GENERAL PURPOSE MODEL OF QUEUE DISSIPATION TIME AT SERVICE FACILITY WITH INTERMITTENT BULK SERVICE SCHEDULE

Shinya KIKUCHI

Civil Engineering Department, University of Delaware, Newark, Delaware 19716, U.S.A.

Yoshio HAMAMATSU

Tamagawa University, Tokyo, Japan

Abstract. An absorbing Markov Chain model is formulated to describe the queue formation and dissipation process at a service facility. The model yields the average time for a queue to dissipate and the probability for the queue to reach a certain length, using the properties of the fundamental matrix derived from the canonical form of the transition probability matrix of the Markov chain. The model is useful in evaluating the time for a queue to dissipate at a facility which provides service intermittently, for example, at a loading point of a transportation facility. At these locations, a vehicle cannot depart until all the waiting passengers (or cargo) are aboard the vehicle. The delay to a user is thus affected not only by the number of persons ahead in the queue but also by the ones behind him in the queue and the ones who join the queue during the boarding process. The total waiting time of the first person in the queue is approximately equal to the vehicle standing time and he experiences the longest delay since he also had waited the longest before the vehicle arrived. The last person in the queue experiences the shortest delay. This paper formulates the general purpose model for calculating delay, queue dissipation time, and queue length fluctuation under such conditions. The model may be applied to a number of queuing situations in which dissipation of the entire queue is the main concern, including problems of the dissipation of traffic back-up at a traffic accident site or road construction site.

Key words and phrases: queuing theory, transportation

1. INTRODUCTION

In designing and operating any service facility, queue formation and dissipation behavior is an important concern since it relates to the determination of the waiting area size, service schedule, and control strategies. Among numerous examples of queuing problems, this paper analyzes the behavior of a queue at a facility which provides services intermittently in a bulk mode.

Units arrive randomly at the service facility, and form a queue; at a given time, the facility opens and begins service; the length of the queue fluctuates due to the

arrivals of more units and the departures of unts; eventually, the queue dissipates, and when no unit arrives for a certain time the service facility closes until the next service time. Typical examples of such a queuing pattern are seen in the passenger boarding process at a shuttle bus terminal or in the vehicle loading process at a ferry boat terminal where units wait for the arrival of a vehicle; once the vehicle arrives, the passengers begin boarding; the vehicle departs when all passengers including ones who arrive during the boarding process complete boarding (the entire queue dissipates).

In the above case, the delay experienced by a unit is the sum of: the time spent for waiting for the vehicle; and, once the vehicle arrived, the waiting time before boarding the vehicle; and the time required to complete boarding of all the units behind it. The longest delay is experienced by the first unit in the queue, since it must wait until all the units behind it board the vehicle, in addition to the time it waited for the arrival of the vehicle, which is also the longest. An interesting feature of this type of queuing problem is that the total delay experienced by a unit is determined not only by the length of the queue in front of it but, also by the length of the queue behind it.

The queuing process described above is schematically shown in Figure 1. In the figure, a dot represents a unit. Its arrival time at the service facility is shown by its location on the horizontal axis. Its location in the queue is shown by the vertical axis. For example, unit (X) is in the queue when the service begins (when the vehicle arrives); it waits for time (a) to reach the front of the queue; it spends time (b) to board the vehicle; and then it waits until the whole queue dissipates; at this time the vehicle departs. Unit (Y), on the other hand, arrives and joins the queue while some other units are waiting or boarding the vehicle and it departs with unit (X) at the same time. To analyze the delay of this type, the conventional queuing theory is not sufficient, since it only deals with the time for a unit to reach the service channel. For the type of problem described above, the dissipation time of the entire queue must be considered.

This paper develops a general purpose model which analyzes queue dissipation time using an absorbing Markov chain. The model describes the dynamics of queue formation and dissipation process, and yields the average queue dissipation time (which is the average delay), and the probability that the queue length reaches a certain value during the queue fluctuation process. In developing and explaining the model, passenger arrival, waiting and boarding processes at a shuttle bus terminal are visualized, only to facilitate the explanation. Many other situations may be considered to illustrate the model, including vehicle loading process at a ferry-boat terminal, train formation process at a freight car marshalling yard. In these cases, the departure of the vehicle (bus, ferry-boat, train) is subject to the clearance of the entire queue formed by the arriving traffic units (passengers, automobiles, freight cars). The vehicle (bus, ferry-boat, or train) standing time at the terminal or station is approximately equal to the time for the entire queue to dissipate, except for the times for departure preparation.

In practical analyses, the standing time of a vehicle at a stop has been computed as the product of the average boarding time per passenger and the number of passengers. This approach may suffice for planning purposes; however, in order to look at the queue formation during the boarding process and to gain a more accurate understanding of vehicle standing time and delay, an approach which enables us to examine the dynamics of queue formation and dissipation must be employed. This is particularly important for the situations in which boarding rate is small and the variation of vehicle standing time is critical to the operation, for example, container loading onto rail freight cars and the train formation at a marshalling yard.

2. PAST STUDIES

In most deterministic models, the queue formation and dissipation processes are usually analyzed by a diagram showing the relationship between the accumulation of the units and time, in which the cumulative arrivals and departures of traffic units are shown on the vertical axis and time on the horizontal axis. The envelop formed by the cumulative arrivals and departures of traffic units represents the total queue and the vertical difference between the two lines is the queue size at a given time and the horizontal distance is the delay of a unit. Among many such examples, Morales (1986) developed programs to compute highway traffic delay by computing the area of the envelop formed by many combinations of traffic arrival and departure patterns. For passenger delays at a transit stop, Vuchic (1970) showed passenger accumulation and vehicle standing time using a similar illustration for his analysis of delay propagation. Hendrickson and Kocur (1981) also used a similar queue dissipation model in analyzing the delay of commuters caused by a traffic bottleneck. These graphical methods are deterministic and that they cannot take the probablistic nature of the phenomena into account; such as the fluctuation of queue size due to the random arrivals of units and variation of service pattern. A problem of marginal delay caused by an additional traffic unit in the queue was analyzed by Wohl (1970), in which he analyzed the effect of additional arrivals to the queue dissipation time.

3. ASSUMPTIONS, NOTATION AND APPROACH

ASSUMPTIONS

- (1) Assuming a close headway operation, passengers arrive randomly at the terminal independent of bus arrivals. Only boarding takes place at the stop, and passengers board through a single door. Those who arrive while the bus is standing are also allowed to board the bus. The capacity of the bus is sufficiently large for the demand so that all waiting passengers and the ones arriving during the boarding process can board the vehicle.
- (2) The events of passenger arrival at the stop and passenger boarding are observed in the increment of a small time interval; in each interval, no more than one passenger arrives, and no more than one passenger boards the bus. The time

interval is designated as Δt , and this unit is used to count time in the Markov chain model.

- (3) Probability of one person arriving in Δt is p. Probability of one person boarding during Δt is m. The events of passenger arrival at the stop and passenger boarding on a small interval (Δt) are assumed to be independent; the arrival rate does not influence the boarding rate. The probability of no one arriving in Δt is q (= 1-p), and the probability of no one boarding in Δt is l (= 1-m).
- (4) The initial queue size (the number of passengers at the stop when the vehicle arrived) is a function of the time elapsed since the departure of the previous vehicle and the passenger arrival rate, p. If the elapsed time is expressed by kΔt, the probability of x persons waiting at the stop when the vehicle arrived (the initial queue size) is expressed as a binomial form as follows:

$$A_x = \text{Prob } (x \text{ passengers at the stop, or initial queue } = x)$$

$$= {k \choose x} p^x q^{k-x}. \tag{1}$$

(5) The vehicle leaves the stop as soon as the passenger queue is cleared (queue dissipates).

NOTATION

p — probability that a person arrives in Δt ,

q — probability that no person arrives in Δt (= 1 - p)

P — Canonical form of the transition probability matrix

N — Fundamental matrix of the transition probability matrix, P

B — Absorption probability vector

 W_1 — Average vehicle standing time at the station (including by-pass situation)

 W_2 — Average vehicle standing time when the vehicle stops at the station

k — number of time intervals elapsed since the departure of the last vehicle

 number of passengers waiting in the queue at the time of vehicle arrival (initial queue length)

 A_x — probability that x persons are waiting at the stop when the bus arrived

m — probability that a person boards the vehicle in Δt

I — probability that a person is not able to board the bus in Δt , (=1-m)

s - queue length (or the state of the system) at a given time

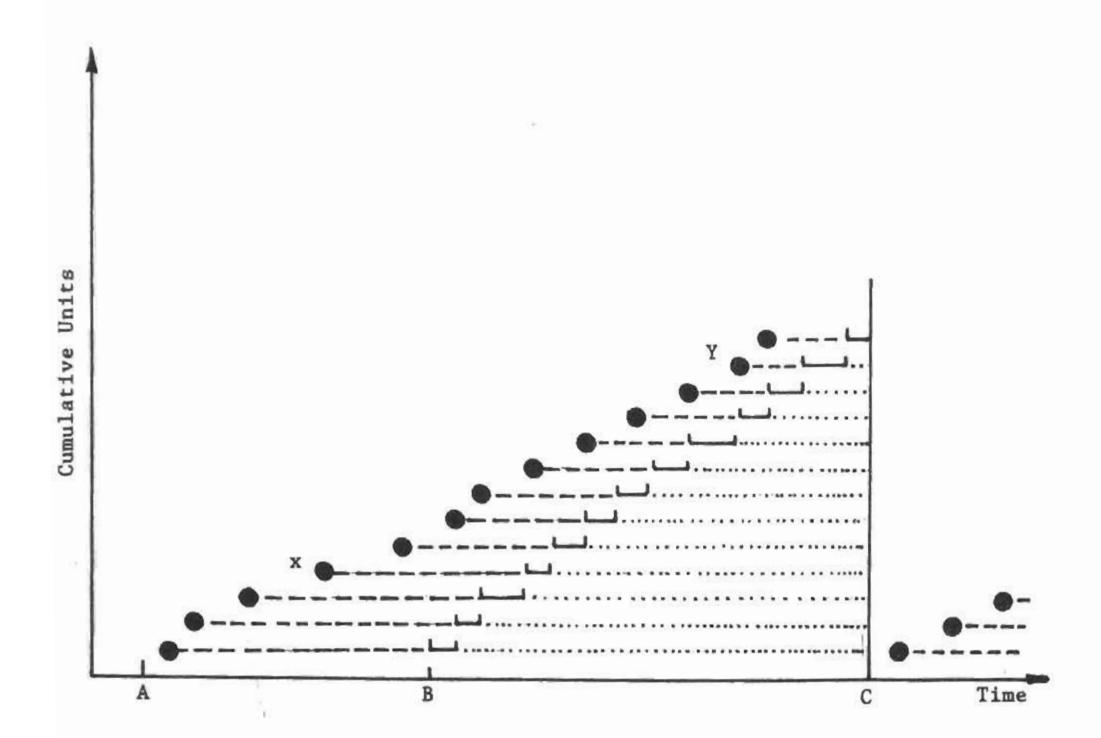
 T_i — Average time in $(\Delta t$'s) to reach the absorbing state from state i

 $h_{i,j}$ — probability that queue length becomes j starting from i

 L_j — probability that the queue length becomes j

4. THE MODEL

Fluctuation of queue length and its eventual dissipation is modelled by an absorbing Markov Chain. In the model, the state is defined by the number of persons



waiting in the queue (before the arrival of the vehicle) boarding

waiting for queue dissipation (waiting in the vehicle after boarding)

A: Time of the previous vehicle's departure

B: Time of vehicle arrival

C: Time of vehicle departure

FIGURE 1. Illustration of the Queuing Process Analyzed in the Model

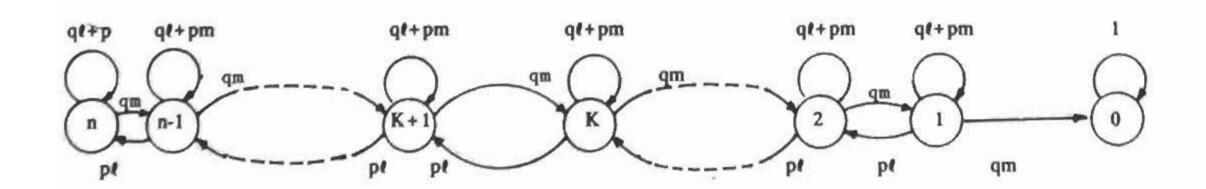


FIGURE 2. Markov Chain Diagram of the Model

in the queue (queue length). Since, in order for the bus to depart, the number of persons in the queue must reduce to zero eventually, the passenger boarding process can be considered as an absorbing Markov chain process. The absorbing state is the state in which the queue has dissipated and thus, the queue length is zero. This section presents the following steps: (1) illustration of the absorbing Markov Chain, (2) transition matrix, canonical form and fundamental matrix of the absorbing chain, (3) expected absorption time and average bus standing time, and (4) the queue length probability.

MARKOV CHAIN DIAGRAM

The Markov chain diagram which depicts the dynamics of the fluctuation of the queue length is shown in Figure 2. Its characteristics are summarized as:

- State (s): number of persons in the queue in Δt ,
- Absorption state (s = 0): no person is in the queue and that the vehicle leaves the stop,
- Time unit (Δt) ; a small time interval in which no more than one person joins the queue, and also in which a maximum of one person can board the vehicle. The state of the system is observed in the increments of Δt .
- Initial state (s = i): the state when the bus arrived at the stop; the number of persons in the queue when the vehicle arrived.
- Transitions: possible transitions of state between two consecutive time periods are:
 - 1. no change in the queue length, no change in state $(s \rightarrow s)$

No change in queue length occurs when either no person arrives and no one completes boarding, or one person arrives and one person boards. The probability of no change in state is ql + pm.

2. queue length increases by one $(s \rightarrow s + 1)$

Queue length increases by one when no one completes boarding and one person joins the queue. The probability of queue length increase by one is pl.

3. queue length decreases by one $(s \rightarrow s-1)$

Queue length decreases by one when one person boards and no one joins the queue. This probability is qm.

It should be noted that in order for the queue to dissipate eventually, qm must be greater than pl, (qm > pl); only under this condition, the process eventually moved to s = 0.

— Queue length limit (n): an artificial limit of the queue length. This limit is necessary to define the size of the transition matrix. Thus, in the early stage of the model development, persons who arrive at the stop when the queue length is n, are assumed not to join the queue. Therefore, at s = n, the possible transitions of state are n to n-1 (no arrival and one boarding), and n to n (no arrival and no boarding). However, in the later stage of the model

development, this constraint will be relaxed by letting n be infinity in order to show the general case.

TRANSITION MATRIX AND FUNDAMENTAL MATRIX

The transition probability matrix for the Markov chain of Figure 2 is shown in the canonical form in Table 1. In the table, the matrix is segmented into four sub-matrices by dashed lines: I, an identity matrix, representing the absorption probability; R, a vector showing the transition probability into absorption; O, a zero vector; and Q, a transition matrix between nonabsorbing states.

TABLE 1. Transition Probability Matrix in Canonical Form

The inverse of matrix I - Q is the fundamental matrix (N), whose elements, (i, j), represents the average number of times the process is in state j if it starts from i. This matrix possesses interesting properties; by simple manipulations, it yields the probability to reach an absorption state(s), and also, the average time to reach the absorption state(s) from a given state. For detailed explanation of the properties of the transition matrix of an absorbing Markov chain, readers are advised to refer to Kemeny and Snell (1976).

The fundamental matrix, N, is obtained as

$$N = I + Q + Q^2 + \dots = \sum_{k=0}^{n} Q^k = (I - Q)^{-1}.$$
 (2)

Elements of the fundamental matrix, $N_{i,j}$, are obtained as

$$N_{i,j} = \begin{cases} \frac{1}{qm} \sum_{k=0}^{j-1} \left(\frac{pl}{qm}\right)^k, & \text{for } i \ge j \\ \frac{1}{qm} \sum_{k=j-i}^{j-1} \left(\frac{pl}{qm}\right)^k, & \text{for } i < j \end{cases}$$
(3)

where 1 < i < n, 1 < j < n, (see Appendix 1) for derivation of (3) and (4).

The probability of reaching the absorption state from state i, B_i , is given by the product of N and R, NR. In this case, there is only one absorption state, and

the process always moves to the absorbing state (s = 0) regardless of the initial starting state, thus, B_i is a unit column vector.

AVERAGE ABSORPTION TIME (QUEUE DISSIPATION TIME FROM A GIVEN STATE)

The average time to reach the absorption state from state i is computed as the sum of the elements of the i-th row of matrix N as shown in (3) and (4):

$$T_{i} = \sum_{j=1}^{n} N_{i,j} = \sum_{j=1}^{i} N_{i,j} + \sum_{j=i+1}^{n} N_{i,j}.$$
 (5)

Using the relationships m + l = 1, and p + q = 1 with (3) and (4), T_i , average absorption time, is obtained as:

$$T_{i} = \begin{cases} \frac{1}{qm} \sum_{k=0}^{n-1} \left(\frac{pl}{qm}\right)^{k}, & i = 1 \\ \frac{1}{qm} \left\{ i \sum_{k=0}^{n-i} \left(\frac{pl}{qm}\right)^{k} + \sum_{k=n-i+1}^{n-1} (n-k) \left(\frac{pl}{qm}\right)^{k} \right\}, & 2 \leq i \leq n. \end{cases}$$
 (6)

These two experessions can be combined and simplified (see Appendix 2):

$$T_{i} = \frac{i(m-p) - qm \left\{ 1 - \left(\frac{pl}{qm}\right)^{i} \right\} \left(\frac{pl}{qm}\right)^{n-i+1}}{(m-p)^{2}}, \qquad 1 \le i \le n.$$
 (8)

So far, the analysis has been based on the Markov Chain which has a limited number of states, n, an artificial bound of the queue size. However, as seen in (3) and (4), the values of $N_{i,j}$ are independent of n, and that n merely defines the size of the transition probability matrix. Therefore, when n is set infinity, B is still 1, which indicates that the Markov chain is still an absorbing chain, and hence, the average absorption time is given by making n infinity in (8). Since qm > pl,

$$T_i = \frac{i}{m - p}. (9)$$

This expression represents the queue dissipation time from a given initial queue length for the general case (no upper bound on the queue length).

AVERAGE QUEUE DISSIPATION TIME

The average time for a queue to dissipate, W_1 , is obtained from the product of the probability of the initial queue length and the average absorption time from the initial queue length:

$$W_1 = \sum_{i=0}^k A_i T_i \tag{10}$$

where A_i is the probability that the initial queue length is i, as given by (1) and also T_i is given by (9). Thus, the average queue dissipation time, W_1 is obtained as:

$$W_{1} = \frac{1}{m-p} \sum_{i=0}^{k} i {k \choose i} p^{i} q^{k-i}. \tag{11}$$

This average includes the case in which the initial queue length is zero, in other words, the case in which the vehicle bypasses the stop because no one is waiting for the bus.

The queue dissipation time only for the case when a queue is actually formed, W_2 , can also be developed as:

$$W_2 = \sum_{i=1}^k A_i T_i \tag{12}$$

where \hat{A}_i is the probability that the process reaches the absorbing state from any non-absorbing state $s = 1, \ldots, k$ (not including s = 0). This probability is obtained using the Bayes rule as follows:

$$\hat{A}_i = \frac{A_i}{\sum_{a=0}^k A_a B_a} \tag{13}$$

where B_i is the absorption probability from state i and it is 1 as discussed before. Therefore, the denominator of the above becomes $1 - A_0$ or $1 - q^K$. Equation (13) is now expressed as:

$$\hat{A}_i = \frac{1}{1 - q^k} \binom{K}{i} p^i q^{K - i}. \tag{14}$$

Therefore, the average queue dissipation time (standing time of the bus) when it stops at the stop is:

$$W_2 = \frac{1}{(1 - q^K)(m - p)} \sum_{i=1}^K i \binom{K}{i} p^i q^{K - i}. \tag{15}$$

PROBABILITY OF QUEUE LENGTH

During the queue dissipation process, the queue length fluctuates. It may increase temporarily if more passengers arrive. The probability that the queue length becomes a given value is also developed here, using the fundamental matrix. According to Kemeny and Snell (1976), the probability that the process begins from state i and reaches j, $h_{i,j}$, is given by:

$$h_{i,j} = (N - I)N_{\mathrm{dg}}^{-1} \tag{16}$$

where N_{dg}^{-1} is a matrix whose diagonal elements are the same as those of N and all other elements zero, and $h_{i,j}$ is computed as follows (see Appendix 3):

$$h_{i,j} = \begin{cases} 1, & i > j \\ (1 - qm) \left[\sum_{k=0}^{j-1} \left(\frac{pl}{qm} \right)^k \right]^{-1}, & i = j \\ \left[\sum_{k=j-i}^{j-1} \left(\frac{pl}{qm} \right)^k \right] \cdot \left[\sum_{k=0}^{j-1} \left(\frac{pl}{qm} \right)^k \right]^{-1}, & i < j \end{cases}$$
(18)

As seen in Figure 2, if i > j, $h_{i,j}$ is 1 since the process must pass j in order to reach the absorbing state; while for i < j, $h_{i,j}$ is less than 1, since it is not certain that the queue length increases during the passenger boarding process. By multiplying the probability of the initial queue length A_i , and $h_{i,j}$, the unconditional probability of queue length, L_j , is derived as:

$$\operatorname{Prob}(L_j) = \sum_{j=1}^K A_i h_{i,j}. \tag{20}$$

5. EXAMPLE

A simple example of the model is presented, in which passenger arrivals follow a Poisson distribution, and passenger boarding is deterministic, (a passenger always boards the bus within Δt , m=1). The transition probabilities of this case are shown in canonical from in Table 2.

TABLE 2. Transition Probability Matrix in Canonical Form for Example Problem

The fundamental matrix is obtained as follows:

$$N = (I - Q)^{-1} = \begin{bmatrix} 1/q & 0 & 0 & \dots & 0 \\ 1/q & 1/q & 0 & 0 & \dots & 0 \\ \vdots & & & & & \\ 1/q & 1/q & \dots & \dots & 1/q \end{bmatrix}$$
(21)

The average queue dissipation time from a given initial state is given by (9) as:

$$T_i = \frac{i}{q}, \qquad i = 1, 2, \dots, K.$$
 (22)

The average vehicle standing time, W_1 , is given by (10):

$$W_1 = \sum_{i=0}^{K} A_i T_i = \frac{1}{q} \sum_{i=1}^{K} i \binom{K}{i} p^i q^{K-i}. \tag{23}$$

The average vehicle standing time only when the vehicle stops at the stop, W_2 is given by (15):

$$W_2 = \frac{1}{q(1-q^K)} \sum_{i=1}^K i \binom{K}{i} p^i q^{K-i}$$
 (24)

Since the fundamental matrix is in rather simple form, the probability of vehicle standing time given the initial passenger queue size can be developed by examining the combination of passenger arrivals and initial states. For example, the probability that the dissipation time is $2\Delta t$, given the initial queue length of 2 occurs when no person arrives for two consecutive time periods, q^2 . Table 3 shows the probabilities under different combinations of initial queue length and the queue dissipation time.

	Queue dissipation Time in Multiples of Δt , μ								
	0	1	2	3	4	5	6	7	
0	1	0	0	0	0	0	0	0	
1	0	q	pq	p^2q	p^3q	p4q	p^5q	p^6q	
2	0	0	q^2	$2pq^2$	$3p^2q^2$	$4p^3q^2$	$5p^4q^2$	$6p^{5}q^{2}$	***
3	0	0	0	q ³	$3pq^3$	$6p^2q^3$	$10p^4q^3$	$15p^{5}q^{3}$	
4	0	0	0	0	494	4pq4	$10p^2q^4$	20p4q4	
5	0	0	0	0	0	q^5	$5pq^5$	$15p^2q^5$	
	0	0	0	0	0	0	p^5q	$6p^{5}q^{2}$	

TABLE 3. Probability of Queue Dissipation Time (in $\mu\Delta t$) for Given Initial Queue Length (x)

In the table, the sum of the products between the initial queue length probability, A_x , (given by (1)), and the corresponding element of column u represents the probability that the queue dissipation time is $u\Delta t$. Such probability, C_u , is summarized in general form as follows:

$$C_{u} = \begin{cases} p^{u} q^{K} \sum_{i=0}^{u-1} {u-1 \choose i} {K \choose i+1}, & u \leq K \\ p^{u} q^{K} \sum_{i=0}^{K-1} {u-1 \choose i} {K \choose i+1}, & u > K \end{cases}$$
(25)

6. CONCLUSION

The fluctuation and eventual dissipation process of a queue was modelled using an absorbing Markov Chain. For given arrival and service rates and the initial queue size at a service facility, the model yields the time for the queue to dissipate, and the probabilities of various queue lengths during the queue size fluctuation process. The expressions for these are derived using the properties of the fundamental matrix developed from the canonical from of the transition probability matrix.

The model is useful in computing boarding or loading delay at a passenger or freight transportation terminal where vehicles arrive intermittently and they depart only when all the waiting passengers or cargo are aboard. In these cases, delay to an individual is the time for the queue to dissipate counting from the time when he joined the queue, this includes the time to reach the service channel and the time for the queue behind him to clear, as well as the boarding time. The average vehicle standing time is the sum of the product of the queue dissipation time from the initial queue length and its probability. The probability of the initial queue length is influenced by the time elapsed from the departure of the last vehicle and the passenger arrival rate.

The model may also be applied to analyze train formation time at a railroad marshalling yard or container loading time onto a ship, in which the departure of the train or the ship is contingent upon the clearance of the queue.

APPENDICES

- 1. Derivation of Fundamental Matrix
- 2. Derivation of Average Queue Dissipation Time
- 3. Derivation of Maximum Queue Size

REFERENCES

- [1] Hendirickson, C. and Kocur, G. (1981), Schedule delay and departure time decisions in a deterministic model, Transportation Science 15 (1), 62-77.
- [2] Kemeny, J. G. and Snell, J. L. (1976), Finite Markov Chains, Springer-Verlag, 43-67.
- [3] Morales, J. M. (1986), Analytical procedure for estimating freeway traffic congestion, Public Roads 50 (2), 55-61.
- [4] Ozekici, S. (1987), Average waiting time in queue with scheduled batch service, Transportation Science 21 (1), 55-61.
- [5] Wohl, M. (1970), The Practicalities of Determining Marginal Delay Times and Costs, The Urban Institute, Washington, D.C.

APPENDIX 1.

DERIVATION OF THE FUNDAMENTAL MATRIX

From Eq (2),
$$N = (I - Q)^{-1}$$
,

$$(I - Q)N = I \tag{A.1}$$

Elements of the fundamental matrix, $N_{i,j}$, are derived from the simultaneous equations represented by (A.1):

$$(pl+qm)N_{1,j} - plN_{2,j} = \begin{cases} 1. & \text{for } j=1\\ 0, & \text{for } 2 < j < k \end{cases}$$
 (A.2)

$$-qmN_{i-1,j} + (pl + qm)N_{i,j} - plN_{i+1,j} =$$

$$= \begin{cases} 1, & \text{for } i = j, \ 2 \le i \le k - 1, \ 2 \le j \le k - 1 \\ 0, & \text{for } i \ne j, \ 2 \le i \le k - 1, \ 1 \le j \le k - 1 \end{cases}$$
(A.4)
(A.5)

$$(A.5)$$

$$-qmN_{k-1,j} + qmN_{k,j} = \begin{cases} 1, & \text{for } j = k \\ 0, & \text{for } 1 < j < k-1 \end{cases}$$
 (A.6)

From Eq (A.5) and (A.7), it can be easily shown that:

$$N_{k,j} = N_{k-1,j} = N_{k-2,j} = \dots = N_{j+1,j} = N_{j,j}.$$
 (A.8)

By introducing $N_{j+1,j} = N_{j,j}$ in Eq (A.4):

$$N_{j,j} = \frac{1}{qm} + N_{j-1,j}. \tag{A.9}$$

From Eq (A.5):

$$N_{j-1,j} = \frac{1}{qm} \left(\frac{pl}{qm} \right) + N_{j-2,j}$$
 (A.10)

$$N_{3,j} = \frac{1}{qm} \left(\frac{pl}{qm}\right)^{j-3} + N_{2,j} \tag{A.11}$$

$$N_{2,j} = \frac{1}{qm} \left(\frac{pl}{qm}\right)^{j-2} + N_{1,j}$$
 (A.12)

By introducing Eq (A.12) in Eq (A.3), $N_{i,j}$, is

$$N_{i,j} = \frac{1}{qm} \left(\frac{pl}{qm}\right)^{j-1} \tag{A.13}$$

From (A.13), (A.12), (A.11), ..., (A.10) and (A.())

$$N_{2,j} = \frac{1}{qm} \left\{ \left(\frac{pl}{qm} \right)^{j-2} + \left(\frac{pl}{qm} \right)^{j-1} \right\}$$
(A.14)

$$N_{3,j} = \frac{1}{qm} \left\{ \left(\frac{pl}{qm} \right)^{j-3} + \left(\frac{pl}{qm} \right)^{j-2} + \left(\frac{pl}{qm} \right)^{j-1} \right\}$$
 (A.15)

$$N_{j,j} = \frac{1}{qm} \sum_{k=0}^{j-1} \left(\frac{pl}{qm}\right)^k.$$
 (A.16)

Finally, from Eq (A.8), (A.13), (A.14), (A.15), ..., (A.16), we have

$$N_{i,j} = \begin{cases} \frac{1}{qm} \sum_{k=0}^{j-1} \left(\frac{pl}{qm}\right)^k, & \text{for } i \ge j \\ \frac{1}{qm} \sum_{k=j-i}^{j-1} \left(\frac{pl}{qm}\right)^k, & \text{for } i < j \end{cases}$$

$$(3)$$

APPENDIX 2.

DERIVATION OF AVERAGE QUEUE DISSIPATION TIME

In order to simplify the equation(7), we let $\left(\frac{pl}{qm}\right)$ be equal to ρ .

$$T_{i} = \frac{1}{qm} \left\{ i \sum_{k=0}^{n-i} \rho^{k} + \sum_{k=n-i+1}^{n-1} (n-k)\rho^{k} \right\}, \qquad 2 \leq i \leq n$$

$$= \frac{1}{qm} \left\{ i \frac{1-\rho^{n-i+1}}{1-\rho} + (i-1)\rho^{n-i+1} + (i-2)\rho^{n-i+2} + \dots + \rho^{n-1} \right\}$$

$$= \frac{1}{qm} \left\{ \frac{i(1-\rho^{n-i+1})}{1-\rho} + \frac{i\rho^{n-i+1}(1-\rho^{i-1})}{1-\rho} - \frac{\rho^{n-i+1}(1-i\rho^{i-1}+(i-1)\rho^{i})}{(1-\rho)^{2}} \right\}$$

$$= \frac{1}{qm} \cdot \frac{i(1-\rho) - (1-\rho^{i})\rho^{n-i+1}}{(1-\rho)^{2}}$$

$$= \frac{i(m-p) - qm \left\{ 1 - \left(\frac{pl}{qm}\right)^{i} \right\} \left(\frac{pl}{qm}\right)^{n-i+1}}{(m-p)^{2}}, \qquad 2 \leq i \leq n.$$

Since Eq (B.2) is the same as Eq (6), if i = 1, Eq (B.2) is valid for 1 < i < n.

APPENDIX 3.

DERIVATION OF MAXIMUM QUEUE SIZE $(h_{i,j})$

As in Appendix 2, let $\frac{pl}{qm}$ be equal to ρ .

$$N = \frac{1}{qm} \begin{bmatrix} 1 & \rho & \rho^2 & \dots \\ 1 & (1+\rho) & (\rho+\rho^2) & \dots \\ 1 & (1+\rho) & (1+\rho+\rho^2) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$
 (C.1)

$$N_{\rm dg} = \frac{1}{qm} \begin{bmatrix} 1 & 1+\rho & \\ & 1+\rho+\rho^2 & \\ & & \ddots \end{bmatrix}$$
 (C.2)

therefore

$$(N-I)N_{dg}^{-1} = \frac{1}{qm} \begin{bmatrix} 1 - qm & \frac{\rho}{1+\rho} & \frac{\rho^2}{1+\rho+\rho^2} & \dots \\ & \frac{1}{1+\rho-qm} & \frac{\rho+\rho^2}{1+\rho+\rho^2} & \dots \\ & & 1 & \frac{1+\rho-qm}{1+\rho+\rho^2} & \dots \\ & & & 1 & \frac{1+\rho+\rho^2-qm}{1+\rho+\rho^2} & \dots \end{bmatrix}$$
(C.3)

Then we have

$$h_{i,j} = \begin{cases} 1, & i > j \\ (1 - qm) \left[\sum_{k=0}^{j-1} \left(\frac{pl}{qm} \right)^k \right]^{-1}, & i = j \\ \left[\sum_{k=j-i}^{j-1} \left(\frac{pl}{qm} \right)^k \right] \cdot \left[\sum_{k=0}^{j-1} \left(\frac{pl}{qm} \right)^k \right]^{-1}, & i < j \end{cases}$$
(18)