The Shuttle Dispatch Problem with Compound Poisson Arrivals: Controls at Two Terminals

Hyo-Seong Lee

M. M. Srinivasan

Department of Industrial and Operations Engineering

The University of Michigan, Ann Arbor 48109-2117

Technical Report 88-6
June 1988

Abstract

We consider the control of an infinite capacity shuttle which transports passengers between two terminals. The passengers arrive at each terminal according to a compound Poisson process and the travel time from one terminal to the other one is a random variable following an arbitrary distribution. The following control limit policy is considered: dispatch the shuttle at terminal i, at the instant that the total number of passengers waiting at terminal i reaches or exceeds a predetermined control limit m_i. The objective of this paper is to obtain the mean waiting time of an arbitrary passenger at each terminal for given control values m_1 and m_2. We also discuss a search procedure to obtain the optimal control values which minimize the total expected cost per unit time under a linear cost structure.

1. Introduction

We consider a transportation system in which a single shuttle transports passengers between two terminals. All passengers who arrive at a given terminal are to be transported to the other terminal. We will term the passengers who arrive at terminal i as *type i passengers*. We assume that type i passengers arrive at terminal i according to an independent compound Poisson process with rate λ_i (i=1,2). The batch size of each arrival of type i passengers is an independent random variable $X_i \geq 1$, following an arbitrary discrete distribution. The interterminal travel time from terminal i to the other terminal is a random variable D_i having an arbitrary distribution, which is assumed to be independent of the number of passengers carried by the shuttle. We assume that the time needed to board the shuttle is negligible. The capacity of the shuttle is assumed to be infinite so that all the waiting passengers at the terminal can be carried all at once.

For this transportation system (see figure 1), we consider the following operating policy, which is usually termed as a *control limit policy*: dispatch the shuttle at terminal i at the instant that the total number of passengers waiting at terminal i reaches or has exceeded a predetermined threshold value m_i (i=1,2); if the number of waiting passengers is equal to or greater than m_i when the shuttle arrives at terminal i, the shuttle immediately leaves terminal i for the other terminal. On the other hand, if the number of passengers at terminal i at that time is below m_i, the shuttle is held until the number of waiting passengers reaches or exceeds m_i. The objective of this paper is to obtain an expression for the mean waiting time of the passengers at each terminal under a given control limit policy.

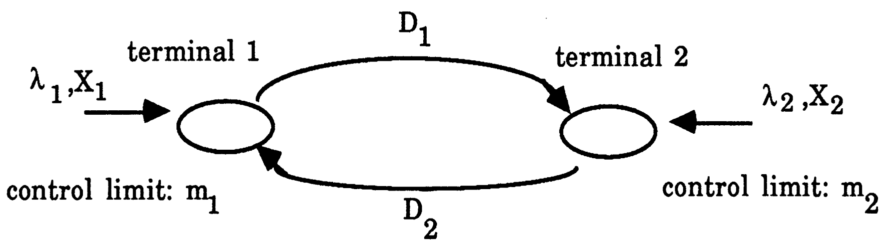

Fig. 1. Dispatching of a shuttle with controls at two terminals.

A considerable amount of work has been done on the shuttle dispatching problem with one or two terminals. Comprehensive surveys on this problem can be found in the papers by Sim and Templeton(1983) and Teghem(1986).

For a single terminal dispatching problem with Poisson arrivals, Weiss(1979) developed an algorithm to find the optimal control value under a linear cost structure assuming that the capacity of the shuttle is infinite. Deb and Serfozo(1973) proved that a control limit policy is an optimal operating policy under a linear cost structure for the case of a single terminal with finite shuttle capacity. Recently, Powell(1985,1986) studied more general strategies for dispatching a shuttle which includes both holding and cancellation strategies. Under an assumption of compound Poisson arrivals, he developed an algorithm to find the mean waiting time of an arbitrary passenger for a given policy.

For the shuttle dispatching problem with two terminals, Ignall and Kolesar(1974) and Weiss(1981) studied infinite capacity shuttle dispatching problems assuming that the dispatcher can hold the shuttle for more passengers at only one of the terminals. In their paper, Ignall and Kolesar proved that the optimal policy is to dispatch the carrier if and only if the *total* number of passengers waiting at both terminals is greater than a threshold value. However, since a shuttle is held only at one terminal, this problem is relatively simple compared to the problem with controls at two terminals. For a shuttle dispatching problem with controls at two terminals, Deb(1978) presented a significant result on the characteristics of the optimal dispatching policy. Under a linear cost structure, he proved that the optimal policy which minimizes the expected total discounted cost over an infinite time horizon has the following form: suppose the shuttle is at one of the terminals with x passengers waiting there and y passengers waiting at the other terminal. Then the optimal policy is to dispatch the shuttle if, and only if, $x \geq G(y)$, where G(.) is a monotone decreasing control function. Unfortunately, however, the explicit determination of the function G(.) is still not known.

Although our operating policy appears to be similar to the policy described by Deb, there is a fundumental difference between our model and Deb's model as explained below:

First, our model assumes that arrival of the passengers follows a compound Poisson process and the capacity of the shuttle is infinite while Deb's model assumes that passengers arrive according to a Poisson process and the capacity of the shuttle is finite.

Second, in Deb's model, it is assumed that the number of passengers at both terminals is always known, so that the information at both terminals is used to decide whether the shuttle is dispatched or not. Although this assumption gives a natural structure for the dynamic programming modeling, in many cases, the information on the number of passengers at one terminal is not available at the other terminal. In our study, we assume that the number of passengers at a terminal is known only when the shuttle is present at that terminal. Consequently, in our model, only the number of passengers who are waiting at the terminal where the shuttle is present is used to decide whether the shuttle is dispatched to the other terminal or not. Although we use the information only at one terminal in deciding whether the shuttle is dispatched or not, the number of passengers at the other terminal must be implicitly considered in the analysis because the sojourn time of the shuttle at one terminal affects the sojourn time of the shuttle at the other terminal. Due to this consideration, the analysis becomes fairly complicated.

One approach to solve the shuttle dispatching problem is to use a semi-Markov decision process. Although this approach has been quite successful in proving the optimality of the control limit policies in some cases(Deb and Serfozo 1973, Deb 1978), it usually needs considerable computational effort. In addition, the Markov decision process approach does not provide some important performance statistics such as the mean waiting time for each type of passenger. Since our objective in this study is to obtain the mean waiting time for each type of passengers, hence in order to solve the problem with controls at two terminals, we adopt a different approach as follows. We decompose the system into two individual terminals and then analyze each individual terminal separately as a single terminal shuttle dispatching problem. Thus, in order to solve the problem with two terminals, we must first be able to solve a single terminal dispatching problem. Although the algorithms developed by Powell(1985,1986) solve a very general single terminal problem, for the infinite capacity shuttle dispatching problem described below, the procedure presented in the next section is much more efficient.

2. The single terminal dispatching problem

We consider here a shuttle which services passengers who arrive at a single terminal. Passengers arrive at this terminal according to a compound Poisson process with rate λ. The batch size of each arriving passenger is a random variable $X \geq 1$, following an arbitrary distribution. To operate the system, we use the following control limit policy: when the shuttle is at the terminal and the number of passengers waiting there is less than the predetermined control limit m, the shuttle waits at the terminal until the total number of waiting passengers reaches or exceeds m. As soon as the number of waiting passengers reaches or exceeds m, the shuttle leaves the terminal to transport the passengers to their destination and then returns to the terminal after a random amount of time V, which we shall call the *intervisit time*. If the number of passengers waiting at the terminal is equal to or greater than m at the instant the shuttle comes back to the terminal, it leaves the terminal immediately carrying all the waiting passengers. On the other hand, if the shuttle finds less than m passengers, it waits at the terminal until the number of waiting passengers reaches or exceeds m.

2.1. Notation

For analytical convenience, for any discrete random variable (r.v.), A, that is used in the analysis, we adopt the following notation throughout the paper:

$a_i = P\{A=i\}$,

$A(z) = \sum_{i=0}^{\infty} a_i z^i$, the probability generating function (p.g.f.) of A,

$a^{(1)} = E(A)$,

$a^{(2)} = E(A(A-1))$.

Let

Q = number of batches that arrive during an intervisit time V, r.v.,

R = number of passengers that arrive during V, r.v.,

I = total number of passengers at the terminal when the shuttle leaves the terminal, r.v.,

J_n =number of passengers who arrive during the sojourn time of the shuttle at the terminal if the number of passengers at the shuttle arrival instant is m-n, $0 \leq n < m$, r.v.,

$x_j^{*(i)}$ = P{i-fold convolution of X is j},

L = mean queue length of the waiting passengers at the terminal,

W = mean waiting time of an arbitrary passenger at the terminal.

2.2. The Analysis

It is shown in Lee and Srinivasan(1987) that if we know the first two moments of I, which is the number of passengers when the shuttle leaves the terminal, the mean waiting time of an arbitrary passenger experienced at the terminal can be obtained as follows:

$$W = \frac{1}{2}\left(\frac{i^{(2)}}{i^{(1)}} - \frac{x^{(2)}}{x^{(1)}}\right) / \lambda x^{(1)}. \tag{2.1}$$

To obtain the first two moments of I, however, we must know both r_j, which is the probability that j passengers arrive during the intervisit time for $j=0,\ldots,m-1$, and the first two moments of R. Suppose we know the values of q_i. Then, r_j can be obtained by

$$r_j = \sum_{i=0}^{j} q_i x_j^{*(i)}, \quad j=0,1,2,\cdots, \tag{2.2}$$

where $x_0^{*(0)}=1$.

To obtain the q_i values, note that by definition,

$$q_i = \int_0^\infty \frac{(\lambda t)^i}{i!} e^{-\lambda t}\, dV(t). \tag{2.3}$$

However, for the computation of q_i, we use the following equivalent expression:

$$q_i = \frac{(-\lambda)^i}{i!} V^{*(i)}(\lambda), \tag{2.4}$$

where $V^{*(i)}(\theta)$ is the i^{th} derivative of the Laplace Stieltjes Transform (LST) of V with respect to θ.

The moments of R are obtained by using the following result, stated as lemma 2.1. Since lemma 2.1 can be easily proved, we omit the proof here.

Lemma 2.1. The probability generating function of R is given by

$$R(z) = V^*(\lambda - \lambda X(z)). \tag{2.5}$$

■

From equation (2.5), we can obtain the first and second factorial moments of R as

$$r^{(1)} = E(R) = \lambda E(V) x^{(1)}, \tag{2.6.a}$$

$$r^{(2)} = E(R(R-1)) = (\lambda x^{(1)})^2 E(V^2) + \lambda x^{(2)} E(V). \quad (2.6.b)$$

We can now obtain an expression for the probability generating function of I from the following result, stated as theorem 2.1.

Theorem 2.1

$$I(z) = R(z) + \sum_{j=0}^{m-1} r_j z^j \{J_{m-j}(z) - 1\}, \quad (2.7)$$

where

$$J_n(z) = X(z) + \sum_{j=1}^{n-1} x_j z^j \{J_{n-j}(z) - 1\}. \quad (2.8)$$

Proof By conditioning on the number of passengers who arrive during an intervisit time, we can obtain a recursive equation that yields P{I=k} as

$$P\{I=k\} = r_k + \sum_{j=0}^{m-1} r_j \{J_{m-j} = k-j\}, \qquad k \geq m. \quad (2.9)$$

From equation (2.9), the probability generating function of I can be obtained as

$$\begin{aligned} I(z) &= \sum_{k=m}^{\infty} P\{I=k\} z^k \\ &= \sum_{k=m}^{\infty} r_k z^k + \sum_{j=0}^{m-1} r_j\, z^j \sum_{k=m}^{\infty} P\{J_{m-j}=k-j\} z^{k-j} \\ &= \sum_{k=m}^{\infty} r_k z^k + \sum_{j=0}^{m-1} r_j\, z^j\, J_{m-j}(z) \\ &= R(z) + \sum_{j=0}^{m-1} r_j\, z^j \{J_{m-j}(z) - 1\}. \quad (2.10) \end{aligned}$$

The term $J_n(z)$, which is used in equation (2.10), can be obtained as follows. Note that, since the number of passengers is m-n when the shuttle arrives at the terminal, the shuttle must stay at the terminal until at least n more passengers arrive. By conditioning on the number of passengers who arrive in the first batch, we can obtain a recursive equation for $P\{J_n = k\}$ as

$$P\{J_n=k\} = \sum_{j=1}^{n-1} x_j P\{J_{n-j}=k-j\} + x_k, \quad k\geq n. \tag{2.11}$$

From (2.11), the p.g.f. of J_n is expressed by

$$\begin{aligned} J_n(z) &= \sum_{k=n}^{\infty} P\{J_n=k\}z^k = \sum_{k=n}^{\infty} \left\{ \sum_{j=1}^{n-1} x_j P\{J_{n-j}=k-j\}+x_k \right\} z^k \\ &= \sum_{j=1}^{n-1} x_j z^j \sum_{k=n}^{\infty} P\{J_{n-j}=k-j\} z^{k-j} + \sum_{k=n}^{\infty} x_k z^k \\ &= \sum_{j=1}^{n-1} x_j z^j \{J_{n-j}(z)-1\}+X(z). \end{aligned} \tag{2.12}$$

■

From I(z), we can obtain the first and second factorial moments of I as

$$i^{(1)} = E(I) = r^{(1)} + \sum_{j=0}^{m-1} r_j j_{m-j}^{(1)}, \tag{2.13.a}$$

$$i^{(2)} = E(I(I-1)) = r^{(2)} + \sum_{j=0}^{m-1} r_j \left(2j\, j_{m-j}^{(1)} + j_{m-j}^{(2)}\right), \tag{2.13.b}$$

where $j_n^{(1)}$ and $j_n^{(2)}$ are obtained from equation (2.8) as

$$j_n^{(1)} = x^{(1)} + \sum_{j=1}^{n-1} x_j j_{n-j}^{(1)}, \tag{2.14.a}$$

$$j_n^{(2)} = x^{(2)} + \sum_{j=1}^{n-1} x_j \left(2 j_{n-j}^{(1)} + j_{n-j}^{(2)}\right). \tag{2.14.b}$$

By substituting $i^{(1)}$ and $i^{(2)}$ obtained in (2.13) into (2.1), we can find the mean waiting time experienced by an arbitrary passenger at the terminal.

3. Analysis for a problem with two terminals

3.1 General Approach

Now using the result obtained for the shuttle dispatching problem with a single terminal, we will analyze the problem with two terminals. As shown in section 2, the mean waiting time of an arbitrary passenger at a terminal can be obtained if we know the first two moments for the number of passengers at the terminal at the instant that the shuttle leaves the terminal. It was also observed that in order to obtain these two moments, we must know the probabilities for the number of passengers (from 0 to m_i-1) present at the terminal when the shuttle arrives at the terminal as well as the first two moments of the number of these passengers. In case of a shuttle dispatching problem with two terminals, however, these values cannot be obtained easily because the sojourn time of the shuttle at one terminal and the travel time from that terminal to the other one affect the sojourn time of the shuttle at the other terminal.

To obtain the probabilities for the number of passengers (from 0 to m_i-1) present at the terminal at the shuttle arrival instant, we will restrict our attention to the following four types of epochs on the path of the shuttle. Let the instant that the shuttle arrives at terminal 1 be an epoch of type 1 and the instant that the shuttle leaves terminal 1 be an epoch of type 2. Similarly, let the instant that the shuttle arrives at terminal 2 be an epoch of type 3 and the instant that the shuttle leaves terminal 2 be an epoch of type 4. Henceforth, we choose to call an epoch of type i as epoch i. Thus, the shuttle moves on from epoch 1 to epoch 4 via epochs 2 and 3, and comes back to epoch 1. At this point, the next cycle is initiated. Note that we must find the steady state probabilities for the number of passengers (from 0 to m_i-1) at epochs 1 and 3. However, for this purpose, we will use epochs 2 and 4 instead of epochs 1 and 3 because by doing so we can reduce the number of states, and hence the computational effort significantly. This advantage comes from the observation that the number of type 1 passengers at epoch 2 and the number of type 2 passengers at epoch 4 are always zero. Once we obtain the steady state probabilities for the queue lengths (from 0 to m_i-1) at epochs 2 and 4, we can easily obtain the steady state probabilities for the queue lengths (from 0 to m_i-1) at epochs 1 and 3.

Define the states at epochs 1 and 4 to be k if the number of type 1 passengers is k for $k<m_1$. If the number of type 1 passengers at these epochs is equal to or greater than m_1, the state is defined to be m_1. Similarly, define the states at epochs 2 and 3 to be k if the number of type 2 passengers at these epochs is k for $k<m_2$. If the number of type 2 passengers is equal to or greater than m_2, the state is defined to be m_2. Thus, the state space at epochs 1 and 4 is $\{0,1,\ldots,m_1\}$ and the state space at epochs 2 and 3 is $\{0,1,\ldots m_2\}$. Let us define I_s^t to be the state at epoch s at the t^{th} cycle and define $q_s(i,j)$ to be

$$q_1(i,j)=P\{I_1^{t+1}=j \mid I_4^t=i\}, \tag{3.1.a}$$

$$q_2(i,j)=P\{I_2^{t+1}=j \mid I_4^t=i\}, \tag{3.1.b}$$

$$q_3(i,j)=P\{I_3^{t}=j \mid I_2^t=i\}, \tag{3.1.c}$$

$$q_4(i,j)=P\{I_4^{t}=j \mid I_2^t=i\}. \tag{3.1.d}$$

So, for instance, $q_4(i,j)$ represents a transition probability that the state at epoch 4 will be j given that the state at epoch 2 is i.

Let $p_2(i,j)$ to be the transition probability, $P\{I_2^{t+1} = j \mid I_2^t = i\}$ and let $p_4(i,j)$ be the transition probability, $P\{I_4^{t+1} = j \mid I_4^t = i\}$. Then, these probabilities can be obtained from the following equations:

$$p_2(i,j) = \sum_{k=0}^{m_1} q_4(i,k)q_2(k,j), \tag{3.2.a}$$

$$p_4(i,j) = \sum_{k=0}^{m_2} q_2(i,k)q_4(k,j). \tag{3.2.b}$$

To represent these equations in a more compact form, we define matrices Q_s and P_s such that

$$Q_s = \left(q_s(i,j) \right) \quad \text{for } s=1 \text{ to } 4,$$

$$P_s = \left(p_s(i,j) \right) \quad \text{for } s=2 \text{ and } 4.$$

Then, equation (3.2) can be represented by

$$P_2 = Q_4Q_2, \tag{3.3.a}$$

$$P_4 = Q_2Q_4. \tag{3.3.b}$$

Suppose we know the matrices Q_s for s=1 to 4. Then, we can find matrices P_2 and P_4 from equation (3.3). Let π_{ik} be the steady state probability that the state at epoch i is k and let π_i be vectors such that

$$\pi_i=(\pi_{i0},\pi_{i1},\ldots,\pi_{im_1}) \text{ for } i=1 \text{ and } 4,$$

$$\pi_i=(\pi_{i0},\pi_{i1},\ldots,\pi_{im_2}) \text{ for } i=2 \text{ and } 3.$$

Since matrices P_2 and P_4 are ergodic, using the transition probabilities $p_2(i,j)$ and $p_4(i,j)$ obtained from (3.3), we can compute π_{ik} for i=2 and 4 by solving the following equations:

$\pi_i = \pi_i P_i,$

$e\pi_i = 1, \qquad \text{for } i=2,4. \qquad (3.4)$

Recall that we want to find the steady state probabilities π_1 and π_3 using π_2 and π_4, which can be easily done from

$$\pi_1 = \pi_4 Q_1, \qquad (3.5.a)$$

$$\pi_3 = \pi_2 Q_3. \qquad (3.5.b)$$

Thus, if we know Q_k for k=1 to 4, then π_2 and π_4 can be obtained from Q_2 and Q_4, and given π_2, π_4, Q_1 and Q_3, then π_1 and π_3 can be obtained. Once π_1 and π_3 are obtained, we can also compute the first two moments of the intervisit times, which yield the first two moments of the number of passengers at the shuttle arrival instant. With all this information, we can finally compute the mean waiting time for each type of passenger for given values of m_1 and m_2. Each of these steps will be described below in detail.

3.2. Notation

The notation used for a discrete random variable in section 2 will continue to be used in this section. In addition to that, in this section, for any continuous random variable C used in the analysis, $C^*(.)$ will denote the LST of C.

We define

I_i = total number of passengers at terminal i when the shuttle leaves terminal i, r.v.,

V_i = intervisit time of the shuttle for terminal i in steady state, r.v.,

U_i = sojourn time of the shuttle at terminal i in steady state, r.v.,

$U_{i,n}$ = sojourn time of the shuttle at terminal i given that the state was m_i-n at the time the shuttle arrives to terminal i (i=1,2), r.v.,

$S_1 \quad = D_1 + U_2$, r.v.,

$S_2 \quad = D_2 + U_1$, r.v.,

$x_{ij}^{*(k)} \quad = P\{$k-fold convolution of X_i is $j\}$,

W_i = mean waiting time of an arbitrary passenger of type i.

3.3. The Analysis

In this section, we will derive an expression for W_i, the mean waiting time for an arbitrary passenger of type i. As described in section 3.1, to obtain W_i, the transition matrices Q_k for k=1 to 4 must be computed. While the matrices Q_1 and Q_3 can be easily obtained, the matrices Q_2 and Q_4 are not easy to obtain because of the interrelationships among the random variables. We now describe how the matrix Q_4, namely, each value of $q_4(i,j)$, can be computed. We will not show how the matrix Q_2 is obtained because it can be computed in exactly the same way as Q_4 is computed. Note that $q_4(i,j)$ can be obtained if we can compute the probabilities for the number of type 1 passengers that arrive during S_1 given that the state at epoch 2 was i. To obtain this, let us define a random variable $S_{1,i}$ such that

$P\{S_{1,i} \le t\} = P\{S_1 \le t \mid \text{the state at epoch 2 was i}\}.$

Since the probability for the number of type 1 passengers that arrive during $S_{1,i}$ can be obtained if we know the LST of $S_{1,i}$, we will first find an expression for the LST of $S_{1,i}$. To this end, we define g_{1n} and a new random variable $D_{1,n}$ as follows:

$g_{1n} = P\{$number of batches of type 2 which arrive during D_1 is $n\}$,

$P\{D_{1,n} \le t\} = P\{D_1 \le t \mid \text{n batches of type 2 have arrived during } D_1\}.$

Then, by conditioning on the number of individual passengers of type 2 which arrive during D_1, we can express the LST of S_{1i} as

$$S^*_{1,i}(\theta) = \sum_{n=0}^{m_2-i-1} g_{1n} \Big[\sum_{j=n}^{m_2-i-1} x_{2j}^{*(n)} \{D^*_{1,n}(\theta) U^*_{2,m_2-(i+j)}(\theta)\} + (1 - \sum_{j=n}^{m_2-i-1} x_{2j}^{*(n)}) D^*_{1,n}(\theta) \Big]$$

$$+ \ \{D^*_1(\theta) - \sum_{n=0}^{m_2-i-1} g_{1n} D^*_{1,n}(\theta)\}. \qquad (3.6)$$

The explanation of equation (3.6) is as follows:

Suppose the number of batches which arrive during D_1 is n. Note that this happens with probability g_{1n} and the LST of the travel time D_1 in this case is $D^*_{1,n}(\theta)$. Since the state at epoch 2 is given as i, if n is greater than, or equal to, m_2-i, the number of type 2 passengers at epoch 3 will reach or exceed the control limit m_2 and hence there will be no sojourn time for the shuttle at terminal 2. This is represented by the second line in equation (3.6). Similarly, suppose the number of batches which arrive during D_1 is less than m_2-i, i.e., n<m_2-i. In this case, if the number of individual passengers, j, who belong to these n batches is less than m_2-i-1, the number of passengers of type 2 at epoch 3 will be i+j, and the shuttle must stay at terminal 2 for $U_{2,m_2-(i+j)}$. On the other hand, if j is equal to or greater than m_2-i-1, since the state at epoch 3 will be m_2, the shuttle will not be delayed at terminal 2, which is represented by the first line in equation (3.6).

From equation (3.6), we see that $S^*_{1,i}(\theta)$ can be obtained if we can compute $D^*_{1,n}(\theta)$ and $U^*_{2,n}(\theta)$. We now explain how these terms can be computed. The following results, stated as lemma 3.1 and theorem 3.1 yield $D^{*}_{1,n}(\theta)$ and $U^{*}_{2,n}(\theta)$ respectively.

Lemma 3.1 The LST of $D_{1,n}$ is expressed by

$$D^{*}_{1,n}(\theta) = \frac{(-\lambda_2)^n}{g_{1n}n!} D_1^{*(n)}(\theta+\lambda_2) = \frac{D_1^{*(n)}(\theta+\lambda_2)}{D_1^{*(n)}(\lambda_2)}. \tag{3.7}$$

Proof Let $d_1(t)$ be a p.d.f. of D_1 and $d_{1n}(t)$ be a p.d.f. of $D_{1,n}$. By applying Bayes' formula, we can obtain the p.d.f. of $D_{1,n}$ as

$$d_{1n}(t) = \frac{1}{g_{1n}} \frac{e^{-\lambda_2 t}(\lambda_2 t)^n}{n!} d_1(t). \tag{3.8}$$

From equation (3.8) and using that $D_1^{*(n)}(\theta) = \int_0^\infty (-t)^n e^{-\theta t} d_1(t)\, dt$, the LST of $D_{1,n}$ is expressed by

$$D^{*}_{1,n}(\theta) = \int_0^\infty e^{-\theta t} d_{1n}(t)\, dt = \frac{(-\lambda_2)^n}{g_{1n}n!} D_1^{*(n)}(\theta+\lambda_2). \tag{3.9}$$

Since $g_{1n} = \frac{(-\lambda_2)^n}{n!} D_1^{*(n)}(\lambda_2)$, by substituting this into equation (3.9), we obtain lemma 3.1. ■

Theorem 3.1

$U^{*}_{i,n}(\theta)$ is expressed recursively as

$$U^{*}_{i,n}(\theta) = B^{*}_i(\theta) \sum_{j=1}^{n-1} x_{ij}\{U^{*}_{i,n-j}(\theta) - 1\} + B^{*}_i(\theta). \tag{3.10}$$

where $B_i^*(\theta) = \dfrac{\lambda_i}{\lambda_i+\theta}$.

Proof Let $Y_{i,n}$ be the number of batch arrivals of type i during the sojourn time of the shuttle at terminal i given that the state is m_i-n at the instant the shuttle arrives at terminal i. Then, $P\{Y_{i,n}=k\}$ can be expressed by

$$\begin{aligned} P\{Y_{i,n}=k\} &= \sum_{j=n}^{\infty} x_{ij} && \text{for } k=1, \\ &= \sum_{j=1}^{n-1} x_{ij} P\{Y_{i,n-j}=k-1\} && \text{for } k>1. \end{aligned} \tag{3.11}$$

Consequently, the p.g.f. of $Y_{i,n}$ can be expressed by

$$\begin{aligned} Y_{i,n}(z) &= \sum_{j=n}^{\infty} x_{ij} z + \sum_{k=2}^{n} \sum_{j=1}^{n-1} x_{ij} P\{Y_{i,n-j}=k-1\} z^k \\ &= \sum_{j=n}^{\infty} x_{ij} z + \sum_{j=1}^{n-1} x_{ij} z \sum_{k=2}^{n} P\{Y_{i,n-j}=k-1\} z^{k-1} \\ &= \sum_{j=n}^{\infty} x_{ij} z + \sum_{j=1}^{n-1} x_{ij} z Y_{i,n-j}(z) \\ &= z \sum_{j=1}^{n-1} x_{ij} \{Y_{i,n-j}(z) - 1\} + z. \end{aligned} \tag{3.12}$$

Let B_i be a random variable denoting the interarrival time of type i batch arrivals. Since B_i is assumed to follow an exponential distribution with mean $\dfrac{1}{\lambda_i}$, the LST of B_i is expressed by

$$B_i^*(\theta) = \frac{\lambda_i}{\lambda_i+\theta}, \quad i=1,2. \tag{3.13}$$

The result now follows by substituting $B_i^*(\theta)$ in place of z in equation (3.12).

Since we have an expression for $S^*_{1,i}(\theta)$, we can now obtain $q_4(i,j)$. Let h_{1in} denote the probability that the number of batches of type 1 which arrive during $S_{1,i}$ is n.

Then, in the same manner in which q_i was computed in equation (2.4), h_{1in} can be computed from (for $n=0,\ldots,m_1-j-1$),

$$h_{1in} = \frac{(-\lambda_1)^n}{n!} S_{1i}^{*(n)}(\lambda_1), \quad n=0,\ldots m_1-j-1, \tag{3.14}$$

where $S_{1i}^{*(k)}(\lambda_1)$ can be obtained as follows:

By taking the k^{th} derivative on both sides of equation (3.6) and setting $\theta=\lambda_1$, we have

$$S_{1i}^{*(k)}(\lambda_1) = \sum_{n=0}^{m_2-i-1} g_{1n} \Big[\sum_{j=n}^{m_2-i-1} x_{2j}^{*(n)} \{\sum_{r=0}^{k} \binom{k}{r} D_{1n}^{*(r)}(\lambda_1)\, U^{*(k-r)}_{2\ m_2-(i+j)}(\lambda_1)\}$$

$$+ (1- \sum_{j=n}^{m_2-i-1} x_{2j}^{*(n)})\, D_{1n}^{*(k)}(\lambda_1)\Big] + \{D_1^{*(k)}(\lambda_1) - \sum_{n=0}^{m_2-i-1} g_{1n}\, D_{1n}^{*(k)}(\lambda_1)\}, \tag{3.15}$$

where

$D_{1n}^{*(k)}(\lambda_1)$ can be obtained from equation (3.7) as

$$D_{1n}^{*(k)}(\lambda_1) = \frac{D_1^{*(n+k)}(\lambda_1+\lambda_2)}{D_1^{*(n)}(\lambda_2)}, \tag{3.16}$$

and $U_{2n}^{*(k)}(\lambda_1)$ is obtained from equation (3.10) as

$$U_{2n}^{*(k)}(\lambda_1) = \sum_{j=1}^{n-1} x_{2j}\{\sum_{r=0}^{k} \binom{k}{r} B_2^{*(r)}(\lambda_1)\, U^{*(k-r)}_{2\ n-j}(\lambda_1)\}$$

$$+ (1 - \sum_{j=1}^{n-1} x_{2j})\, B_2^{*(k)}(\lambda_1). \tag{3.17}$$

Since we know h_{1in}, we can finally obtain $q_4(i,j)$ as

$$q_4(i,j) = \sum_{k=0}^{j} h_{1ik}\, x_{1j}^{*(k)}, \qquad i=0,\ldots m_2, \quad j=0,\ldots,m_1-1,$$

$$q_4(i,m_1) = 1 - \sum_{j=0}^{m_1-1} q_4(i,j), \qquad i=0,\ldots\ldots m_2. \tag{3.18}$$

We have now computed the transition matrix Q_4. Since Q_2 can be computed in the same way as Q_4, we can obtain P_2 and P_4 using equation (3.3), and then, π_2 and π_4 using equation (3.4).

We are now in a position to find π_1 and π_3, and from them, W_1 and W_2. Since π_2 and W_2 can be obtained in the same way as π_1 and W_1, we will only present how π_1 and W_1 can be computed. As shown in equation (3.5), to compute π_1, we need to know Q_1, namely, each value of $q_1(i,j)$. Let f_{1j} be the probability that the number of individual passengers of type 1 who arrive during D_2 is j. Then, $q_1(i,j)$, can be computed as

$$\begin{aligned} q_1(i,j) &= 0 && \text{if } j<i \\ &= f_{1\,j-i} && \text{if } i\leq j,\ j<m_1 \\ &= 1 - \sum_{k=0}^{m_1-i-1} f_{1k} && \text{if } i\leq j,\ j=m_1. \end{aligned} \tag{3.19}$$

Since Q_1 has now been obtained, by applying equation (3.5), we can obtain π_1. Once π_1 has been obtained, the only other information needed to obtain $i_1^{(1)}$ and $i_1^{(2)}$ are the first two moments of the number of type 1 passengers at epoch 1. Note that these values can be obtained from equation (2.6) if we know the first two moments of the intervisit time, V_1. The first two moments of V_1 can be easily obtained because $V_1 = S_1+D_2$ and S_1 and D_2 are independent. The independence holds here because in this vacation, D_2

follows S_1. (Note, however, that D_2 affects the realization of the S_1 which follows D_2.) Thus, we have

$$E(V_1) = E(S_1) + E(D_2), \tag{3.20.a}$$

$$E(V_1^2) = E(S_1^2) + 2E(S_1)E(D_2) + E(D_2^2). \tag{3.20.b}$$

Since $E(D_2)$ and $E(D_2^2)$ are given values, in order to obtain $E(V_1)$ and $E(V_1^2)$, we should determine $E(S_1)$ and $E(S_1^2)$. This is done as described below. Since π_{2k} is the steady state probability that the state at epoch 2 is k, the LST of S_1 is expressed by

$$S_1^*(\theta) = \sum_{k=0}^{m_2} \pi_{2k} S_{1,k}^*(\theta). \tag{3.21}$$

By differentiating $S_1^*(\theta)$ with respect to θ, we obtain

$$E(S_1) = - \sum_{k=0}^{m_2} \pi_{2k} S_{1k}^{*(1)}(0), \tag{3.22.a}$$

$$E(S_1^2) = \sum_{k=0}^{m_2} \pi_{2k} S_{1k}^{*(2)}(0), \tag{3.22.b}$$

where $S_{1k}^{*(1)}(0)$ and $S_{1k}^{*(2)}(0)$ can be obtained using equation (3.15).

Since we now know the first two moments of the intervisit time as well as π_{1k} for $k=0,\dots,m_1-1$, we can calculate $i_1^{(1)}$ and $i_1^{(2)}$ using the procedure presented in section 2, and hence, can obtain W_1 using equation (2.1).

Finally, let us analyze the expected length of a cycle when m_1 and m_2 are used as control limits. Obviously, the expected length of a cycle, χ, can be obtained by

$$\chi = E(V_1) + E(U_1) = E(V_2) + E(U_2). \tag{3.23}$$

However, since the expected number of type i passengers who arrive during one cycle is $i_i^{(1)}$ and the arrival rate is $\lambda_i x_i^{(1)}$, we can obtain the expected length of a cycle more simply from

$$\chi = \frac{i_1^{(1)}}{\lambda_1 x_1^{(1)}} = \frac{i_2^{(1)}}{\lambda_2 x_2^{(1)}} . \qquad (3.24)$$

4. The Optimization Procedure

The motivation for the analysis in this paper comes from studies of the optimal control policies when a cost structure is imposed on our two-terminal shuttle system. Suppose a linear holding cost of c_i per unit time is incurred for each passenger of type i (i=1,2) who waits at terminal i and a cost c_s is incurred each time the shuttle leaves a terminal and returns, again, to that terminal. Since the expected length of a cycle is χ, the expected operating cost for the shuttle per unit time is expressed as $\frac{c_s}{\chi}$. Similarly, from Little's rule, the mean number of passengers of type i waiting at terminal i at an arbitrary time is $\lambda_i x_i^{(1)} W_i$. Thus, the expected waiting cost per unit time is $c_1 \lambda_1 x_1^{(1)} W_1 + c_2 \lambda_2 x_2^{(1)} W_2$. From this, the expected total cost per unit time, $\boldsymbol{C}$, is calculated as

$$\boldsymbol{C} = \frac{\lambda_1 x_1^{(1)} c_s}{i_i^{(1)}} + c_1 \lambda_1 x_1^{(1)} W_1 + c_2 \lambda_2 x_2^{(1)} W_2. \qquad (4.1)$$

This expression enables us to search for the optimal control values which minimize the expected cost per unit time in the long run.

Due to the complex form of the equations, we could not prove any properties (such as convexity) for the cost function which could be useful in the search procedure. However, our computational experience suggests that for a fixed value of one control variable, the cost function is unimodal with respect to the other control variable. If we make use of this observation and the recursive nature of the functions, we can find the optimal control values with reasonable computational effort. In particular, we observed that, for a symmetric system(in the sense that two queues are same in every respect), the cost function shows a unimodality with respect to the control value m_i. The following example illustrates this:

<example>

$\lambda_1=\lambda_2=0.2$,

$x_{11}=x_{21}=0.4$, $x_{12}=x_{22}=0.3$, $x_{13}=x_{23}=0.3$,

$D_1=D_2=10$, with probability 1,

$c_s=500$,

$c_1=c_2=1$.

The results of the policy comparisons for the example are presented in table 1. In the table, the values of χ, W_i and $\boldsymbol{C}$ are shown for each value of m_i. Note that in this example, $\boldsymbol{C}$, which is the expected total cost incurred per unit time, shows a unimodality with respect to m_i as was stated.

Table 1.

Result of the Example

m_i	χ	W_i	C
11	36.105	14.307	24.722
12	38.597	15.083	24.418
13	41.129	15.874	24.222
14	43.687	16.676	24.119
15*	46.264	17.484	24.095
16	48.855	18.297	24.140
17	51.455	19.113	24.243
18	54.062	19.931	24.397
19	56.674	20.751	24.593

(* means the optimal policy)

5. Conclusions

In this paper, a control limit policy is presented for the shuttle dispatching problem with two terminals. Under the assumptions that controls are made at both terminals and the number of passengers is known only when the shuttle is staying at that terminal, a procedure to calculate the mean waiting time for each type of passenger is presented for given control values. This procedure can be directly extended to the system with more than two terminals. However, since the number of states needed to analyze the system increases with number of terminals, the procedure developed in this paper is not practicable if the number of terminals is greater than two. Thus, for this case, good approximation methods are desired to be developed.

References

Cooper, R.B., *Introduction to Queueing Theory*, Second Ed., Elsevier, North-Holland, New York, (1981).

Deb, R.K., "Optimal Dispatching of a Finite Capacity Shuttle," *Management Science,* v 24 (1978), pp. 1362-1372.

Deb, R.K. and Serfozo, R.F., "Optimal Control of Batch Service Queues," *Advances Appl. Probability,*, v 5 (1973), pp. 340-361.

Ignall, E and Kolesar, P., "Optimal Dispatching of an Infinite Capacity Shuttle : Control at a single terminal," *Operations Research*, v 22 (1974), pp. 1008-1025.

Lee, H-S and Srinivasan, M.M., " Control Policies for the $M^x/G/1$ Queueing System," Technical Report 87-20 (1987), Dept. of Industrial and Operations Engr.,The Univ. of Michigan, summitted to Management Science.

Powell, W.B., "Analysis of Vehicle Holding and Canellation Strategies in Bulk Arrival, Bulk Service Queues," *Transportation Science,* v 19 (1985), pp 352-377.

Powell, W.B., "Iterative Algorithms for Bulk Arrivals, Bulk Service Queues with Poisson and non-Poisson Arrivals," *Transportation Science,* v 20 (1986), pp.65-79.

Sim, S.H. and Templeton, J.G.C., "Computational Procedures for Steady State Characteristics of Unscheduled Multi-Carrier Shuttle Systems", *European Journal of Operations Research*, v 12 (1983), pp. 190-202.

Teghem, J., Jr., "Control of the Service Process in a Queueing System," *European Journal of Operations Research*, v 23 (1986), pp. 141-158.

Weiss, J., "The Computation of Optimal Control Limits for a Queue with Batch Services," *Management Science* v 25 (1979), pp. 320-328.

_______,"Further Results on an Infinite Capacity Shuttle with Control at a Single Terminal," *Operations Research*, v 29 (1981), pp.1212-1217.