots < \Lambda_k \). The system manager posts a price \( p_n \in \{p_1, p_2, \ldots, p_k\} \), \( n = 0, 1, \ldots, N - 1 \), when there are \( n \) customers in the system under this policy is a non-decreasing function of \( n \).

---

*Eugene A. Feinberg is with the Department of Applied Mathematics and Statistics, State University of New York at Stony Brook, Stony Brook, NY 11794-3600, USA. Email:Eugene.Feinberg@sunysb.edu
†Fenghsu Yang is with the Department of Applied Mathematics and Statistics, State University of New York at Stony Brook, NY 11794-3600, USA. Email:fyang@ic.sunysb.edu

Copyright is held by author/owner(s).
n customers in the system. In addition, when a customer sees there are n customers in the system and pays the price, the system incurs a holding cost \( c_n \) as a lump sum to the customer. He assumed 0 ≤ \( c_0 = c_1 = \ldots = c_{k-1} \leq c_k \leq \ldots \leq c_{N-1} < p_k \). In addition, Low [6] considered that the system manager rejects arrivals only when the system is full. He proved that, under any optimal policy, the price the system manager posts at state \( n \) is a non-decreasing function of \( n \).

Giloni et al. [2] studied the problem we are considering in this paper. They showed that, when the holding costs do not depend on customer types, the total arrival rates under an optimal policy are non-increasing with respect to the states. When the holding costs depend on customer types, they provided an example showing that the total arrival rates under an optimal policy may be increasing at some states. The results in [2] indicate the existence of an optimal TRP for this problem.

In this paper, we provide descriptions of all stationary optimal policies including policies optimal for more selective criteria: canonical policies, bias optimal policies, and Blackwell optimal policies. Any stationary optimal policy becomes a TRP (Theorems 3.1 and 3.2) after decisions are made at transient states. Such policies are called essential trunk reservation policies (ETRPs) [1]. We show that there is a unique bias optimal policy for this problem. It is also Blackwell optimal and is the optimal TRP that selects the lowest optimal price \( P_{n,} \) for \( n = 0, 1, \ldots, N \), when there are \( n \) customers in the system.

2. PROBLEM FORMULATION

The problem considered in this paper is modelled as a CTMDP with the state space \( S = \{0, 1, \ldots, N\} \), where state \( n \) means there are \( n \) customers in the system. Let \( A(n) \) be the set of actions available at state \( n \) for \( n = 0, 1, \ldots, N \), where an action \( a \in A(n) \) means the system manager posts the price \( a \) to arrivals at state \( n \). Let \( A = \bigcup_{n=0}^{N} A(n) \). Since the system manager is aware of the rewards in the market, we can restrict \( A(n) \), \( n = 0, 1, \ldots, N-1 \), to the finite set \{\( r_0, r_1, \ldots, r_N \)\}. Because the system manager has to reject all arrivals when the system is full, we let \( A(N) = \{r_0\} \).

Note that, when \( a = r_0 \) is selected, the arriving rate \( \lambda_0 = 0 \). Each action \( a = r_j \in A(n) \), \( n = 0, 1, \ldots, N \), defines a transition intensity \( \lambda_n^a = \sum_{j=0}^{\infty} \lambda_j \) and an expected revenue per unit time \( R(n, a) = (a - h_n) \lambda_n^a \).

Since \( S \) and \( A \) are finite, there exists a stationary optimal policy for this problem. Thus, in this paper, we focus on the set of all stationary policies and let \( F \) be this set. For any policy \( f \in F \), we let \( f = (f_0, f_1, \ldots, f_N) \), where \( f_x, \quad n = 0, 1, \ldots, N \), is the action that \( f \) chooses from \( A(n) \) and \( \lambda_n^a \) be the transition intensity defined by action \( f_x \). Moreover, the policy \( f \) defines a vector of expected revenue per unit time \( R(f) \) whose \((n+1)\)-st component is \( R(n, f_x) \).

Furthermore, let \( P^f \) be the steady probability matrix under \( f \in F \). We denote \( v_n^f \) the revenue, given that the process starts at state \( n \) under \( f \), and \( V^f \) the vector whose \((n+1)\)-st element is \( v_n^f \) for \( n = 0, 1, \ldots, N \). Then \( V^f = P^f(f) R(f) \). The objective is to maximize the vector \( V^f \). Note that there is only one recurrent class under any policy from \( F \). Such CTMDPs are called unichain. Since the CTMDP is unichain, all rows of \( P^f \) are equal to the unique steady state probability vector \( (\pi_n^f(f), \pi_1^f(f), \ldots, \pi_N^f(f)) \). Thus, all elements of \( V^f \) are equal and \( v_n^f = v^f = \sum_{s=0}^{N-1} \pi_s^f(f) R(s, f_s) \) for \( n = 0, 1, \ldots, N \). A policy \( f \in F \) is optimal if \( v^f \geq v^d \) for all \( \phi \in F \). Let \( F^* \) be the set of stationary optimal policies.

For a unichain CTMDP, a policy \( f \) is called canonical if and only if there exists a function \( y^f \) such that

\[
 v^f = R(i, f_i) + \sum_{j=0}^{N} q(j|i, f_i) y^f_j 
\]

where \( f^f ) = 0 \text{ for } i = 0, 1, \ldots, N \). For a unichain CTMDP with finite states and action sets, a canonical policy always exists, is optimal, and Howard’s [3] policy iteration algorithm computes it; see [8] for detail.

3. OPTIMAL PRICING POLICIES

In order to compute a policy \( f \in F^* \) for this problem, we consider the following equations for a CTMDP [7]:

\[
 V^f = P^*(f) \times R(f),
\]

\[
 R(f) + Q(f) \times y^f = V^f,
\]

\[
 P^*(f) \times y^f = 0,
\]

where \( y^f \) is the bias vector and \( f \in F \). We denote by \( y_n^f \) the \((n+1)\)-st element of \( y^f \), \( n = 0, 1, \ldots, N \). Let \( y_n H_n = y_n^f - y_n^{f-1} \) for \( n = 0, 1, \ldots, N-1 \). For any \( f \in F \) and \( n = 0, 1, \ldots, N \), we define

\[
 H^f(n, z) = (f_n - h_n - z) \lambda_n^a,
\]

where \( H^f(n, z) = 0 \) if \( f_n = r_0 \). Also, we let \( \nabla y_n^f = 0 \) for any \( f \in F \). Then, according to Miller [7], the following equation is the same as (3):

\[
 H^f(n, \nabla y_n^f) + \mu_n \times \nabla y_n^{f-1} = v^f, \quad n = 0, 1, \ldots, N.
\]

Below is the policy iteration algorithm for this problem:

**Algorithm 1 Policy Iteration Algorithm**

1. Choose a policy \( f \in F \).

2. For policy \( f \), compute \( v^f \) and \( \nabla y_n^f, n = 0, 1, \ldots, N-1 \), by solving linear equations (5).

3. a. If \( H^f(n, \nabla y_n^f) = \max_{a \in A(n)} \{(a - h_n - \nabla y_n^f) \lambda_n^a\} \) for \( n = 0, 1, \ldots, N-1 \), then \( f \) is optimal. Stop.

   b. For all \( n = 0, 1, \ldots, N-1 \) such that \( H^f(n, \nabla y_n^f) < \max_{a \in A(n)} \{(a - h_n - \nabla y_n^f) \lambda_n^a\} \), change \( f_n \) to \( a' \) such that \( (a' - h_n - \nabla y_n^f) \lambda_n^a = \max_{a \in A(n)} \{(a - h_n - \nabla y_n^f) \lambda_n^a\} \). Go to Step 2.

The following example shows that there may exist a canonical policy that is not a TRP.

**Example 3.1.** Consider an M/M/2/5 queue with three customer types, where \( \lambda_1 = \lambda_2 = \mu = 1, \lambda_3 = 2, r_1 = 10, r_2 = 6, r_3 = 4, h_0 = h_1 = 0, h_2 = h_3 = 6, \) and \( h_4 = 8 \). In addition, we let \( r_0 = 20 \) be the price to reject arrivals. Consider a policy \( f = (10, 10, 20, 20, 20) \). By solving (5), we have \( v^f = 8 \). Now consider another policy \( d = (10, 10, 20, 10, 20) \). By solving (5), we have \( v^d = 8 \). Both \( f \) and \( d \) are canonical policies, but only policy \( f \) is a TRP.
The following formula describes an arbitrary canonical policy for Example 3.1:

$$f_n \in \begin{cases} 
\{10, 6, 4\}, & \text{if } n = 0; \\
\{1\}, & \text{if } n = 1; \\
\{20, 10\}, & \text{if } n = 2, 3; \\
\{20\}, & \text{if } n = 4, 5.
\end{cases}$$

Let $n^f = \min(n = 1, \ldots, N : f_n = r_0)$, where $f \in F$. For a policy that admits customers at transient states, Feinberg and Yang [1] introduced the definition of the restriction of a policy and the definition of an essential trunk reservation policy. Below we modify these definitions for this problem.

**Definition 3.1.** We call $f^R \in F$ the restriction of policy $f \in F$ if $f^R = f_n$ for $n = 0, 1, \ldots, n^f$ and $f^R_n = r_0$ for $n = n^f + 1, n^f + 2, \ldots, N$. A policy $f \in F$ is called an essential trunk reservation policy (ETRP) if $f^R \in F$ is a TRP.

Let $C^*$ be the set of all canonical policies, $E^*$ be the set of all optimal ETRPs, and $T^*$ be the set of all optimal TRPs. We need the following condition and definition before we describe our results for $T^*$, $C^*$, $E^*$, and $F^*$.

**Condition 3.1.** (a) $h_0 = h_0 = \ldots = h_{k-1} < h_k < \ldots < h_{N-1}$ or (b) $h_0 = h_1 = \ldots = h_{N-1}$.

**Definition 3.2.** A price $P_n \in \{r_0, r_1, \ldots, r_m\}$ is called an optimal price at state $n$ if there exists a policy $f \in C^*$ such that $f_n = P_n$.

After obtaining a canonical policy $f$ from Algorithm 1, the set of optimal prices, $\theta_n$, at state $n$, $n = 0, 1, \ldots, N$, can be found by the formula:

$$\theta_n = \{r_j, 0 \leq j \leq m : (r_j - h_n - \nabla y_n^j) \sum_{i=0}^{j} \lambda_i = H^f(n, \nabla y_n^j)\}$$

for $n = 0, 1, \ldots, N$. Let $P^\theta_n = \max_{P \in \theta_n} P$ and $P^\theta_n = \min_{P \in \theta_n} P$ for $n = 0, 1, \ldots, N$. The following theorems describe our results for $T^*$, $C^*$, $E^*$, and $F^*$.

**Theorem 3.1.** If Condition 3.1 holds, then: (a) $\emptyset \neq T^* = C^* \subseteq E^* \subseteq F^*$; (b) $P^\theta_n \leq P^\theta_{n+1}$ for $n = 0, 1, \ldots, N - 1$; (c) $f \in T^*$ if and only if $f_n \in \theta_n$ for $n = 0, 1, \ldots, N$.

**Theorem 3.2.** If Condition 3.1 does not hold, then: (a) $\emptyset \neq T^* \subseteq C^* \subseteq E^* \subseteq F^*$; (b) $f \in T^*$ if and only if $f_n \in \theta_n$ for $n = 0, 1, \ldots, N$ and $f_n \leq \theta_{n+1}$ for $n = 0, 1, \ldots, N - 1$.

4. **Bias Optimality**

If there exist two or more stationary optimal policies, bias optimality can be applied to find the policy that maximizes the bias vector among all stationary optimal policies. Below are the definitions of bias optimality and Blackwell optimality.

**Definition 4.1.** A policy $g$ is bias optimal if $g$ is optimal and $y_{g_n}^j \geq y_{f_n}^j$, $n = 0, 1, \ldots, N$, for every policy $f \in F^*$.

**Definition 4.2.** A policy $f$ is Blackwell optimal if there exists a number $\alpha^* > 0$ such that $\phi_{\alpha}^f \geq \phi_{\alpha}^f$, for all $\alpha \in (0, \alpha^*)$ and all $f \in F$, where $\phi_{\alpha}^f$ is the total discounted reward under policy $f$ with discount rate $\alpha$.

The lemmas below state the properties of a bias optimal policy for this problem.

**Lemma 4.1.** If $g$ is a bias optimal policy, then $g \in C^*$.

To compute the bias vector for a policy $f \in F$, Feinberg and Yang [1] transformed (4) into the following equation:

$$y_0^f = \sum_{z=1}^{N} \pi_z^f(\sum_{n=0}^{z-1} \nabla y_n^f).$$

In addition, $y_0^f = y_0^g - \sum_{n=0}^{z-1} \nabla y_n^f$, $n = 1, 2, \ldots, N$. We use (6) to prove the following lemma.

**Lemma 4.2.** If $g$ is a bias optimal policy, then $g_n = P_n^f$ for $n = 0, 1, \ldots, N$.

The following theorem describes the bias optimal policy $g$ for this problem.

**Theorem 4.1.** There exists a unique bias optimal policy $g$ and it is also Blackwell optimal. This policy is the TRP that selects the lowest optimal price at each state. In particular, the bias optimal policy $g$ is a TRP with $g_n = P_n^f$ for $n = 0, 1, \ldots, N$.

From Theorem 4.1, the bias optimal policy $g$ can be found by the following formula after obtaining a canonical policy $f$ from Algorithm 1:

$$g_n = \min_{j=0, 1, \ldots, m} \{r_j : (r_j - h_n - \nabla y_n^j) \sum_{i=0}^{j} \lambda_i = H^f(n, \nabla y_n^j)\}$$

for $n = 0, 1, \ldots, N - 1$.

**Acknowledgement.** This research was partially supported by NSF grant CMMI-0928490.

5. **REFERENCES**


