STRUCTURAL RESULTS ON A BATCH ACCEPTANCE PROBLEM FOR CAPACITATED QUEUES

Eren Başar Çil, E. Lerzan Örmeci and Fikri Karaesmen

Department of Industrial Engineering

Koç University

34450, Sarıyer, İstanbul, TURKEY

ecil@ku.edu.tr, lormeci@ku.edu.tr, fkaraesmen@ku.edu.tr,

April 5, 2006

Structural Results on a Batch Acceptance Problem for Capacitated Queues

Eren Başar Çil, E. Lerzan Örmeci, Fikri Karaesmen

Department of Industrial Engineering, Koç University

34450, Sarıyer, İstanbul, TURKEY

ecil@ku.edu.tr, fkaraesmen@ku.edu.tr, lormeci@ku.edu.tr

Abstract

The purpose of this paper is to investigate the structural properties of the optimal batch acceptance policy in a Markovian queueing problem where different classes of customers arrive in batches and the buffer capacity is finite. We prove that the optimal policy can possess certain monotonicity properties under the assumptions of a single-server and constant batch sizes. Even though our proof cannot be extended to cases where either one of the assumptions is relaxed, we numerically observe that the optimal policy can still possess the same properties when only the single-server assumption is relaxed. Finally, we present counter examples that show the non-monotone structure of the optimal policy when the constant batch size assumption is relaxed.

1 Introduction

Buffer capacity control in production and service systems addresses optimal allocation of fixed buffer resources to different demand segments. Since this objective can be achieved by admission control policies which determine when to accept or reject different segments, this class of control problems has received a lot of attention in the queueing literature. An interesting problem within this class is the case where the customers arrive in batches. In the models with batch arrivals, two types of control policies can be employed: a partial acceptance policy in which some of the jobs in a batch can be admitted while the rest are rejected or a batch acceptance policy where the system can either admit or reject the whole batch. In our study, we focus on a batch acceptance problem in a system where the buffer capacity (i.e. the waiting space) is finite and several classes of customers differing in their rewards arrive in batches. We note that we will refer to systems with finite buffer capacity as capacitated systems. Moreover, the system incurs a holding cost per customer per unit time. The objective of the problem is to maximize the expected total discounted profit over an infinite horizon as well as the expected long-run average profit.

Admission control problems of queueing systems have been studied extensively in the literature. For comprehensive reviews on queueing control problems and their applications on communication networks, we refer to Stidham (2002) and Altman (2002), respectively. Most of the earlier studies focus on systems where customers arrive individually. These systems can further be grouped into two, as systems with no waiting room (which we will refer to as loss systems) and systems with infinite waiting room (which will be referred to as uncapacitated systems). Altman et. al. (2001), Örmeci et. al. (2001), Savin et. al. (2005), Gans and Savin (2004) and Örmeci and van der Wal (2006) fall into the former group, whereas Stidham (1978), Ghoneim and Stidham (1985), Stidham (1985) and Blanc et. al. (1992) consider systems in the latter group. All these studies investigate mainly the structure of optimal admission control problems and prove the optimality of threshold policies.

Admission control problems in queueing systems receiving batch arrivals have been studied as a natural extension of the single arrivals case. Moreover, considering batch arrivals allows to observe the system where customers request more than one resource. When partial acceptance is employed as the control policy, the optimality of threshold policies is shown in several studies; see Örmeci and Burnetas (2004) and Örmeci and Burnetas (2005) for loss systems, see Kulkarni and Tedijanto (1998) for capacitated systems, and see Stidham (1978), Langen (1982) and Helm and Waldmann (1984) for uncapacitated systems. On the other hand, the structure of an optimal batch acceptance problem is analyzed only for an uncapacitated system, where Artiges (1995) shows the optimality of a threshold policy for a discrete-time queueing system with constant batch size. In this paper, our aim is to classify the set of continuous-time Markovian capacitated systems for which this result can be extended. It is well-known that boundaries have a strong effect on the structure of optimal policies. In particular, Örmeci and Burnetas (2004) provide an example of a loss system using batch acceptance policy, which does not possess any monotonicity property. In a related control problem, Kim and van Oyen (1998) denote that even though the well-known $c\mu$ rule is shown to be optimal for many dynamic scheduling problems of uncapacitated systems, it may not be optimal for equivalent scheduling problems of capacitated systems. We note that the $c\mu$ rule is a simple index rule, which always serves an available job with the largest $c\mu$ index; where c is the waiting cost per unit time and μ is the service rate.

As a result, we examine the structure of batch acceptance problems in several capacitated systems. We first investigate the optimal policy in a single-server queue with identical batch sizes for each class and show the optimality of a threshold policy. Unfortunately, our proof cannot be extended to other capacitated systems when any one of the two assumptions, i.e., the single server and the constant batch size, is relaxed. However, we observe through many numerical examples that threshold policies are still optimal for multi-server systems. Finally, we present counter-examples to underline that such monotonicity properties may not exist whenever the constant batch size assumption is relaxed. Hence, we conclude that optimal threshold policies exist for capacitated systems only if the batch sizes are identical for all customer types.

The rest of the paper is organized as follows: In the next section, we build the corresponding Markov decision process (MDP) model of the single-server system with constant batch size. We present the structural properties of the model and the optimal actions in the third section. In section 4, we examine the problem when the single-server assumption is relaxed. Section 5 presents counter examples to the monotonicity property of the optimal policy when the constant batch size assumption is relaxed. Finally, we conclude in the last section.

2 MDP Model

In this section, we build a discrete-time Markov decision process (MDP) for a system employing a batch acceptance policy. We consider a single server queue with waiting room capacity (including the server), K, and N classes of customers. Arrivals occur according to a Poisson process with rate λ. At each arrival epoch, the probability that an arriving batch consists of class-j customers with $j = 1, \ldots, N$ is p_j and the batch size, B, is the same for all classes of customers. Whenever a class-j batch is admitted to the system, it brings a reward of $R_j > 0$ upon its arrival. Without loss of generality, we assume that rewards are ordered as $R_1 \geq R_2 \geq \cdots \geq R_N$. The service times of all admitted customers are exponentially distributed with mean $1/\mu$, regardless of the class of the customer. Moreover, the queue owner incurs a holding cost per unit time as a function of the queue size. We are interested in dynamic admission policies that maximize the total expected discounted profit with a continuous discount rate β over an infinite horizon as well as the long run average profit.

Under any given batch acceptance policy, π, the system evolves as a continuous time Markov chain with state $X(t)$, where $X(t)$ is the number of customers in the system at time t. Then, the state space, S, is the set of non-negative integers less than or equal to K, i.e., $S = \{x : 0 \leq x \leq K\}$. If we denote the number of class-j customers admitted to the system under policy π until time t by $N_j(t)$, the expected total discounted profit of the system starting in state x is given as follows:

$$\mathbb{E}_x^\pi \left[\sum_{j=1}^{N} \int_0^\infty e^{-\beta t} R_j d(N_j(t)) - \int_0^\infty e^{-\beta t} h(X(t)) dt \right], \tag{1}$$

where $h(X(t))$ is the holding cost per unit time when there are $X(t)$ customers in the system.

The objective of the problem is to find the optimal policy π^* that maximizes $\mathbb{E}_x^\pi[.]$. To achieve this aim, we first build the discrete time equivalent of the original system. Since we can interpret discounting as exponential failures with rate β, the arrivals occur according to a Poisson process with rate λ, and the mean service rate is μ, the maximum rate of transition is $\lambda + \mu + \beta$, which is finite. Therefore, we can use uniformization (Lippman, 1975) and normalization to build the discrete time equivalent of the original system. After uniformization and normalization, we assume that the time between two consecutive transitions is exponentially distributed with rate $\lambda + \mu + \beta$, and using the appropriate time scale, assume that $\lambda + \mu + \beta = 1$.

We use the event-based dynamic programming framework introduced by Koole (1998) to define the value function of the system and show some structural properties of the value function. This formalism establishes that if certain event operators, T, satisfy some structural properties under given assumptions, then the value function of the models which can be constructed by using these operators will also satisfy the same structural properties under the same assumptions. Therefore, we build our model as the combination of the operators: departure (T_{DEP}) and batch arrival ($T_{B\_ARR(j)}$). Since the class of the arriving batch is random at each arrival epoch we let the batch arrival operator depend on the class of the arriving batch. In this operator, the states reached as a result of the action taken are the same for all types but the rewards are different for each type of batches.

As we mentioned before, we will describe the structural properties of a system which operates over an infinite horizon. For this purpose, we first prove the structural properties with the objective of maximizing the expected total β-discounted reward for a finite number of transitions, n. The finite horizon problems allow us to use induction to prove the structural properties for all finite n. To start the induction we specify the initial function v^0 as $v^0(x) = 0$ for all x. We denote the maximum expected total β-discounted reward of a system starting in state x when n transitions remain by $v^n(x)$ and present the optimality equation of the finite horizon problem as,

$$v^n(x) = \mu T_{DEP} v^{n-1}(x) + \lambda \sum_{j=1}^{N} p_j T_{B\_ARR(j)} v^{n-1}(x) - h(x) \tag{2}$$

where,

$$T_{DEP} v(x) = \begin{cases} v(x-1) & \text{if } x > 0 \\ v(x) & x = 0, \end{cases}$$

$$T_{B\_ARR(j)} v(x) = \begin{cases} \max\{R_j + v(x+B), v(x)\} & \text{if } x \leq K - B \\ v(x) & \text{if } x > K - B, \end{cases}$$

where $h(x)$ is increasing and convex in x.

We first concentrate on maximizing the total expected β-discounted reward over an infinite horizon. Since both the state and action spaces are finite and the rewards are bounded, there is always an optimal stationary policy due to Theorem 6.2.10 of Puterman (1994). Moreover, this policy can be computed by the value iteration algorithm. Then, all our results for finite horizon problem can be extended to infinite horizon problem with discounting. Specifically, $v(x)$ denotes the value function for the infinite horizon expected discounted reward. Thus for β >0,

$$v(x) = \lim_{n \to \infty} v^n(x).$$

Besides the discounted reward criterion, we can also consider the criterion of maximizing the expected long-run average reward. In this case, we need to define the relative value function, $v'(x)$, and the optimal expected revenue per unit time, g^*. Then, the optimality equation for the average reward criterion is,

$$g^* + v'(x) = \mu T_{DEP} v'(x) + \lambda \sum_{j=1}^{N} p_j T_{B\_ARR(j)} v'(x) - h(x)$$

In addition to the finite state and action spaces and bounded rewards, the corresponding model is unichain since the state 0 is reachable under all possible policies and it is aperiodic due to the fictitious service completions in $x = 0$. Thus, Theorem 8.4.5 of Puterman (1994) guarantees the existence of an optimal stationary policy for the long-run average problem and the validity of the value iteration algorithm to find this policy. Furthermore, Weber and Stidham (1987) establish that the long-run average problem can be obtained as the limit of the infinite horizon discounted problem as $\beta \to 0$ under some specific conditions. The model considered in this section satisfies all of these conditions, so that all results for the infinite horizon problem with discounting hold also for the long-run average problem.

3 Single-server Case with Constant Batch Size: Structure of the Optimal Policy

The first structural property on which we focus is the monotonicity of the value function in x for all states x, explicitly $v(x) \geq v(x+1)$. We can iterate on this property to obtain $v(x) \geq v(x+B)$ which means that admitting an arriving batch induces a positive burden in the system. This is a natural consequence of collecting the rewards immediately upon admitting a batch: The customers already in the system do not generate any additional reward, instead they prevent the system to accept new customers who will bring some reward. We refer to this burden, $v(x) - v(x+B)$, as the opportunity cost of admitting a new batch. In many queuing system problems, the opportunity cost affects the optimal decisions, while the monotonicity of the opportunity cost implies optimal threshold policies in the admission control problems. Therefore, we also work on the monotonicity of the opportunity cost. To simplify the notation, we define $\Delta_B f(x)$ as $\Delta_B f(x) = f(x) - f(x+B)$ for any function f defined on S. We first prove that the operators preserve the monotonicity properties of both $v(x)$ and $\Delta_B v(x)$ in order to prove the properties for the whole system. The following lemma summarizes our results on the monotonicity properties preserved by the operators, and its proof is placed in the appendix.

Lemma 1 *Assume that $v(x)$ is non-increasing and $\Delta_B v(x)$ is non-decreasing in x. Then, $Tv(x)$ is non-increasing in x and $\Delta_B Tv(x)$ is non-decreasing in x for T=T_{DEP} and T=$T_{B\_ARR(j)}$ for all j=1,...,N.*

Now, we observe that the same monotonicity properties hold for the value function: We first prove the properties for $v^n(x)$ for all n by induction. As a result of the specification that $v^0(x) = 0$ for all x, the initial condition of the induction holds for all x. Then, we assume that $v^{n-1}(x) \geq v^{n-1}(x+1)$ and $\Delta_B v^{n-1}(x) \leq \Delta_B v^{n-1}(x+1)$ and prove the same inequalities for $v^n(x)$. Since we use event-based dynamic programming, these inequalities hold for $v^n(x)$ due to Lemma 1 and the monotonicity and concavity of $h(x)$. Hence, we complete the proof of the monotonicity properties for $v^n(x)$ and $\Delta_B v^n(x)$ for all n. Since $v(x) = \lim_{n\to\infty} v^n(x)$, we conclude that the value function, $v(x)$, in this model is also non-increasing in x, and similarly $\Delta_B v(x)$ is non-decreasing in x. The same results are also valid for the average reward criterion.

Finally, we state the effects of the monotonicity of the value functions and the opportunity costs on the optimal policies: Let $l_j^* = \min\{x : v(x) - v(x+B) \geq R_j\}$, where we set $l_j^* = K - B + 1$ if there is no such x, so that, it is optimal to reject a class-j batch in state l_j^*. Since the opportunity cost of an arriving batch, $\Delta_B v(x)$, is non-decreasing in x, the opportunity cost of admitting a class-j batch will continue to exceed the reward obtained by admitting this batch in all states $x \geq l_j^*$ and thus, it is not worth admitting these batches. Therefore, the optimal batch acceptance policies for any class j are of threshold type. Moreover, if the reward obtained by admitting a class-j batch is higher than the reward of a class-i batch, then the optimal threshold of class j will be also higher than that of class i as a result of the definition of l_j^*. The following theorem summarizes our structural results for the discounted problem, which are also valid for the average reward criterion.

Theorem 1 *In the given model, $v(x) \geq v(x+1)$ and $\Delta_B v(x) \leq \Delta_B v(x+1)$, and for every j, there exists a threshold value, $0 \leq l_j^* \leq K - B + 1$, such that an arriving class-j batch is admitted if and only if $x < l_j^*$. Moreover, l_j^*'s are monotone in j, i.e., $l_1^* \geq l_2^* \geq \cdots \geq l_N^*$.*

4 Multi-server Case with Constant Batch Size: Numerical Study

In this section, we investigate the structure of optimal batch acceptance policies when we relax the single-server assumption so that the system has c identical parallel servers. In this case, our methodology to prove the monotonicity of $\Delta_B T_{DEP} v(x)$ requires the concavity of $v(x)$, i.e., $v(x) - v(x+1) \leq v(x+1) - v(x+2)$; and unfortunately we have seen many examples in which $v(x)$ is not concave: Consider a system with 6 identical and parallel servers, 2 classes of batches and no holding costs, i.e., $h(x) = 0$ for all x, where the parameters are set, before normalization, as $K = 15$, $B = 5$, $R_1 = 5$, $R_2 = 10$, $p_1 = 0.75$, $p_2 = 0.25$ and $\lambda = 100$, $\mu = 17$ and $\phi = 0$. The value function $v(x)$ is not concave for this system. As a result, our methodology fails to prove the monotonicity of $\Delta_B T_{DEP} v(x)$. This implies that we cannot prove the monotonicity of the opportunity cost, $\Delta_B v(x)$, and so the existence of an optimal threshold policy. On the other hand, we have not observed any counter-examples to the monotonicity of $\Delta_B v(x)$ in our comprehensive numerical study.

In the numerical study, we consider a multi-server queue with 2 classes of customers and no holding costs, where we fix $K = 15$, $\lambda = 100$, $B = 5$, and $R_2 = 10$. Then we let μ change in range [1,250], R_1 in range [1,50], the number of servers, c in range [2,6], all with increments of 1, and p_1 in a range [0.01,0.80] with 0.01 increments. In this way, we generate 5,000,000 different instances, and observe that in all these instances $\Delta_B v(x)$ is non-decreasing in x and the optimal policy is a threshold policy. Our intuition also agrees with these observations: Whenever the batch sizes are identical, all classes use the available capacity in the same way, so that the only criterion to compare different classes is their rewards, which naturally induce the same order in admission control. Based on this intuition and the numerical evidence, despite the lack of proof, we can state the following conjecture:

Conjecture 1 *In a multi-server queue with constant batch sizes, there exists a threshold value for every j, $0 \leq l_j^* \leq K - B + 1$, such that an arriving class-j batch is admitted if and only if $x < l_j^*$. Moreover, l_j^*'s are monotone in j, i.e., $l_1^* \geq l_2^* \geq \cdots \geq l_N^*$.*

5 Random Batch Size: Counter Examples

In this section, we present counter examples to show that optimal batch acceptance policies do not have any structural properties when the constant batch size assumption is relaxed. Differences in batch sizes bring a new criterion to evaluate the classes, namely how the classes use the available resources. This criterion is quite different than the rewards, as the value of classes may change drastically when most of the resources are being used, i.e., when the system state is close to the boundary. In all capacitated systems, boundary effects are observed. In systems with constant batch sizes, these effects are identical for all classes, so that they affect optimal admission policies only in a monotone way. For systems with random batch sizes, on the other hand, the boundary effects on different classes depend on their batch sizes, which, in turn, induce non-monotonicity on optimal admission policies. The following examples show these non-monotone effects first in a single-server system, then for a multi-server system.

5.1 Single-server Case

In this example, we observe the optimal batch acceptance policy of a single-server queue with two classes of customers, and no holding costs. Class-1 customers arrive in 5-unit batches whereas class-2 customers arrive in 1-unit batches. The parameters are: $K = 8$, $\lambda = 10$, $p_1 = 0.7$, $p_2 = 0.3$, $R_1 = 50$, $R_2 = 10$, $\mu = 1/2$, $\phi = 1$.

As it can be seen in Table 1, the optimal policy accepts both classes whenever it can, except for state 3. In this state, the policy rejects the class-2 customers in order to wait for a 5-batch, which will fully use the resources. If one small job is accepted in state 3, most of the capacity will stay idle until many class-2 customers arrive at the system. Therefore, the optimal batch acceptance policy

# of customers in the system	Optimal policy for class 1	Optimal policy for class 2
0	1	1
1	1	1
2	1	1
3	1	0
4	0	1
5	0	1
6	0	1
7	0	1
8	0	0

Table 1: The optimal batch acceptance policy for single-server queue and random batches

does not possess any monotonicity property when we consider a single-server queue and random batches.

5.2 Multi-server Case

In this case, we only change the number of servers from 1 to 3 and the waiting room capacity from 8 to 20 in the previous example and observe that the system rejects class-2 customers only in state 15. In other words, the optimal policy still does not possess any monotonicity property. The intuition behind this result is the same as the previous one.

6 Conclusion

The aim of this study has been to characterize the set of capacitated systems for which there exist optimal batch admission policies that are of threshold type. We are able to show the existence of optimal threshold policies for systems with single servers and constant batch sizes, while we have strong numerical evidence that this result extends to multi-server systems with constant batch sizes. Finally, when the batch sizes differ, we know that optimal batch admission policies do not have any monotone behavior in general as evidenced by the counterexamples. Hence, we can conclude that the set of capacitated systems with monotone batch admission policies is restricted to those receiving batches with identical batch sizes. As a result, this paper presents a complete analysis of capacitated systems which employ batch admission policies addressing all monotonicity questions regarding optimal policies.

References

E. Altman (2002). Applications of Markov Decision Processes in Communication Networks. *Handbook of Markov Decision Processes: Methods and Applications.* Eds. E. A. Feinberg and A. Shwartz Kluwer.

E. Altman, T. Jimenez, and G. M. Koole (2001). On optimal call admission control in a resource-sharing system. *IEEE Transactions on Communications*, 49:1659–1668.

D. Artiges (1995). Optimal routing into two heterogeneous service stations with delayed information. *IEEE Transactions on Automatic Control*, 40:1234–1236.

J. P. C. Blanc, P. R. de Waal, P. Nain, and D. Towsley (1992). Optimal control of admission to a multiserver queue with two arrival streams. *IEEE Transactions on Automatic Control*, 37:785-797.

N. Gans and S. Savin (2004). Pricing and capacity rationing in rentals. Working Paper, Wharton School of Business, University of Pennsylvania.

H. A. Ghoneim and S. Stidham (1995). Control of arrivals to two queues in series. *European Journal of Operations Research*, 21:399–409.

W. E. Helm and K. H. Waldmann (1984). Optimal control of arrivals to multiserver queues in random environment. *Management Science*, 21:602–615.

E. Kim and M. P. Van Oyen (1998). Beyond the $c\mu$ rule: Dynamic scheduling of a two-class loss queue. *Mathematical Methods of Operations Research*, 48:17–36.

G. Koole (1998). Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems*, 30:323–339.

V. G. Kulkarni and T. E. Tedijanto (1998). Optimal admission control of Markov-Modulated Batch Arrivals to a Finite Capacity Buffer. *Stochastic Models*, 14:95–122.

H. J. Langen (1982). Applying the method of phases in the optimization of queueing systems. *Advances in Applied Probability*, 14:122–142.

S. A. Lippman and S. M. Ross (1971). The streetwalker's dilemma: A job shop model. *SIAM Journal of Applied Mathematics*, 20:336–342.

S. A. Lippman (1975). Applying a new device in the optimization of exponential queuing systems. *Operations Research*, 23:687–710.

E. L. Örmeci and A. Burnetas (2004). Admission control with batch arrivals. *Operations Research Letters*, 32:448–454.

E. L. Örmeci and A. Burnetas (2005). Dynamic admission control for loss systems with batch arrivals. *Advances in Applied Probability*, 37:915–937.

E. L. Örmeci, A. Burnetas, and J. van der Wal (2001). Admission policies for a two class loss system. *Stochastic Models*, 17:513–540.

E. L. Örmeci and J. van der Wal (2006). Admission policies for a two class loss system with general interarrival times. *Stochastic Models*, 22:37–53.

M. L. Puterman (1994). *Discrete Stochastic Dynamic Programming*. John Wiley and Sons, USA.

S. Savin, M. Cohen, N. Gans, and Z. Katalan (2005). Capacity management in rental businesses with heterogeneous customer bases. *Operations Research*, 53:617–631.

S. Stidham (1978). Socially and individually optimal control of arrivals to a GI/M/1 queue. *Managemet Science*, 24:1598–1610.

S. Stidham (1985). Optimal control of admission to a queuing system. *IEEE Transactions on Automatic Control*, 30:705–713.

S. Stidham (2002). Analysis, design and control of queuing systems. *Operations Research*, 50:197–216.

R. R. Weber and S. Stidham (1987). Optimal control of service rates in network queues. *Advanced Applied Probability*, 19:202–218.

Appendix

A Proof of Lemma 1

A.1 Monotonicity of $T_{DEP}v(x)$ and $\Delta_B T_{DEP}v(x)$

Due to the definition of the departure operator, the operator preserves the structural properties of $v(x)$ whenever $x > 0$. Therefore, the proof of Lemma 1 for T_{DEP} is trivial whenever $x > 0$. However, we have to show that equations (3) and (4) also hold for $x = 0$ in order to complete the proof of the lemma:

$$T_{DEP}v(0) \geq T_{DEP}v(1) \tag{3}$$

$$T_{DEP}v(0) - T_{DEP}v(B) \leq T_{DEP}v(1) - T_{DEP}v(B+1) \tag{4}$$

Equation (3) can be rewritten as $v(0) \geq v(0)$ by using the definition of the operator, which is obviously true. Similarly, we can write equation (4) as $v(0) - v(B-1) \leq v(0) - v(B)$ and it holds by the monotonicity of $v(x)$. Therefore, both equation (3) and (4) hold and the proof of Lemma 1 for the departure operator is completed. □

A.2 Monotonicity of $T_{B\_ARR(j)}$

In this proof, we show the monotonicity of $T_{B\_ARR(j)}v(x)$ when $v(x)$ is non-increasing in x. In other words, we want to prove the following inequality for all $j = 1, \ldots, N$.

$$T_{B\_ARR(j)}v(x) \geq T_{B\_ARR(j)}v(x+1) \tag{5}$$

We first compare the states x and $x+1$, so that we define an optimal action vector $\bar{a}_j$ such that $\bar{a}_j = (a_x, a_{x+1})$ where a_x and a_{x+1} are optimal actions for a given state x and $x+1$ when a class-j batch arrives, respectively. An optimal action can be either admitting the whole batch, 1, or rejecting the whole batch, 0. Although there are 4 permutations of two consecutive optimal actions, due to the monotonicities of $v(x)$ and $\Delta_B v(x)$ the action $\bar{a}_j = (0, 1)$ cannot be optimal, so that it is enough to consider the cases $(1, 1)$, $(1, 0)$, and $(0, 0)$. We study these cases first for the condition $x < K - B$, then for the condition $x = K - B$, and finally for the condition $x > K - B$ in order to examine the boundary effect.

Cases	$\bar{a}_j = (a_x, a_{x+1})$	Rewritten form of equation (5)
Case I	(1,1)	$R_j + v(x+B) \geq R_j + v(x+B+1)$
Case II	(1,0)	$R_j + v(x+B) \geq v(x+1)$
Case III	(0,0)	$v(x) \geq v(x+1)$

Table 2: Possible optimal actions in states x and $x+1$ for the condition $x < K - B$

Cases	$\bar{a}_j = (a_x, a_{x+1})$	Rewritten form of equation (5)
Case I	(1,0)	$R_j + v(K) \geq v(K-B+1)$
Case II	(0,0)	$v(K-B) \geq v(K-B+1)$

Table 3: Possible optimal actions in states x and $x+1$ for the condition $x = K - B$

A.2.1 $x < K - B$

For this condition, we can write equation (5) for each case as in Table 2. Cases I and III hold by the monotonicity of $v(x)$. However, case II holds not only by the monotonicity of $v(x)$ but also due to the optimal action in state x: Since the optimal policy admits an arriving batch in state x, we have that $R_j + v(x+B) \geq v(x)$. Moreover, $v(x) \geq v(x+1)$ by the monotonicity of $v(x)$. By combining these two results, equation (5) holds, i.e., $R_j + v(x+B) \geq v(x+1)$, when $\bar{a}_j = (1,0)$. Therefore, the batch arrival operator, $T_{B\_ARR(j)}$, is non-increasing in x whenever $x < K - B - 1$.

A.2.2 $x = K - B$

In this condition, we only need to study the cases $(1,0)$ and $(0,0)$ because the only feasible action for $x > K - B$ is rejecting an arriving batch. We can rewrite equation (5) for theses cases as in Table 3. Case II holds by the monotonicity of $v(x)$ and case I holds by both the monotonicity of $v(x)$ and the optimal action in state x, as in the case II of the previous condition. Thus, $T_{B\_ARR(j)}$ is non-increasing in x for $x = K - B$.

A.2.3 $x > K - B$

The condition $x > K - B$ is trivial because the only possible action for states $x > K - B$ is to reject the arriving batches and monotonicity of the batch arrival operator follows from the monotonicity of $v(x)$.

Thus, we complete the proof of the monotonicity of the batch arrival operator for all states x.□

Cases	$\bar{a}_j = (a_x, a_{x+1}, a_{x+B}, a_{x+B+1})$	Rewritten form of equation (6)
Case I	(1,1,1,1)	$v(x+B) - v(x+2B) \leq v(x+B+1) - v(x+2B+1)$
Case II	(1,1,1,0)	$v(x+B) - v(x+2B) \leq R_j$
Case III	(1,1,0,0)	$R_j \leq R_j$
Case IV	(1,0,0,0)	$R_j \leq v(x+1) - v(x+B+1)$
Case V	(0,0,0,0)	$v(x) - v(x+B) \leq v(x+1) - v(x+B+1)$

Table 4: Possible optimal actions in states x, $x+1$, $x+B$ and $x+B+1$ for the condition $x < K - 2B$

A.3 Monotonicity of $\Delta_B T_{B\_ARR(j)}$

In this proof, we show the monotonicity of $\Delta_B T_{B\_ARR(j)} v(x)$ when $\Delta_B v(x)$ is non-decreasing in x. In other words, we want to prove the following inequality for all $j = 1, \ldots, N$.

$$\Delta_B T_{B\_ARR(j)} v(x) \quad \leq \quad \Delta_B T_{B\_ARR(j)} v(x+1) \tag{6}$$

To show the monotonicity of $\Delta_B T_{B\_ARR(j)} v(x)$, we compare the states x, $x+1$, $x+B$, and $x+B+1$. Therefore, the optimal action vector $\bar{a}_j$ is defined as $\bar{a}_j = (a_x, a_{x+1}, a_{x+B}, a_{x+B+1})$ where a_x, a_{x+1}, a_{x+B} and a_{x+B+1} are the optimal actions for states x, $x+1$, $x+B$, and $x+B+1$ when a class-j batch arrives, respectively. Although there are 16 optimal action permutations for these 4 states, it is enough to consider the cases $(1,1,1,1)$, $(1,1,1,0)$, $(1,1,0,0)$, $(1,0,0,0)$, and $(0,0,0,0)$ due to the assumptions on $v(x)$ and $\Delta_B v(x)$. As in the monotonicity of the operator, we prove the monotonicity of $\Delta_B T_{B\_ARR(j)} v(x)$ first for the condition $x < K - 2B$, then for the condition $x = K - 2B$ and finally for the condition $K - 2B < x \leq K - B - 1$ to observe the boundary effect. Since our state space is bounded by K, it is not necessary to consider the monotonicity of $\Delta_B T_{B\_ARR(j)} v(x)$ for $x \geq K - B$.

A.3.1 $x < K - 2B$

For this condition, we can rewrite equation (6) for each case as in Table 4. Case III is obvious and cases I and V hold by the monotonicity of $\Delta_B v(x)$. In case II, $v(x+B) \leq R_j + v(x+2B)$ as a result of the optimal action in state $x+B$ and thus, case II also holds. Similarly, Case IV holds by the optimal action in state $x+1$ in this case. Therefore, we finish the proof of the monotonicity of $\Delta_B T_{B\_ARR(j)} v(x)$ for the condition $x < K - 2B$.

A.3.2 $x = K - 2B$

For this condition, we do not need to study all of the five cases mentioned for $x < K - 2B$. Since the only feasible action in states $x \geq K - B + 1$ is 0 (i.e., rejecting an arriving batch) because of the capacity, we only focus on the cases: $(1, 1, 1, 0)$, $(1, 1, 0, 0)$, $(1, 0, 0, 0)$, and $(0, 0, 0, 0)$. The proofs of these cases are similar to the cases II, III, IV, and V in **A.3.1**.

A.3.3 $K - 2B < x \leq K - B - 1$

Similar to the previous condition, we only need to study the cases, $(1, 1, 0, 0)$, $(1, 0, 0, 0)$, and $(0, 0, 0, 0)$. The proofs of these cases are similar to the cases III, IV, and V in **A.3.1**.

Hence, $\Delta_B T_{B\_ARR(j)} v(x)$ is non-decreasing in x for all states . □