Pricing and Queueing

Christian Borgs (MSR-NE) Jennifer T. Chayes (MSR-NE)
Sherwin Doroudi (CMU-Tepper) Mor Harchol-Balter (CMU-CS) Kuang Xu (MIT-LIDS)

ABSTRACT

We consider a pricing in a single observable queue, where customers all have the same valuation, V, and the same waiting cost, v. It is known that earning rate is maximized in such a model when state-dependent pricing is used and an admissions threshold is deployed whereby arriving customers may not join the queue if the total number of customers exceeds this threshold. This paper is the first to explicitly derive the optimal threshold. We use our explicit formulation to obtain asymptotic results on how the threshold grows with V.

Keywords

Optimal threshold; state-dependent pricing; observable queue; maximizing revenue

1. INTRODUCTION

We consider the classical setting of Naor (1969) and Chen and Frank (2001), wherein a firm is selling a service to stochastically arriving customers. The customers gain utility from receiving the service, but this utility decreases as the delay experienced by customers increases. The goal of the firm is to price its service and choose an admission policy so as to maximize its earning rate.

Customers are assumed to be identical in the value they obtain from receiving service (their reservation value), V, and their waiting cost per unit time, v. They differ only in their arrival times, which occur according to a Poisson process with rate λ, and their service requirements, which are assumed to be independent identically distributed exponential random variables with rate $\mu = 1$.

We assume that customers can see the current state of the queue (a.k.a. observable queue) and can use this in determining whether they want to join the queue. Given that the length of the queue is observable, and customers experience a greater net value (before paying a price) when the queue is shorter, it will be advantageous for a profit maximizing firm to charge different prices at different queue states. That is, the firm charges $p(n)$ when there are $n \geq 0$ customers in the system (including the queue and the customer being served, if any). When the queue is shorter, customers are willing to pay more to join the queue, so $p(n)$ is higher for small n and lower for large n.

When a customer arrives, she observes the state of the system, n (total number of jobs in the system). It only pays for the customer to join the queue if:

$$V - v \cdot \mathbb{E}[T|n] - p(n) \geq 0 \tag{1}$$

where $\mathbb{E}[T|n]$ is the expected response time given that there are n customers in the system, not including the new arrival. Since the average service time is 1 for all customers, $\mathbb{E}[T|n] = n + 1$. Setting $v = 1$, we can rewrite (1) to find that the customer joins the system if and only if:

$$V - (n+1) - p(n) \geq 0 \tag{2}$$

There is a long history of studying the question of how the firm should set its prices, so as to maximize its earning rate, $\mathcal{R}$. This history can be summarized in the work of Chen and Frank (2001), who find that the earning rate is maximized by charging prices that are as *high as possible* in each state (i.e., set $p(n) = V - (n+1)$) while imposing an appropriate *threshold*, k^*, such that arrivals may not join the queue beyond this state. While this threshold can be computed numerically, none of the prior work explicitly derives a closed form for this threshold.

It may not be obvious why a threshold is a good idea, since, by definition, it means that some customers are turned away (a customer seeing k^* jobs in the system is refused service). However, imposing a threshold forces the queue to stay short, which is exactly the region where prices are highest; thus, if λ is high, a threshold can end up greatly increasing the earning rate, despite forgoing some potential revenue.

The purpose of our paper is to explicitly derive the optimal threshold, k^*, which maximizes the earning rate:

$$\mathcal{R}(k, \lambda, V) = \lambda \sum_{n=0}^{k-1} p(n) \cdot \pi_n^{(k)} \tag{3}$$

where $\pi_n^{(k)}$ represents the limiting probability that the state of the queue is n, given that the queue has a threshold of k, and $p(n) = V - (n+1)$ is the price charged when the state is n. While we consider the case where $\mu = 1$ and $v = 1$, our results easily generalize to other values of μ and v.

2. RESULTS SUMMARY

Our contributions are as follows:

- We find that optimal threshold is given by the formula

$$k^* = \left\lceil (1-\lambda)V + \frac{1}{1-\lambda} - \frac{1}{\ln(\lambda)} \cdot X - 2 \right\rceil$$

where

$$X = W\left(\frac{\ln(\lambda)\cdot\lambda^{(1-\lambda)V+\frac{1}{1-\lambda}}}{1-\lambda}\right)$$

where W is the greater branch of the Lambert W function for $\lambda < 1$, and the lesser branch for $\lambda > 1$. In the remaining case where $\lambda = 1$,

$$k^* = \left\lceil \frac{1}{2}\left(\sqrt{1+8V} - 3\right) \right\rceil$$

The Lambert W (product logarithm) function is a well understood non-elementary function (see Corless et al. (1996) for an overview of the theory and applications of the Lambert W function). Although to our knowledge, this function has not appeared in the analysis of queueing models with pricing, it has been applied in related areas involving queueing. In Libman and Orda (2002), the Lambert W function is used in computing the optimal time to wait before retrying an action under M/M/1-induced delay, and Gupta and Weerawat (2006) use the function to compute the optimal inventory level in a two stage queueing model.

- We find the following strikingly simple approximation for the optimal earning rate

$$\mathcal{R}(k^*, \lambda, V) \approx V - (k^* + 1)$$

- We find the following asymptotic characterizations for k^* as $V \to \infty$:

$$k^* \sim \begin{cases} (1-\lambda)V & \text{if } \lambda > 1 \\ \sqrt{2V} & \text{if } \lambda = 1 \\ \log_\lambda(V) & \text{if } \lambda > 1 \end{cases}$$

and these asymptotic results can be used together with the aforementioned approximation to obtain the following asymptotic characterizations of $\mathcal{R}(k^*)$:

$$\mathcal{R}(k^*) \sim \begin{cases} \lambda V & \text{if } \lambda < 1 \\ V & \text{if } \lambda \geq 1 \end{cases}$$

- Finally, we demonstrate some simple examples where the optimal threshold can produce a many-fold increase in earning rate.

3. THE ANALYSIS

In the interest of space, we only show the analysis for our first result, the derivation of the closed form expression for the optimal threshold, k^*, which minimizes the earning rate.

We start by deriving the earning rate, which, after some lengthy algebra, can be expressed as follows (assuming $\lambda \neq 1$):

$$\begin{aligned}\mathcal{R}(k,\lambda,V) &= \lambda \sum_{n=0}^{k-1} p(n)\cdot \pi_n^{(k)} \\ &= \frac{\lambda}{1-\lambda^{k+1}}\cdot\frac{1}{1-\lambda} \\ &\quad \cdot \left[V(1-\lambda^k)(1-\lambda) + \lambda^k(1+k-k\lambda) - 1\right]\end{aligned} \tag{4}$$

In the case where $\lambda = 1$, it is straightforward to compute:

$$\mathcal{R}(k,1,V) = k\cdot\left(\frac{V}{k+1} - \frac{1}{2}\right) \tag{5}$$

3.1 The forward difference technique

We seek to find the optimal threshold

$$k^* = \arg\max_{k\in\mathbb{Z}_+}\{\mathcal{R}(k,\lambda,V)\}.$$

We know that at least one such optimal threshold exists, see Chen and Frank (2001). We consider the function $\Delta\mathcal{R}(\cdot)$, defined by $\Delta\mathcal{R}(x) = \mathcal{R}(x+1) - \mathcal{R}(x)$, also known as the *forward difference* of $\mathcal{R}(\cdot)$. This function loosely captures the idea of a "discrete derivative" which we can use with the modified first order condition $\Delta\mathcal{R}(x) = 0$. The following theorem, stated without proof, formalizes this approach.

THEOREM 3.1. *The optimal threshold k^* is achieved at $k^* = 0$, or $k^* = \lceil x \rceil$ for some $x \in \mathbb{R}_+$ that satisfies*

$$\mathcal{R}(x+1,\lambda,V) - \mathcal{R}(x,\lambda,V) = 0 \tag{6}$$

We will refer to x as the unrounded optimal threshold. Solving the forward difference equation for x, we have:

$$\begin{aligned}0 &= \mathcal{R}(x,\lambda,V) - \mathcal{R}(x+1,\lambda,V) \\ &= \frac{\lambda}{1-\lambda^{x+1}}\cdot\frac{1}{1-\lambda} \\ &\quad \cdot [V(1-\lambda^x)(1-\lambda) + \lambda^x(1+x-x\lambda) - 1] \\ &\quad - \frac{\lambda}{1-\lambda^{x+2}}\cdot\frac{1}{1-\lambda} \\ &\quad \cdot \left[V(1-\lambda^{x+1})(1-\lambda) + \lambda^{x+1}(1+(x+1)-(x+1)\lambda) - 1\right]\end{aligned}$$

Multiplying both sides by $\frac{1-\lambda}{\lambda}\cdot(1-\lambda^{x+1})(1-\lambda^{x+2})$, and performing some straightforward, if lengthy, algebra, we have:

$$\begin{aligned}0 &= (1-\lambda^{x+2})\cdot[V(1-\lambda^x)(1-\lambda) + \lambda^x(1+x-x\lambda) - 1] \\ &\quad - (1-\lambda^{x+1}) \\ &\quad \cdot \left[V(1-\lambda^{x+1})(1-\lambda) + \lambda^{x+1}(2+x-x\lambda-\lambda) - 1\right] \\ &= -V(1-\lambda)^3\lambda^x + \lambda^x \\ &\quad \cdot \left[(1-2\lambda)(1-\lambda) + x(1-\lambda)^2 + \lambda^{x+2}(1-\lambda)\right]\end{aligned}$$

Dividing both sides by $\lambda^x(1-\lambda)$ we have:

$$0 = -V(1-\lambda)^2 - 1 + 2(1-\lambda) + x(1-\lambda) + \lambda^{x+2}$$

Dividing both sides again by $(1-\lambda)$ and using $\tilde{x} = x + 2$ yields:

$$0 = -V(1-\lambda) - \frac{1}{1-\lambda} + \tilde{x} + \frac{\lambda^{\tilde{x}}}{1-\lambda}$$

We now define:

$$G(\lambda, V) = V(1-\lambda) + \frac{1}{1-\lambda} \tag{7}$$

Rewriting to use $G(\lambda, V)$ we have:

$$G(\lambda, V) - \tilde{x} = \frac{\lambda^{\tilde{x}}}{1-\lambda}$$

Multiplying both sides by $\lambda^{G(\lambda,V)-\tilde{x}}$, we have:

$$(G(\lambda, V) - \tilde{x})\, \lambda^{G(\lambda,V)-\tilde{x}} = \frac{\lambda^{G(\lambda,V)}}{1-\lambda}$$

Expressing $\lambda^{G(\lambda,V)-\tilde{x}}$ as $e^{\ln(\lambda)\cdot((G(\lambda,V)-\tilde{x})}$ and multiplying both sides by $\ln(\lambda)$, we have

$$\ln(\lambda) \cdot (G(\lambda, V) - \tilde{x})\, e^{\ln(\lambda)\cdot(G(\lambda,V)-\tilde{x})} = \frac{\ln(\lambda) \cdot \lambda^{G(\lambda,V)}}{1-\lambda} \tag{8}$$

Our goal is to solve this equation for $\tilde{x}$, and hence $x = \tilde{x} - 2$.

Definition 3.1. *For all $z > -1/e$, the* Lambert W func*tion (also known as the product logarithm, or productlog function) is defined as either one of two real-valued functions (branches) giving the solution to:*

$$W(z)e^{W(z)} = z$$

We refer to the specific branches as W_0 and W_{-1}, with $W_0(z) > W_{-1}(z)$ for all $z > -1/e$.

By definition of the Lambert W function, the solutions to equations of the form $Xe^X = Y$ are $X = W(Y)$. Observing that (8) has this form, and rearranging terms, we have:

$$\tilde{x} = G(\lambda, V) - \frac{1}{\ln(\lambda)} \cdot W\left(\frac{\ln(\lambda) \cdot \lambda^{G(\lambda,V)}}{1-\lambda}\right) \tag{9}$$

Finally, recalling that $\tilde{x} = x + 2$, we have:

$$x = G(\lambda, V) - \frac{1}{\ln(\lambda)} \cdot W\left(\frac{\ln(\lambda) \cdot \lambda^{G(\lambda,V)}}{1-\lambda}\right) - 2 \tag{10}$$

Substituting back (7), we obtain a closed form for the unrounded optimal threshold:

$$x = (1-\lambda)V + \frac{1}{1-\lambda} - \frac{1}{\ln(\lambda)} \cdot X - 2 \tag{11}$$

where

$$X = W\left(\frac{\ln(\lambda) \cdot \lambda^{(1-\lambda)V + \frac{1}{1-\lambda}}}{1-\lambda}\right)$$

Therefore, by the analysis above, we may conclude that x is a solution to $\mathcal{R}(x+1) - \mathcal{R}(x) = 0$. By Theorem 3.1, it follows that the optimal threshold, k^*, satisfies

$$k^* = \left\lceil (1-\lambda)V + \frac{1}{1-\lambda} - \frac{1}{\ln(\lambda)} \cdot X - 2 \right\rceil \tag{12}$$

where

$$X = W\left(\frac{\ln(\lambda) \cdot \lambda^{(1-\lambda)V + \frac{1}{1-\lambda}}}{1-\lambda}\right)$$

3.2 Selecting the correct branch of the Lambert W function

As stated earlier, the Lambert W function has two real-valued branches, $W_0(x)$, and $W_{-1}(x)$. It follows that k^* must be given in terms of the correct branch, which we find depends on λ. Theorem 3.2 unambiguously states k^* in closed form by indicating which branch should be selected. The proof relies on the observation that, when $\lambda < 1$, only the $W_0(x)$ branch yields a positive value for k^* in (12). Similarly, when $\lambda > 1$, only the $W_{-1}(x)$ branch yields a positive value for k^*.

Theorem 3.2. *For all $\lambda \neq 1$,*

$$k^* = \left\lceil (1-\lambda)V + \frac{1}{1-\lambda} - \frac{1}{\ln(\lambda)} \cdot X_i - 2 \right\rceil \tag{13}$$

where

$$X_i = W_i\left(\frac{\ln(\lambda) \cdot \lambda^{(1-\lambda)V + \frac{1}{1-\lambda}}}{1-\lambda}\right)$$

where

$$i = \begin{cases} 0 & \text{if } 0 < \lambda < 1 \\ -1 & \text{if } \lambda > 1 \end{cases}$$

In the remaining case where $\lambda = 1$,

$$k^* = \left\lceil \frac{1}{2}\left(\sqrt{1+8V} - 3\right) \right\rceil$$

Proof. Omitted due to lack of space. □

4. DISCUSSION

While the statement of the problem is very simple and has been considered in multiple prior papers, there was seemingly no interest in looking for a closed-form solution to the optimal cutoff, or in understanding the asymptotic behavior of this cutoff in terms of how it scales with V and λ. It is possible that the algebraic manipulations needed to find these forms were off-putting and obscured the underlying simple result in (13). This simple formula leads to the elegant asymptotics described in Section 2.

There are many more general customer models that we are currently considering, involving multiple classes of customers, with different V and v values.

This work is funded by a Computational Thinking Gift from Microsoft as well as NSF grant CNS-1116282.

References

Chen, H. and Frank, M. (2001). State dependent pricing with a queue. *IIE Transactions*, 33(10):847–860.

Corless, R., Gonnet, G., Hare, D., Jeffrey, D., and Knuth, D. (1996). On the Lambert W function. *Advances in Computational Mathematics*, 5(1):329–359.

Gupta, D. and Weerawat, W. (2006). Supplier-manufacturer coordination in capacitated two-stage supply chains. *European Journal of Operational Research*, 175(1):67–89.

Libman, L. and Orda, A. (2002). Optimal retrial and timeout strategies for accessing network resources. *Networking, IEEE/ACM Transactions on*, 10(4):551–564.

Naor, P. (1969). The regulation of queue size by levying tolls. *Econometrica: journal of the Econometric Society*, pages 15–24.