



Transportation Research Forum

A Simplified Method for Performance Evaluation of Public Transit Under Reneging Behavior of Passengers

Author(s): Md Kamrul Islam, Upali Vandebona, Vinayak Dixit, and Ashish Sharma

Source: *Journal of the Transportation Research Forum*, Vol. 54, No. 3 (Fall 2015), pp. 23-42

Published by: Transportation Research Forum

Stable URL: <http://www.trforum.org/journal>

The Transportation Research Forum, founded in 1958, is an independent, nonprofit organization of transportation professionals who conduct, use, and benefit from research. Its purpose is to provide an impartial meeting ground for carriers, shippers, government officials, consultants, university researchers, suppliers, and others seeking exchange of information and ideas related to both passenger and freight transportation. More information on the Transportation Research Forum can be found on the Web at www.trforum.org.

Disclaimer: The facts, opinions, and conclusions set forth in this article contained herein are those of the author(s) and quotations should be so attributed. They do not necessarily represent the views and opinions of the Transportation Research Forum (TRF), nor can TRF assume any responsibility for the accuracy or validity of any of the information contained herein.

A Simplified Method for Performance Evaluation of Public Transit Under Reneging Behavior of Passengers

by Md. Kamrul Islam, Upali Vandebona, Vinayak V. Dixit and Ashish Sharma

This paper develops a model based on the Markov Chain technique to evaluate performance of a public transport route. The model addresses a special situation where a passenger left behind by a bus leaves the system without any further waiting to make an alternative travel arrangement. Such renegeing behavior is indicative of an infinite penalty associated with further waiting from a passenger viewpoint. Apart from the theoretical derivations for the various attributes of interest, numerical examples to analyze the system performance (such as expected number of passengers served, expected number of abandoned passengers, and expected amount of unused space on the transit system) are presented. This provides insights for optimum selection of fleet size and size of vehicles

INTRODUCTION

Two basic problems often faced by analysts of transportation systems are related to estimation of vehicle size and frequency of service. In the case of public transit systems, the use of smaller buses with a relatively high service frequency lowers the average waiting time and increases operating speed, but is not suitable for high passenger demand conditions as it costs more to operate per seat provided. On the other hand, comparatively larger buses are usually associated with lower operating cost to operators, but lead to low service frequencies and long average waiting time for passengers. From the operator perspective, it is desirable to use large vehicles that maximize productivity. From the passengers' perspective, the frequency of service is the matter of concern. This poses a dilemma to transit designers in selection of service configuration to meet user needs and desired service levels in terms of service frequency.

The analysis in this paper offers insights to the problem faced by transit system designers; namely with regard to fleet size as well as what should be the size of the vehicles that should be part of the fleet. To address this, a stochastic model using Markov Chain Technique is formulated for a bus transit system with multiple stops, where carriers with a regular headway serve all waiting passengers under a capacity constraint. Markov Chain is a method used to model sequential events of bus operation at a stop where randomly arrived passengers wait for a bus and board the bus depending on space availability after alighting of passengers. This paper sheds light on the performance of transit using the following metrics: number of passengers served by the system, number of passengers that were unable to use the service because of space unavailability, and number of unused space throughout the transit operations.

LITERATURE REVIEW

There were a number of studies performed on bus size and frequency related to urban public transport systems with the aim of improving system performance and enhancing efficiency.

Jansson (1980) paid significant attention to operating cost and user cost while arguing that previous models underestimated user cost and overestimated operator cost. He proposed a model that minimizes total social cost, which includes operator cost, passenger waiting time, and travel

time during peak capacity conditions. He concluded that the vehicle size required to minimize social cost is smaller than the vehicle size found in practice, where the number of vehicles is set for a given vehicle size to achieve an average occupancy rate (the mean occupancy rate is the ratio of the mean passenger flow to the product of the service frequency and the bus size). Oldfield and Bly (1988) provided an analytical model to determine optimal bus size, which considered elastic demand that is affected by trip cost. They showed that the optimal size varies with the square root of demand and with the unit cost to the power of 0.1 to 0.2. In a case study in the United Kingdom for typical urban operating conditions, they found that the optimal bus size lies between 55 and 65 seats for a monopoly service.

Jansson (1993) included the variability of value of time to optimize the vehicle size, frequency, and journey price simultaneously. However, because it required large amounts of data, this model was found to be difficult to implement compared with other models. Lee et al. (1995) developed a model that is able to find the bus size for different periods of day in addition to optimal bus size for each route. They also attempted to find suitable conditions to use one bus or alternative uses for a mixed-size fleet. They determined the threshold ratio of peak demand to off-peak demand for multiple-route operation is 1.92 and showed that mixed-fleet operation is preferable on multiple route operation in case of high variation in demand between peak and off-peak period. Rietveld et al. (2001) have extended Mohring's (1972, 1976) basic "square root model" for frequencies and derived a general formulation under an alternative regime of welfare and profit optimization for frequency, vehicle size, and cost of building a railway system. It was observed that in rail transport, the average occupancy rates were low. Chien (2005), with the objective to minimize total cost, developed a methodology integrating both analytical and numerical techniques to optimize headway, vehicle size and passenger route choice for a feeder bus service. The methodology is then applied to analyze a non-stop feeder bus service connecting a selected rail station and a recreation center (Sandy Hook, NJ). It was shown that the optimal fleet size is a function of the demand multipliers. If the demand multiplier is less than 0.7, the optimal fleet size is three buses. The optimal fleet size is four and five buses if the demand multiplier is between 0.7 and 1.1, and greater than 1.1 but less than 1.6, respectively. Dell'Olio et al. (2008) have presented a model to solve the problem of optimizing frequencies and bus size on a transit network consisting of 15 routes. They claimed that their model can designate different types of buses on each route taking into account the reciprocal influence of each route in addition to optimization of the capacity of the bus. However, their model differs from the more commonly held idea that smaller buses are more profitable in a majority of cases.

These earlier studies did not consider the situation where passengers are unwilling to wait for the next bus when they have been unable to board an earlier bus. The current study considers such situations, where buses are full, the transit operators lose a proportion of passengers. This deteriorates the level of service of the transit system.

For the objective of estimating the expected number of passengers served by transit systems, allowing for lost passengers, a queuing theory technique known as the Markov Chain method has been applied. In queuing theory, serving more than one customer at a time is a case known as bulk service. The service capacity is referred to as variable capacity, as it is not the same at all instances of service. In a transportation system analogy, customers are passengers and servers are vehicles, respectively. Bulk service for a single station was first addressed by Bailey (1954) and Downton (1955) applying imbedded Markov Chain technique for a transportation system when the average number of passengers that can be served was constant. Jaiswal (1961) extended Bailey's model for a series of stations where number of passengers served were not constant, rather based on the number of passengers waiting at a stop and vehicle capacity. Arguing that these models are very complex to solve and numerical answers to these models are difficult to find, Giffin (1966) developed a model for a transit system with a series of stations where the simplification was made in assuming there are no en-route departures of customers and all customers remain onboard the vehicle until it reaches the terminal. A. Grosfeld-Nir and Bookbinder (1995) and Lam et al. (2009) considered departure of

passengers from buses at on-route destinations but did not link with the performance measures of transit systems. This paper demonstrates how a Markov Chain technique can be utilized to calculate numerical answers to a model bulk service transit system of multiple stops served by a fixed-size vehicle fleet.

MODELING FRAMEWORK

This section presents a description of operation of a bus route and the associated notations along with the modeling framework used to evaluate performances of bus transit systems.

Description of A Bus Route Operation

The operation of bus transit in a route that contains multiple stops is illustrated with the aid of the schematic diagram shown in Figure 1 (a) and Figure 1 (b). From the point of view of bus operation, empty buses are dispatched from the dispatch station and travel along the route allowing passengers to alight and board at stops. This process of serving, boarding, and alighting passengers continues stop after stop. The bus movement is presented as a trajectory diagram in Figure 1(a). The vertical axis shows the distance travelled by buses along the route stopping at designated bus stops, and the horizontal axis indicates the time. Inclined lines between two stops show bus travel from one stop to the next stop. Inclined lines between two stops show bus travel from one stop to the next stop. The dotted lines between two inclined lines show the time spent at stops allowing mainly for boarding and alighting of passengers. As shown in Figure 1 (b), there are five events related to passengers. These events are (a) passengers arrive at a bus stop, (b) passengers board a bus upon arrival if space is available, (c) passengers wait for a bus, (d) passengers leave a stop if space is not available, and (e) a bus departs with passengers for next stop.

Figure 1 (a): Trajectory of Buses Along a Route for Regular Headway of Buses

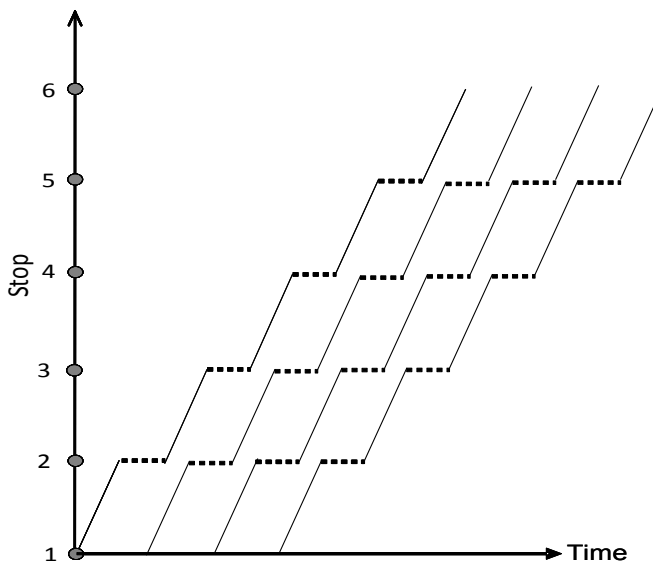
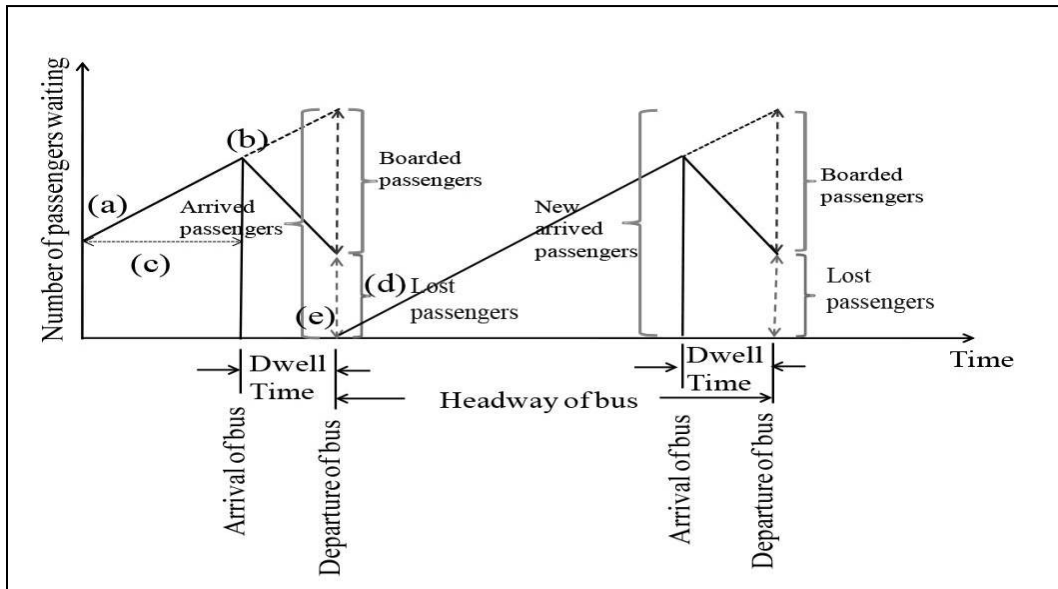


Figure 1 (b): Schematic Presentation of Dynamic Interaction Between Bus and Passengers at Stops



Assumptions

The following assumptions are made in the model formulated to describe the bus operation in Figure 1(a).

a) Operational:

- Bus size: Buses dispatched in this service are assumed be of the same size.
- Service process: Buses serve passengers up to their passenger capacities after passengers alight from bus at given stop. If there is no passenger boarding at a stop, buses skip the stop. It should be noted that no time stop is considered here.
- Capacity of stop: The capacity of the bus stop to accumulate arriving passengers is assumed to be infinite.
- Time table: To simplify the model, buses are assumed to arrive at stops according to a fixed timetable and early or late arrivals of buses are not allowed.
- Headway of bus: Bus headways are assumed to be less than 12 minutes to justify the adoption of random arrival of passengers at stops. Evidences from several empirical studies demonstrated that this assumption is reasonable. Jolliffe and Hutchinson (1975) provided a behavioral explanation of the association between bus and passenger arrivals at a bus stop. They presented passenger arrival pattern in three categories: (i) proportion of passengers (q) arrive coincidentally with the bus (see and run to stop), (ii) proportion of passengers (1-q) who arrive at stop at optimal time based on published timetable and experience, and (iii) proportion of passengers (1-q)(1-p) who arrive randomly. Bowman and Turnquist (1981) used the term “unaware” of schedule for the passengers who arrive at stops randomly. Moreover, they reported that 12 to 13 min. headways is transition from random to coordinated passenger arrivals at stops as concluded by Okrent (1971). O’Flaherty and Mangan (1970) also suggested 12 min. in Leeds as the transition. Furthermore, Seddon and Day (1974) showed by empirical research that passengers arrive at stops randomly for headway at less than 10-12 minutes. Fan and Machemehl (2009) identified 11-minute vehicle headway as the transition from random passenger arrivals to non-random arrival.

- Number of alighting passengers: No passenger disembarks from a bus at the first stop, as the bus arrives empty at the first stop. At other stops, except the last one, the number of alighting passengers depends on the arrival occupancy. Passenger alighting probability is assumed to be the same during the period of study. Passenger behavior is assumed to be independent of each other. Thus, the number of alighting passengers at each stop is assumed to follow a binomial distribution as suggested by Andersson and Scalia-Tomba (1981). At the last stop, all on-board passengers alight to ensure the bus is empty.
- Alighting and boarding process: It is assumed that boarding of passengers on a bus starts after the completion of alighting of passengers. A fraction of onboard passengers alight upon arrival at a stop according to the above assumption. Also, a bus picks up all passengers while adhering to a capacity constraint.
- Boarding and alighting times: For simplicity, boarding and alighting times are assumed to be negligible.
- Travel time: Travel time between stops is assumed constant and remains unchanged during the operation period.

b) Demand :

- Arrival of passengers: It is assumed that passengers arrive randomly at a stop according to a Poisson process. The number of passengers waiting is a function of passenger arrival rate at a stop and the time interval between two consecutive arrivals of buses. Furthermore, passenger arrival at one stop is independent of arrivals at any other stop. It is assumed that passenger demand does not change over the period of interest.
- Passenger waiting behavior: It is assumed that an alternative mode of transportation is available for “impatient passenger.” An impatient passenger is defined as a person reluctant to wait for another bus when he/she is denied entry to the first bus to arrive due to inadequacy in space. Hence, passengers rejected to be served by the next bus are considered lost from the system.

This kind of model can be an approximation for transit operators serving suburbs where the population may have a low tolerance to waiting time. For example, passengers who look for an alternative mode of transportation to be on time at work during rush hour (i.e., 6.00 am to 9.00 am) can be considered impatient passengers. Also, this analysis can be viewed as a particular situation of a large system where there is substantial one-way traffic from suburban areas to the city center. Another scenario where the analysis applies is when passengers wait for a bus on a street just outside their homes (where they have a personal auto immediately available). If the bus has no capacity to accept passengers, then they would travel by their autos.

Notations: The following notations are adopted in the formulations

Table 1: Notations

N	Total number of stops
V	Number of buses provided per hour
n	Index number of stops to be served
λ_n	Passenger arrival rate at stop n
d_n	Alighting proportion of passengers at stop n
H_n	Headway of bus at stop (i.e., Time between two consecutive buses at stop n)
x_n	Number of passengers on bus leaving stop n
P_{jn}	Probability of j number of passengers on a bus leaving stop n
L_{e_n}	Distance between two stops n and $n+1$
C	Bus size in terms of seats and standing passengers allowed
g_{kn}	Probability that k passengers arrive at stop n in the time interval of H_n and given by $\frac{e^{-\lambda_n H_n} (\lambda_n H_n)^k}{k!}$

Mathematical Formulation for the First Stop

Important performance measures for a bus transit system can be developed from the probability distribution of space available on buses as they progress along a fixed route. This probability distribution can be obtained by carefully monitoring the four events mentioned in the previous section.

Consider P_n as the probability vector of number of onboard passengers on a bus leaving stop n and p_{in} indicates each element of the vector where value of i ranges from 0 to C . This means the probability vector P_n has $C+1$ element, which can be represented as $(p_{0n}, p_{1n}, p_{2n}, \dots, p_{Cn})$.

The first stop, the first entry of the probability vector, p_{01} , indicates the probability of no onboard passenger when the bus leaves stop 1. For example, in this particular case of stop 1, when no passenger arrives between the two consecutive buses, there will be no passengers waiting, and therefore no passengers will be onboard after stop 1. Thus, the probability of no passenger arriving at the stop, (or probability of no onboard passenger in the bus departing the first stop) can be derived as $g_{01} = e^{-\lambda_1 H_1} \frac{(\lambda_1 H_1)^0}{0!} = e^{-\lambda_1 H_1}$. Other entries $p_{11}, p_{21}, \dots, p_{C-1,1}$ can be found in a similar way by varying the value of k from 1 to $C-1$ in the expression $e^{-\lambda_1 H_1} \frac{(\lambda_1 H_1)^k}{k!}$ respectively. The last entry, p_{C1} , can be found by simply applying total law of probability as $1 - \sum_{i=0}^{C-1} e^{-\lambda_1 H_1} \frac{(\lambda_1 H_1)^i}{i!}$ or by summing the probability of passenger arrivals as $\sum_{i=C}^{\infty} e^{-\lambda_1 H_1} \frac{(\lambda_1 H_1)^i}{i!}$, since no passengers alight at the first stop, the probability vector of the number of onboard passengers leaving the first stop p_1 can be found as follows:

$$(1) \quad P_i^1 = \begin{cases} \frac{e^{-\lambda_1 H_1} (\lambda_1 H_1)^i}{i!} & \text{for } 0 \leq i < C \\ 1 - \sum_{j=0}^{C-1} \frac{e^{-\lambda_1 H_1} (\lambda_1 H_1)^j}{j!} & \text{for } i = C \end{cases}$$

Since passengers can alight at stops other than the first, the method to find the entries of the probability vector of onboard passengers on the buses at other stops $p_{01}, p_{11}, p_{21}, \dots, p_{C1}$ is different from above and is discussed in the next section.

Mathematical Formulation for the Second and Subsequent Stops

When a bus approaches the second stop, the number of available spaces in the bus is reduced by the number of passengers picked up at the first stop. However, the number of available spaces on the bus at the second stop can increase if one (or more) onboard passengers alight at the second stop. Then the waiting passengers at the second stop are allowed to board the bus until the bus is full. Passengers left behind in the event of inadequate capacity are considered lost from the system, as mentioned earlier. Since passenger arrivals at stops and alighting are random, passenger boarding and alighting at a stop can be modeled as a stochastic process. Thus, there are two sets of probability arrays required to describe the algebra related to transit route operation:

- i. A probability vector representing the number of onboard passengers at a stop after the alighting process has completed.
- ii. A probability vector representing the number of onboard passengers at a stop after the boarding process is completed.

Details of these vectors are described below:

Probability vector representing the number of onboard passengers at a stop after the alighting process has completed upon arrival at the second stop. If there are i passengers on board a bus approaching the second stop, the state of the process is defined as E_i . Here, i could be any value between 0 to C , i.e., there will be a possibility that there are no onboard passengers, only one passenger on board, two passengers on board, three passengers on board, and up to a maximum of C number of passengers on board the approaching bus. If any of the passengers alight at the second stop and there are still some of the j passengers remaining on board, the state E_i will be changed to E_j . In other words, the process makes a transition from state E_i to state E_j . This transition can be represented by a “Transition Probability Matrix,” which is described further.

Now, let us consider, A_n is a one-step transition matrix related with alighting of passengers on a bus arriving at stop n and $a_{ij n}$ is the cell entry for row i , column j of the matrix A_n . The $a_{ij n}$ represents the conditional probability that j passengers will remain on board after $(i-j)$ number of passengers alight at stop n . In other words, the conditional probability that j number of passengers will remain on board after alighting at stop n given that the bus approached from stop $(n-1)$ to stop n with i onboard passengers. These quantities can be denoted as A_2 and $a_{ij 2}$ for the second stop. In the event of a bus approaching the second stop empty, i.e., there are no onboard passengers approaching the second stop, probability of alighting zero passengers is one. Thus, the first entries in the first row of matrix A^2 , $a_{00 2}$ is 1 and 0 for other entries in the first row.

If $(i-j)$ passengers alight at the second stop, i onboard passengers will be reduced from i to j . If the fraction of onboard passengers alighting at the second stop is d_2 , then $(d_2)^{i-j}$ is the probability of $(i-j)$ passengers alighting at the second stop and $(1 - d_2)^{i-j} = (1 - d_2)^j$ is the probability of not alighting $(i-j)$ passenger at the stop. According to the principle of binomial distribution, $\binom{i}{i-j} (d_2)^{i-j} (1 - d_2)^j$

is the probability of alighting ($i-j$) passengers from i number of onboard passengers at stop 2. Then, for $i \geq j$, the entries for a_{ij2} are $\binom{i}{i-j} (d_2)^{i-j} (1-d_2)^j$. Since, a transition from i onboard passengers to some number greater than j is not possible; thus, for $i < j$, the entries for a_{ij2} are zero.

Let us consider P'_2 is the probability vector of onboard passengers after the alighting process has completed at the second stop, then P'_2 can be readily obtained by multiplying P_1 (the probability vector of onboard passenger departing the first stop) by A_1 (transition probability matrix of passenger alighting at the second stop), i.e., $P'_2 = P_1 A_1$.

Probability vector representing the number of onboard passengers at a stop after the boarding process is completed: When the alighting process at the second stop is finished, if space is available, the bus is allowed to pick up waiting passengers. If, after the alighting process, there are k passengers remaining on board, the process is said to be in state E_k . Here, k could be any value between zero to C , i.e., there is also a possibility of having onboard passengers anywhere from zero to a maximum of C that remained on board the bus. Additional boarding of passengers transforms the state from E_k to E_l . Therefore, another transition probability matrix is defined whose elements are the conditional probability that $(l-k)$ number of waiting passengers at the second stop are allowed to board the bus given that the bus already has k onboard passengers after the alighting process has finished. Now, consider B_n as a one step transition probability matrix associated with waiting passengers boarding at bus at stop n ($n \geq 2$) and $b_{kl n}$ as a conditional probability that $(l-k)$ new passenger will board the bus given that it has already k passengers on board after the alighting process has finished. This is the cell entry for row k , column l of the matrix B_n .

In general, k passengers on board the bus will go to l , if $(l-k)$ new passengers board the bus at the second stop. If λ_2 is the arrival rate of passengers at the second stop, then the probability of arrival of $(l-k)$ passengers is $g_{(l-k)2} = e^{-\lambda_2 H_2} \cdot \frac{(\lambda_2 H_2)^{(l-k)}}{(l-k)!}$. Hence, the entries of the matrix B^2 for each bus is given by,

$$(2) \quad b_{kl2} = \begin{cases} e^{-\lambda_2 H_2} \frac{(\lambda_2 H_2)^{(l-k)}}{(l-k)!}, & \text{for } 0 \leq l < C \\ 1 - \sum_{i=0}^{C-1} e^{-\lambda_2 H_2} \frac{(\lambda_2 H_2)^{(i-k)}}{(i-k)!} & \text{for } l = C \end{cases}$$

In addition, for a bus the transition from k onboard passengers to a number lesser than k is clearly impossible. For this reason, $b_{kl2} = 0$ for $l < k$.

The probability vector of onboard passengers in a bus leaving the second stop (denoted by p_2) can be readily obtained by multiplying p_2 (the probability vector of onboard passengers after completion of passenger alighting at the second stop) and B_2 (the transition probability matrix of passengers boarding at the second stop), i.e., $P_2 = P'_2 B_2 = P_1 A_2 B_2$.

Since the system is Markovian, using the Chapman Kolmogorov Equations (Ross 2003), the probability vector for the number of onboard passengers in the bus at any stop n , P_n for $n \geq 1$, can be found as $p_n = p_{n-1} A_n B_n$. In such an approach, the probability statements regarding the onboard passengers leaving a stop can be derived.

PERFORMANCE MEASURES

Once it is possible to calculate the probability vector of onboard passengers leaving a stop, then it is easy to calculate the number of passengers served, number of lost passengers, the load factor, and the unused space on a bus along a route. These performance measures are derived below.

Expected Number of Passengers Served Along a Transit Route

The revenue earned by transit operators mostly comes from the number of passengers. Therefore, the expected number of passengers served along a complete transit route is an important performance measure. To evaluate this measure, the expected number of passengers remaining on a bus at a stop after alighting of others can be found as $\sum_{i=0}^C x_{in} P_{i,n-1} A_n$. x^n is the number of onboard passengers on a bus leaving stop n . It ranges from zero to C . Since passengers are allowed to board the bus after the alighting process has ended and the bus leaves the stop, the expected number of passengers on a bus leaving a stop is $\sum_{i=0}^C x_{in} P_{i,n-1} A_n B_n$. Hence, it is possible to find the expected number of passengers served at a stop from the difference between these two quantities, i.e., expected number of passengers served per bus at stop n ,

$$(3) \quad E(S_n) = \left(\sum_{i=0}^C x_{in} P_{i,n-1} A_n B_n - \sum_{i=0}^C x_{in} P_{i,n-1} A_n \right) = \left(\sum_{i=0}^C x_{in} P_{in} - \sum_{i=0}^C x_{in} P_{in}' \right)$$

This expression gives the expected number of passenger served at an individual stop. The number of total served passengers (TSP) per bus along a transit route can be calculated by summing the number of passengers served at each individual stop as $\sum_{k=1}^N E(S_k)$.

Expected Number of Abandoned Passengers Along a Transit Route

For any given configuration of transit systems, the number of passengers unable to board a bus due to shortage of spaces is another measure of effectiveness. On average, the system will be in equilibrium in that the expected number of passengers arriving for service by each bus at a stop will be the sum of the number of passengers served at a stop and the expected number of passengers turned away from a stop. i.e.,

Expected number of passengers requesting service at a stop = Expected number of passengers served at a stop + Expected number of abandoned passengers at a stop.

(4) Expected number of abandoned passenger per bus at a stop,

$$E(Ab_n) = \lambda_n H_n - E(S_n)$$

This expression provides the expected number of abandoned passengers at an individual stop among the system of N total stops. The number of total abandoned passengers (TAP) per bus along a transit route can be calculated by summing the number of passengers lost at each individual stop as

$$\sum_{k=1}^N E(Ab_k).$$

Number of Total Unused Capacity in Bus Along a Transit Route

When service demand is low compared with the supply, some bus spaces remain unused during the service period, and transit operators can calculate the amount of unused capacity in buses defined as the capacity remaining after the passenger boarding process is finished and buses begin leaving a stop. These unused spaces can be calculated by subtracting the number of passengers onboard from the bus size, i.e.

$$(5) E(U_n) = C - E(S_n)$$

In the case of a bus fully congested, no unused spaces are left, i.e., number of unused spaces on bus is zero. The number of total unused capacity (TUC) per bus along the complete route can be found by summing the expected unused spaces in buses at each stop as $\sum_{k=1}^{N-1} E(U_k)$.

Moreover, the variance of these measures can be calculated at each stop as follows:

(i) Variance of served passengers per bus at stop n

$$(6) \text{var}(s_n) = \left(\sum_{i=0}^C x_{in}^2 p_{in} - \left(\sum_{i=0}^C x_{in} p_{in} \right)^2 - \sum_{i=0}^C x_{in}^2 p'_{in} + \left(\sum_{i=0}^C x_{in} p'_{in} \right)^2 \right)$$

(ii) Variance of abandoned passengers per bus at a stop:

$$(7) \text{var}(L_n) = \lambda_n H_n - \text{var}(s_n)$$

Work Utilization Coefficient

Another performance indicator namely “Work Utilization Coefficient (WUC)” is adopted from Vuchic (2005) as the output or quantity of offered or utilized service on a transit line expressed as “*transportation work W*”. When all buses of a fleet run on entire lengths of routes, *offered work* W_0 , expressed in space-kilometer during one hour is:

$$(8) W_0 = V \times C \times \sum_{m=1}^{N-1} L_{e_m}; \quad (W_0 \text{ in spaces} - km/h)$$

Passenger-kilometer traveled along the transit route is called *Utilized work* W_u . The procedure to calculate this value is shown below.

Let us consider a matrix, “passengers-kilometer matrix,” denoted by PK of dimension $(m \times n)$ is constructed, where m is the origin stop of a passenger and n is destination stop and pk_{mn} is the element of passengers-kilometer matrix indicates the value of passenger-distance travelled from origin stop m to destination stop n where $m = 1, 2, \dots, N$ and $n = 1, 2, 3, \dots, N$. The matrix PK can be presented as follows:

$$(9) PK = [pk_{mn}] = \begin{cases} 0 & \text{for } m \geq n \\ E(S_m)d_{m+1}L_{e_m} & \text{for } n = m+1 \text{ and } n \neq N \\ E(S_m)d_n(n-m)L_{e_m} \prod_{r=m+1}^{n-1} (1-d_r) & \text{for } m+1 < n < N \end{cases}$$

Formulation of each element of the matrix PK is explained here. In general, no passengers alight at the same stop they board the bus. This formulation neglects those passengers who board in error and cannot get down at the same stop. Thus, the entries in the main diagonal of the matrix PK are zero, i.e, $pk_{mn} = 0$; for $m = n$. Furthermore, since it is obviously impossible that the numerical value of a destination stop n less than the origin stop m , the entries below the main diagonal are zero. i.e, $pk_{mn} = 0$; for $m > n$.

In the event of passengers alighting at a stop immediately after their origin stop, the passenger-distance travelled can be found straightforwardly by multiplying the distance between two stops (L_e) by number of passengers alighting at the stop. Number of passengers alighting can be calculated by multiplying the onboard passenger in a bus leaving the origin stop $E(S_m)$ by the alighting proportion at the destination stop d_n . Thus, for $n= m+1$ and $n \neq N$, the entries of the matrix PK are $E(S_m)d_{m+1}L_{e_m}$ for $n= m+1$ and $n \neq N$.

In the case of a passenger alighting at a stop other than that immediately following the origin stop, the entries are $E(S_m)d_n(n-m)L_{e_m} \prod_{r=m+1}^{n-1} (1-d_r)$ for $m + 1 < n < N$. Then, total passenger-kilometers (TPK) can be found by summing all the elements of the PK matrix as follows:

$$(10) \quad TPK = \sum_{m=1}^N \sum_n^N pk_{mn} = \sum_{m=1}^{N-1} E(S_m)d_{m+1}L_{e_m} + \sum_{m=1}^{N-1} \sum_{n=m+2}^N E(S_m)d_n(n-m)L_{e_m} \prod_{r=m+1}^{n-1} (1-d_r)$$

Now, WUC can be defined as the ratio of *utilized work* to the *offered work* as per Vuchic (2005) as:

$$(11) \quad WUC = \frac{W_u}{W_o} = \frac{TPK}{V \times C \times \sum_{m=1}^{N-1} L_m} \quad (WUC \text{ in passenger - km / sp - km})$$

Ratio of Lost Work to Demanded Work (Lw/Dw)

A performance measure, the ‘‘Ratio of Lost work to Demanded work (L_w/D_w)’’ (Islam et al. 2014), is used to investigate the effect of bus size and frequency of service on the performance of a transit system. Here, ‘‘Demanded Work (D_w)’’ is defined as passenger-km that is demanded by passengers to travel along a route based on their origin and destination. However, some passengers may leave the system in case of space unavailability on the bus without further waiting. The passenger-km anticipated by abandoned passengers is termed as ‘‘Lost Work (L_w).’’ Thus, the ratio of lost work to demanded work (L_w/D_w) reflects the amount of lost transportation work in relation to demanded transportation work that awaits for service in the system. Mathematically, the demanded work can be found by replacing $E(S_m)$ by $E(\lambda_m H_m)$ (i.e., number of arrivals at stop m between two consecutive buses) in Equation 12 as follows:

$$(12) \quad D_w = \sum_{m=1}^{N-1} E(\lambda_m H_m)d_{m+1}L_{e_m} + \sum_{m=1}^{N-1} \sum_{n=m+2}^N E(\lambda_m H_m)d_n(n-m)L_{e_m} \prod_{r=m+1}^{n-1} (1-d_r)$$

Similarly, the lost work (L_w) can be found by replacing $E(S_m)$ by $E(\lambda_m H_m)$ (i.e., number of abandoned passengers at stop m per bus) in Equation 13 as follows:

$$(13) \quad L_w = \sum_{m=1}^{N-1} E(Ab_m)d_{m+1}L_{e_m} + \sum_{m=1}^{N-1} \sum_{n=m+2}^N E(Ab_m)d_n(n-m)L_{e_m} \prod_{r=m+1}^{n-1} (1-d_r)$$

The term “Ratio of Lost work to Demanded work (L_w/D_w)” is used here to investigate the adequacy of supplied capacity in response to demand for service, whereas the term “Work Utilization Coefficient (WUC)” indicates utilization of supplied capacity to the system. Analytically, the higher the value of the ratio L_w/D_w reflects the higher amount of lost transportation work and vice versa. For the hypothetical bus route presented in the fifth section, the values of WUC and L_w/D_w for different sizes of buses against service frequency is shown in Figure 4.

NUMERICAL EXAMPLES

The way in which a transit system designer can use the developed model is illustrated here by a series of examples. The impacts of simple policy decisions are also described. A numerical example is presented using the same demand data from Hickman (2001). Suppose an operator wants to examine the role of bus size and number of buses on a route with 10 stops, where each stop has infinite waiting room and stops are equally spaced. It is assumed that passengers arrive at all stops following the Poisson distribution with mean arrival rate specified in column 2 of Table 2. Table 2 also shows the proportion of passengers who alight at stops. Since empty buses start their journeys from the first stop, no passengers alight there. However, passengers alight at the second through the ninth stop, if these are their destinations and all remaining passengers alight upon the arrival of a bus at the tenth stop. To investigate the role of bus sizes on transit system performance, bus capacity is increased in a stepwise manner from 20-passenger bus to 60-passenger bus. In addition, the bus headways are assumed to be less than 12 minutes as mentioned earlier. In this analysis, bus capacity includes standing passenger spaces and available number of seats. With such a fixed configuration, a transit designer will be confronted with two basic problems, determining bus size and fleet size.

Table 2: Parameters of Example Bus Route (Hickman 2001)

Stop	Arrival Rate (passenger/minutes)	Alighting Proportion
1	0.75	0
2	1.5	0
3	0.75	0.1
4	3	0.25
5	1.5	0.25
6	1	0.5
7	0.75	0.5
9	0	0.75
10	0	1

Numerical analysis of performance measures of the transit system with respect to passengers served by the system and those abandoned due to unavailability of space on buses as a function of different service frequencies are shown in Figures 2 and 3.

Based on the figures, it is observed that the number of passengers served by the system increases as the bus size and frequency increases. This also implies that as the bus size and frequency increase, the number of abandoned passengers decreases. However, unused capacity increases with an increase in bus size and frequency. This suggests that when a transit agency runs large buses, the number of passengers that can be accommodated on the bus increases for a given service demand, and cuts down the number of abandoned passengers.

Figure 2: Passengers Served Per Hour by Different Sizes of Buses Against Service Frequency

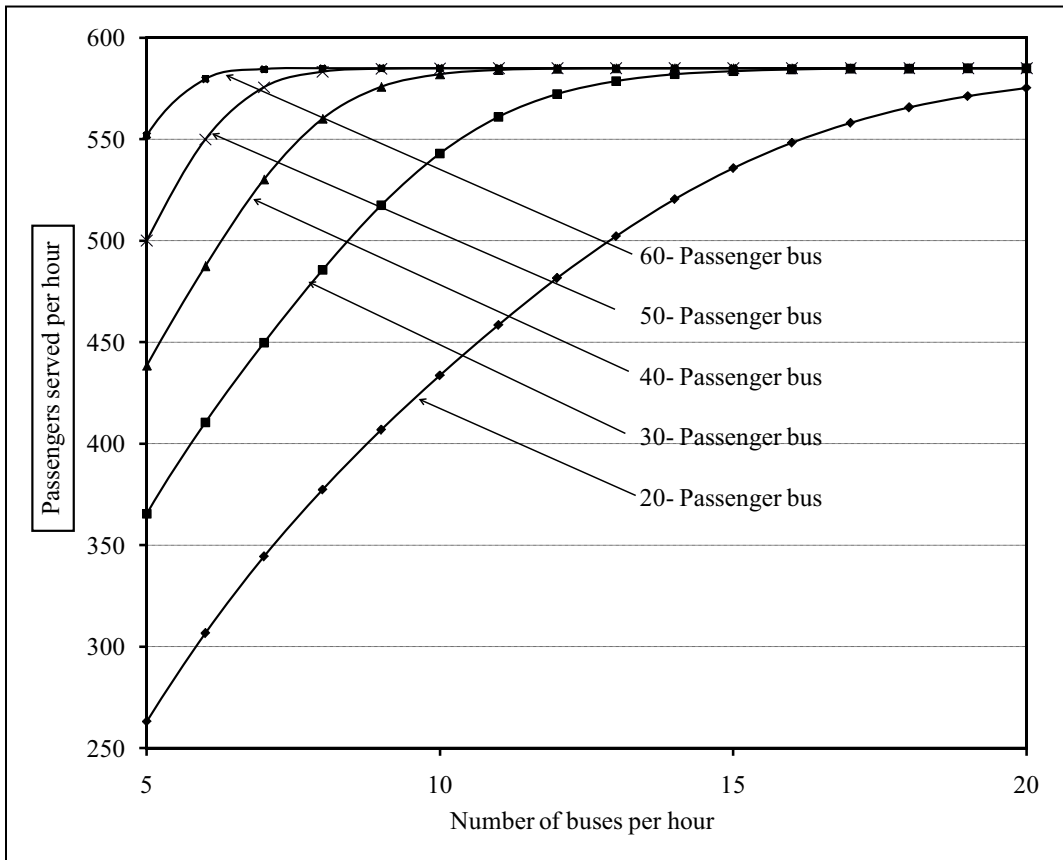


Figure 3: Abandoned Passengers per Hour by Different Sizes of Buses Against Service Frequency

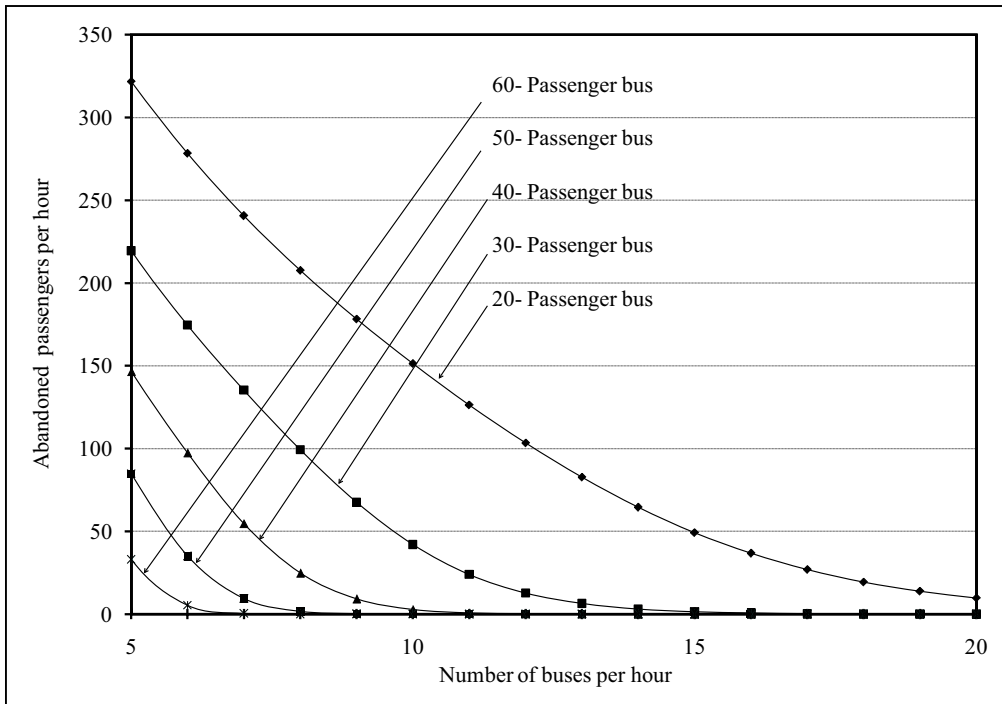


Figure 4 (a): Work Utilization Coefficient for Different Sizes of Buses Against Service Frequency

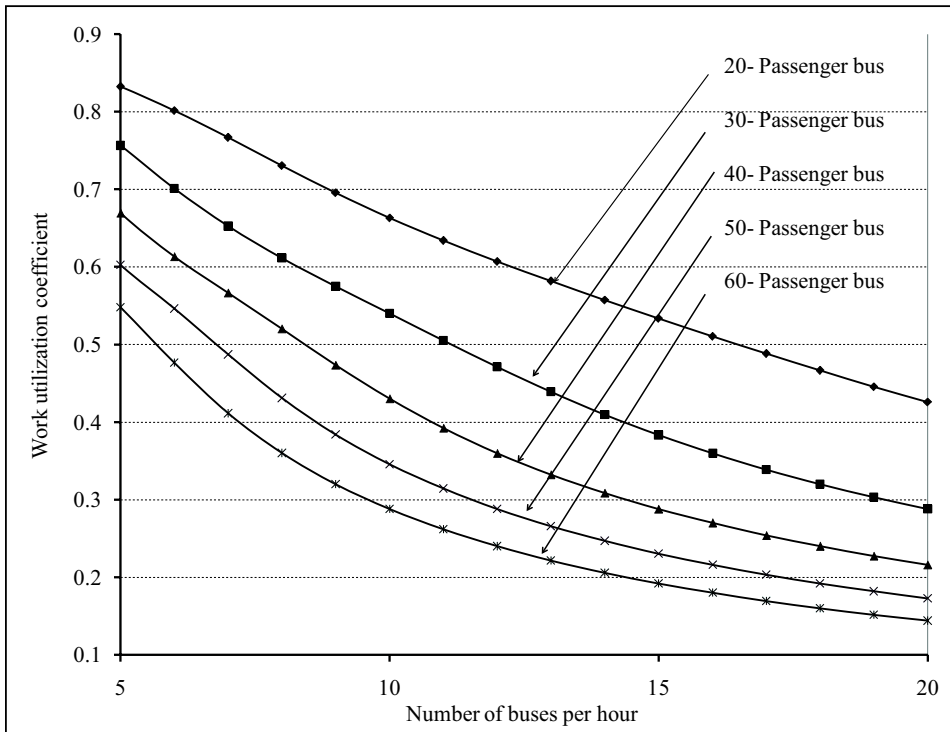
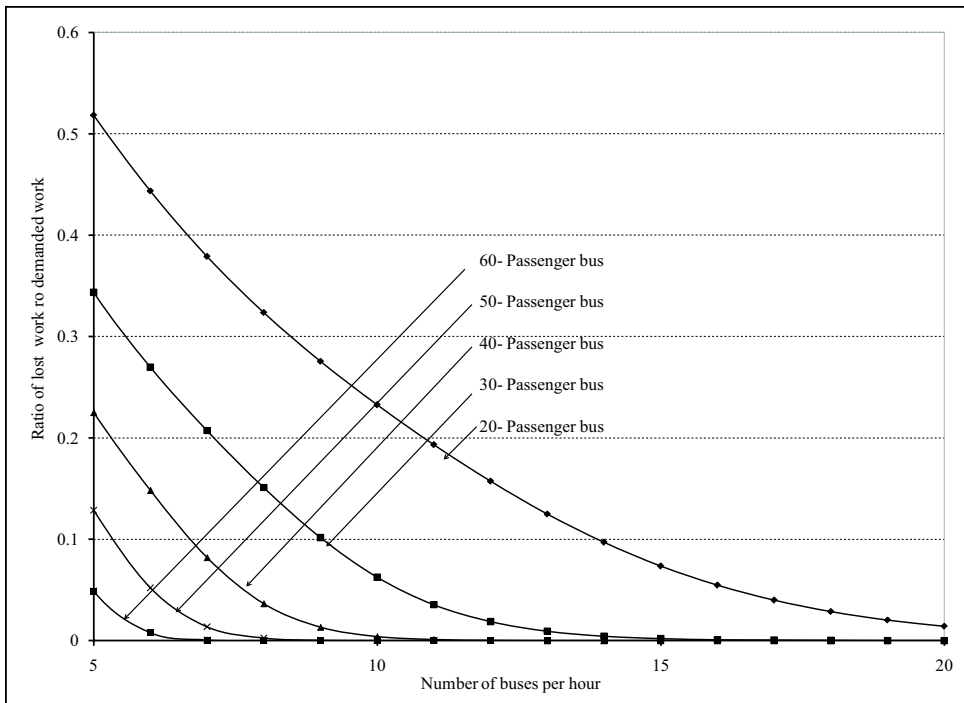


Figure 4 (b): Ratio of Lost Work to Demanded Work for Different Sizes of Buses Against Service Frequencies



The utilization of system capacity is shown in Figure 4 (a) and (b) using the WUC and the ratio of lost work to demanded work (L_w/D_w), respectively. It is observed that both WUC and L_w/D_w decrease with increases in bus size and frequency. A decrease in L_w/D_w indicates that the amount of lost work will be reduced if large sizes of buses with higher frequency are supplied [Figure 4 (b)]. However, in such cases, WUC will be reduced [Figure 4 (a)] indicating lower utilization of systems, which is not desirable to transit operators. Thus, the problem of the transit system designer is to select the appropriate total system capacity, which trades off between bus size and frequency.

Figure 5: Comparison of Policy to Examine System Behavior by Providing a Few Larger Buses or Number of Smaller Buses

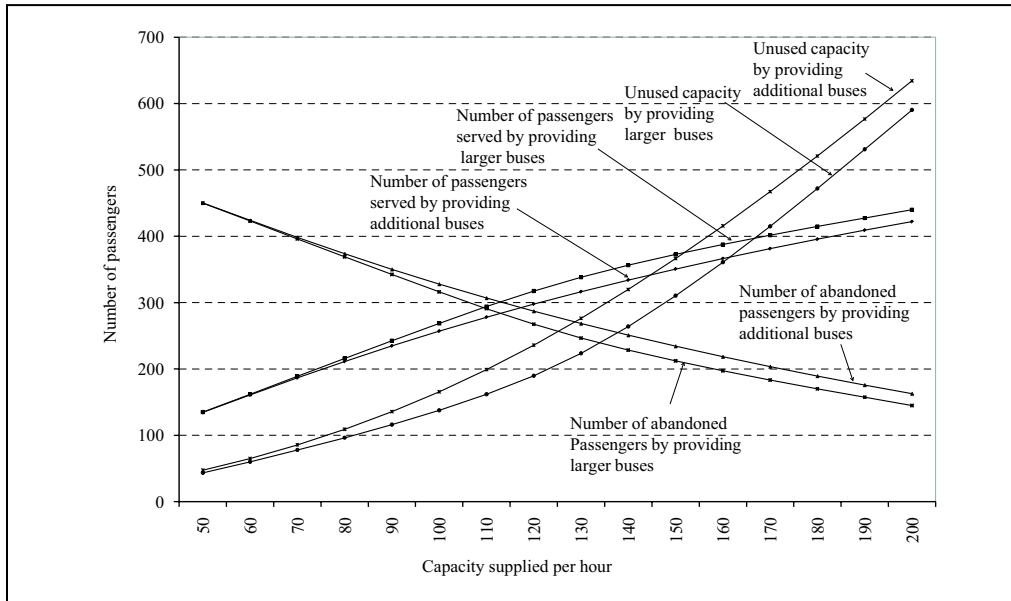


Figure 5 shows the trade-off between providing larger or smaller buses as a function of spaces supplied per hour to the system. If it is assumed that the transport planner has the freedom to use any number of buses of a particular size, then the same level of system capacity can be achieved by providing a small number of large buses at a smaller frequency or a large number of smaller buses with a high frequency. For example, a total system capacity of 100 spaces per hour can be supplied by five 20-passenger buses per hour. The effect of such policy choice is explored graphically using the numerical example in Figure 5. It is observed that there is a smaller number of abandoned passengers as well as lower unused capacity when there are a smaller number of large buses as compared with a larger number of small buses. Therefore, it can be said that under the assumptions made here, the best operating strategy is to select the largest bus size. Moreover, in Figure 5 the intersection points between the curves representing unused capacity and abandoned passengers seem that they should indicate a “best” policy for this application. That is, they identify several “optimal” points with respect to unused capacity and abandoned passengers. Looking at this figure, supplying between 125 and 140 capacity per hour would result in an optimal trade-off between unused capacity and abandoned passengers. Hence, this figure indicates possible strategies for selecting bus sizes based on unused capacity and abandoned passengers. Thus, it can be viewed as a simple application of the proposed models.

Figure 6 (a): Standard Deviation of Passengers Served Per Bus by Different Sizes of Buses Against Service Frequency at Stop 5

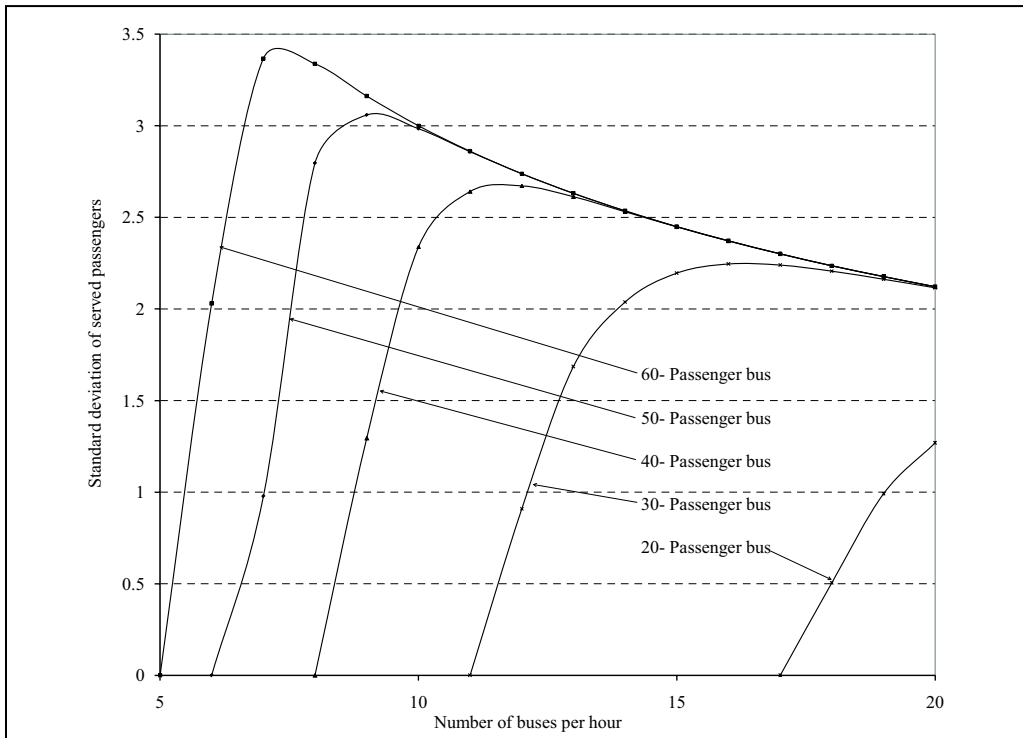
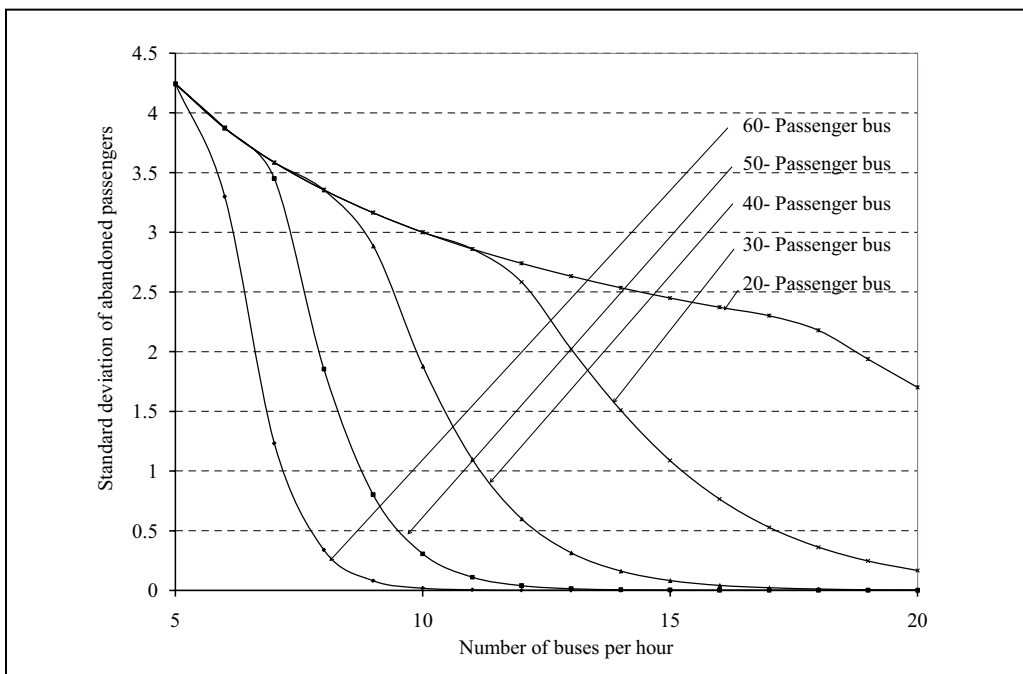


Figure 6 (b): Standard Deviation of Abandoned Passengers per Bus for Different Sizes of Buses Against Service Frequency at Stop 5



In addition to the above performance measures, transit operators are interested in the variance of passengers served, as well as the variance of abandoned passengers as performance indicators. Figure 6 (a) and 6 (b) show the variance of served passengers and variance of abandoned passengers for different sizes of bus and frequency at stop 5, which is the stop with maximum load on the route.

Figure 6 (a) shows that for a 20-passenger bus, the variation of served passengers at this stop is zero for frequencies between five to 17 buses per hour. This indicates that a 20-passenger bus will be full with respect to its carrying capacity, and passengers will be abandoned during the operation for most service times. Similarly, zero variance of served passengers for other bus sizes and frequencies signifies inability to satisfy passenger demand along the route. The figure shows that there is a point for each bus size after which the variance of served passengers decreases with an increase in bus frequency. Figure 6 (b) shows that variance of abandoned passengers decreases with the increase in bus size and frequency. The lower variance of abandoned passengers indicates the high probability of receiving transit service and vice versa.

CONCLUSIONS

In this paper, passenger renegeing behavior is modeled in relation to bus sizes and frequencies used in transit operation. Passengers are described as “impatient” when they leave stops without further waiting, once they are unable to board a bus due to capacity constraints. This behavior can be seen as an approximation of a particular situation where substantial one-way suburban commuters look for an alternative mode of transportation at rush hour to be on time at the workplace. Another situation where this analysis applies is the possibility of balking when a passenger waiting for a bus on a street outside his/her home has a vehicle ready to use in case of inability to board the preferred bus. Using the Markov Chain technique, the stochastic elements of the bus transit system and its performance measures are derived. This model is then demonstrated using numerical examples to illustrate the impacts on transit policy. This model can be viewed as a simplified means to evaluate transit system performance under different levels of supply and demand. This simplified model is able to provide practitioners quantitative insights to problems regarding vehicle sizes and frequencies quickly and effectively.

References

- A. Grosfeld-Nir and J.H. Bookbinder. “The Planning of Headways in Urban Public Transit.” *Annals of Operations Research* 60 (1-4), (1995): 145-60.
- Andersson, P.Å., and G.P. Scalia-Tomba. “A Mathematical Model of an Urban Bus Route.” *Transportation Research Part B: Methodological* 15, (1981): 249-66.
- Bailey, N.T. J. “On Queuing Processes with Bulk Service.” *Journal of the Royal Statistical Society. Series B (Methodological)* 16, (1954): 80-87.
- Bowman, L. A., and M. A. Turnquist. “Service Frequency, Schedule Reliability and Passenger Wait Times at Transit Stops.” *Transportation Research A* 15(6), (1981): 465–471.
- Chien, S. “Optimization of Headway, Vehicle Size and Route Choice for Minimum Cost Feeder Service.” *Transportation Planning and Technology* 28 (5), (2005): 359–80.
- Dell’Olio, L., A. Ibeas, and J.L. Moura, “Headway and Vehicle Size Optimization in Urban Bus Transit Networks.” Third International Symposium of Transport Simulation (ISTS08), Monash University, Monash, 2008.

- Downton, F. "On Limiting Distributions Arising in Bulk Service Queues." *Journal of the Royal Statistical Society Series B* 17, (1955): 256-61.
- Fan, W. and R. Machemehl. "Do Transit Users Just Wait for Buses or Wait With Strategies? Some Numerical Results That Transit Planners Should See." *Transportation Research Record* 2111, (2009): 169-76.
- Giffin, W.C. "A Simplified Model for Bulk Service at a Series of Stations." *Journal of Industrial Engineering* 8 (17), (1966): 430-36.
- Hickman, M.D. "An Analytic Stochastic Model for the Transit Vehicle Holding Problem." *Transportation Science* 35 (3), (2001): 215-37.
- Islam, M.K., U. Vandebona, V.V. Dixit, and A. Sharma. "A Bulk Queue Model for the Evaluation of Impact of Headway Variations and Passenger Waiting Behavior on Public Transit Performance." *IEEE Transactions on Intelligent Transportation Systems* 15 (6), (2014): 2432-2442.
- Jansson, J.O. "A Simple Bus Line Model for Optimisation of Service Frequency and Bus Size." *Journal of Transport Economy and Policy* 14 (1), (1980): 53-80.
- Jaiswal, N.K. "A Bulk Service Queueing Problem with Variable Capacity." *Journal of the Royal Statistical Society Series B* 23, (1961): 143-48.
- Jansson, K. "Optimal Public Transport Prices, Service Frequency and Transport Unit Size." Paper Presented at the Selected Proceedings of The Sixth World Conference on Transport Research, 1993.
- Jolliffe, J.K. and T.P. Hutchinson. "A Behavioral Explanation of the Association Between Bus and Passenger Arrivals at a Bus Stop." *Transportation Science* 9(4), (1975): 248-282.
- Lam, S.W., L.C. Tang, T.N. Goh, and T. Halim. "Multiresponse Optimization of Dispatch Rules for Public Bus Services." *Computers & Industrial Engineering* 56(1), (2009): 77-86.
- Lee, K.K.T., S.H.F. Kuo, and P.M. Schonfeld. "Optimal Mixed Bus Fleet for Urban Operations." *Transportation Research Record* 1503, (1995): 39-48.
- Mohring, H. "Optimization and Scale Economies in Urban Bus Transportation." *American Economic Review* 62, (1972): 591-604.
- Mohring, H. *Transportation Economics*. Ballinger, Cambridge MA, 1976.
- Oldfield, R.H. and P.H. Bly. "An Analytic Investigation of Optimal Bus Size." *Transportation Research Part B: Methodological* 22 (5), (1988): 319-37.
- O'Flaherty, C.A. and D.O. Mangan. "Bus Passenger Waiting Time in Central Areas." *Traffic Engineering and Control* 11(9), 1970: 419-421.
- Okrent, M. "Effects of Transit Service Characteristics on Passenger Waiting Time." M. S. Thesis, Northwestern University, Department of Civil Engineering, Evanston. Illinois, 1974.
- Ross, S.M. *Markov Chains. Introduction to Probability Models*. Academic, Orlando, FL, USA, 2003.

Rietveld, P., J. Dekkers, and S. Woudenberg. "Choice of Frequency and Vehicle Size in Rail Transport: Implications for Marginal External Costs." TI 2001-108/3 Tinbergen Institute Discussion Paper (2001).

Seddon, P.A. and M.P. Day. "Bus Passenger Waiting Times in Greater Manchester." *Traffic Engineering and Control* 15, (1974):422–445.

Vuchic, Vukan R. *Urban Transit: Operations, Planning and Economics*. John Wiley & Sons, 2005.

Md. Kamrul Islam received a B.Sc. degree in civil engineering from Bangladesh Institute of Technology, Gazipur, Bangladesh (Currently known as Dhaka University of Engineering and Technology, Gazipur, Bangladesh) and a M. Eng. degree from the University of Tokyo, Tokyo, Japan, and a Ph.D. degree at the School of Civil and Environmental Engineering, University of New South Wales, Australia.

He is an assistant professor at the department of civil engineering at Dhaka University of Engineering and Technology, Gazipur, Bangladesh. His research interests are in modeling of public transit reliability and control strategies for transit operations.

Upali Vandebona received a B.Sc. in civil engineering from the University of Ceylon in 1975, M. Eng. from Asian Institute of Technology, Thailand, in 1977, and a Ph.D. in transport engineering from Monash University, Australia, in 1986.

He was a senior lecturer of the School of Civil and Environmental Engineering at University of New South Wales, Australia. He was the 2008 PUTRUM Fellow (Public Transit for Urban Mobility Program) at the University of Calgary. His research publications are mainly related to analytical modeling and simulation applications for urban public transport and freight systems.

Vinayak V. Dixit received an M.Tech degree in mathematics and computing from Indian Institute of Technology, Delhi, India, in 2005, and a Ph.D. in transportation engineering from University of Central Florida in 2008.

He is a senior lecturer at the School of Civil and Environmental Engineering at the University of New South Wales (UNSW), Australia. His key research interest lies in studying risk in the transportation infrastructure system as it relates to highway safety, travel time uncertainty, as well as natural and man-made disasters.

Ashish Sharma received a bachelor's degree from the University of Roorkee, Roorkee, India, a master's degree from Indian Institute of Technology, Delhi, India, and a Ph.D. degree from Utah State University.

He is a professor at the School of Civil and Environmental Engineering at the University of New South Wales, Australia, along with being an Australian Research Council Future Fellow up to 2015. Ashish's research interests are related to characterization and use of stochastic techniques to various aspects of civil and environmental engineering.